Trade, Trees, and Health*

Xinming Du

Lei Li

Eric Zou

July 2024

Abstract

This paper shows a cascading mechanism through which international trade-induced deforestation results in a widespread decline of health outcomes in cities distant from where trade activities occur. We examine Brazil, which has ramped up agricultural export over the last two decades to meet rising global demand. Using a shift-share research design, we first show that export shocks cause substantial local agricultural expansion and a virtual one-for-one decline in forest cover. We then construct a dynamic area-of-effect model that predicts where atmospheric changes should be felt – due to loss of forests that would otherwise serve to filter out and absorb air pollutants as they travel – downwind of the deforestation areas. Leveraging quasi-random variation in these atmospheric connections, we establish a causal link between deforestation upstream and subsequent rises in air pollution and premature deaths downstream, with the mortality effects predominantly driven by cardiovascular and respiratory causes. Our estimates reveal a large telecoupled health externality of trade: over 700,000 premature deaths in Brazil over the past two decades. This equates to $0.18 loss in statistical life value per $1 agricultural exports over the study period.

Keywords: natural capital, trade, deforestation, air pollution, health

JEL Codes: F18, O13, Q23, Q53

* Du: Department of Economics, National University of Singapore (email: xdu@nus.edu.sg); Li: Department of Economics, University of Mannheim (email: lei.li@uni-mannheim.de); Zou: Ross School of Business, University of Michigan and NBER (email: ericzou@umich.edu). We thank Philipp Ager, David Autor, Ricardo Dahis, David Dorn, Rafael Dix-Carneido, Lisandra Flach, Bård Harstad, Karen Helene Ulltveit-Moeandre, Andreas Moxnes, Jan Schymik, Arthur Seibold, Joseph Shapiro, Camille Urvoy, Andreas Kotsadam, Ulrich Wagner, Kathrine von Graevenitz, and participants at various seminars for helpful comments. Raphaël Pérot and Zihao Chen provided excellent research assistance. All errors are our own.

1. Introduction

Globalization encompasses complex interactions between demand, production, and downstream environmental consequences that all occur in locations far apart.[1] This paper demonstrates the importance of such telecoupling for understanding the health burden of international trade. We study Brazil, where agricultural export value has quadrupled over the last two decades due to rising global demand. This surge is accompanied by substantial forest losses to accommodate the agricultural expansion. We show that these export-driven deforestations have far-reaching health ramifications, as the lost forests would otherwise serve to absorb harmful air pollutants as they travel, influencing air quality in cities hundreds or even thousands of miles downstream. We find that this trade-forest-health mechanism is considerable in significance when compared to the nominal value of Brazil's agricultural trade over the study period.

We are motivated by the long-standing debate on the trade-environment relationship: what are the sources of trade's environmental effects, and how to quantify them (Copeland, Shapiro, and Taylor, 2022). This is especially relevant in developing country context, where the reliance on exporting resource-intensive products plays a key role in achieving development goals (Frank and Schlenker, 2016). In current literature, the environmental costs of deforestation are predominantly associated with climate change and disruptions to ecosystems (Copeland, Shapiro, and Taylor, 2022; Balboni et al., 2023). These include projected alterations in global temperature and precipitation patterns (Bonan, 2008), biodiversity (Dasgupta, 2021), and the risk of tipping point (Franklin and Pindyck, 2018). Though expected to be profoundly important, these effects are hard to empirically pin down due to their very long-run nature. Our paper instead examines human health effects that can be directly measured in the data *today*. The main message of our work is that the environmental costs of trade-induced deforestation are considerable, even when considering immediate cardio-respiratory health impacts *alone*.

[1] Economists have extensively studied how trade enhances growth (Grossman and Helpman, 1990), productivity (Alcalá and Ciccone, 2004), and institutional quality (Levchenko, 2007). Empirical studies show international trade could affect the environment in positive or negative ways (e.g., Antweiler et al., 2001; Frankel and Rose, 2005; Managi et al., 2009; Abman and Lundberg, 2020), and emphasize the large spatial ranges of these impacts. Incorporating multiple locations and environmental consequences is important in assessing the overall welfare effects of globalization.

As a first step of our analysis, we estimate the causal effect of agricultural export on local forest land coverage. Over the two decades of our study period (1997-2019), Brazil has substantially expanded cattle ranching and soybean farming, coinciding with a one-for-one decline in forest land. This pattern of deforestation is commonly attributed to growing export demands, particularly from China and the European Union, driving farmers to employ slash-and-burn methods for rapid land clearing to boost agricultural production. We estimate the causal link between agricultural export and deforestation using a standard shift-share research design, which exploits heterogeneity in a region's exposure to export shocks. Our estimates show each 1,000 BRL increase in export per capita reduces forest cover in the area by 0.174 percentage points. Applying the same econometric framework, we show that areas of forest retreat are converted into agricultural use.[2]

Next, we estimate the impact of forest losses on downstream environmental and health outcomes. To do this successfully, we need to identify the possible extent of cross-boundary pollution spillover: the impact of a pollution source – or in our case, forests as a source that would have reduced pollution – does not obey administrative boundaries and extends far beyond the source (e.g., Heo, Ito, and Kotamarthi, 2023). This concern is particularly relevant in our study context when considering large-scale deforestation. We develop a simple area-of-effect (AoE) model, where we predict the geo-temporal influence of a given city's forest change by solving a dynamic model that tracks the evolution of airflow across space and time. From this model, we obtain information on the intensities of upwind-downwind linkages, or the lack thereof, of all pairs of cities in Brazil in all months. Our model is simplistic enough to be computationally feasible, but at the same time it is sufficiently realistic: we show that the model's output linearly predicts the correlation between upwind and downwind cities' observed pollution levels (which we did *not* use in training the model), and when our model predicts a lack of up-downwind link, there is indeed a lack of correlation between two cities' pollution levels. Put differently, our model appears to predict the potential extent of pollution transport very well.

Building on this data, we develop a quasi-experimental econometric equation to estimate the impact of forest losses on downwind environmental quality and health. For each sender city,

[2] We find a large increase in land use for soybean farming by 0.427 percentage points per 1,000 BRL export per capita. We found positive effects of increase in sugar cane farming, though the effect is smaller in magnitude and less statistically precise.

potential receiver city, and month, we model the relationship between the sender's forest quantity and the receiver's environmental and health outcomes, allowing this relationship to vary according to the intensity of the sender-receiver linkage, as suggested by the AoE model above. This estimation thus exploits plausibly exogenous variations in the upwind-downwind link over time and space, conditional on high-dimensional controls (such as source-destination pair by month-of-year fixed effects). Our findings suggest a significant effect of upwind forests on downwind air quality and mortality. The mortality effect is driven almost entirely by cardio-respiratory causes. A set of placebo tests – using plausibly irrelevant outcomes such as accident-related mortality or city pair-periods with an absence of upwind-downwind relationships – indicates no effect, lending support to the validity of the empirical design.[3]

The increase in cardio-respiratory mortality downwind results from a combination of two factors: the pollution generated by upwind forest burning and the diminished air-purification capabilities due to the loss of forests. Using remote-sensing data tracking fire incidents, we provide evidence that the latter channel is likely dominant. That is, deforestation causes a reduction in the forest system's ability to filter and absorb pollutants, which would have otherwise cleansed the air before it reached downwind areas.

Together, these estimates can be readily integrated to compute the overall impact of trade-induced forest losses. This calculation reveals that agricultural exports over the study period are responsible for over 3.6 million hectares of forest losses and 732,000 premature deaths. This is a substantial figure compared to the total number of deaths reported in the entire country in 2018, which is 1.28 million. Using a Value of Statistical Life (VSL) approach, we calculate that the life value loss associated with these excess deaths is around 513 billion USD. This is about 18 percent of the total agricultural value of Brazil over the entire study period – thus our headline number. This estimate is likely conservative, as our study focuses on monthly mortality responses, and the literature indicates that the long-term effects of pollution are usually more significant (Ebenstein et al., 2017). Moreover, individuals who do not die from pollution exposure may still experience other significant, albeit non-fatal, consequences, such as reduced productivity (Graff Zivin and Neidell, 2012; Borgschulte, Molitor, and Zou, 2022).

[3] The pollution-mortality elasticity implied from our estimation aligns with established estimates in the existing literature, further enhancing the plausibility of our conclusions.

Our work provides a quantitative basis to discuss the environmental damage of agricultural trade. Domestic environmental regulations are among the limited number of tools the government can use to mitigate the impacts of trade on natural capital depletion (Copeland, Shapiro, and Taylor, 2022). Such policymaking must be grounded in rigorous economic science that monetizes the value of natural assets, and our study describes a feasible framework to do just that.[4] In addition to providing an overall estimate of the export-health trade-off, our estimates may also be useful for targeting, e.g., identifying areas where the same amount of deforestation would lead to the largest downwind environmental health damage.[5] It is also noteworthy that our research is almost entirely done using publicly available data maintained by the Brazilian government and other public entities. This highlights the potential for our framework to be applied in different settings, given the availability of similar datasets in other countries or empirical scenarios. Our end-to-end analysis of the entire trade-forest-pollution-health chain also presents fewer concerns regarding external validity.

While the trade and environment literature has traditionally focused on the role of industrial emissions, there is a growing body of work that examines the environmental impact of trade through its interaction with natural resources. To the best of our knowledge, this paper provides among the first causal estimates on the impact of exports on deforestation. The most closely related work is Carreira, Coasta, Pessoa (2024), who study the impact of agricultural productivity changes and China shocks – and their interactions – on forest cover in Brazil. An earlier work such as Ferreira (2004) also documents a negative relationship between trade and forest cover, though mostly based on cross-country comparison without causal identification.

By further connecting to health outcomes, this paper adds to the literature on the environmental and health effects of forests. This literature is nascent in itself, but emerging applications have found important implications of forest presence and losses on weather, climate, agriculture, health, among other social economic outcomes such as property values (Jones and Goodkind, 2019; Druckenmiller, 2020; Han et al., 2021; Arujo, 2023; Grosset, Papp, and Taylor,

[4] See, for example, Recommendation 3 of *National Strategy to Develop Statistics for Environmental-Economic Decisions*, Office of Science and Technology Policy, Office of Management and Budget, Department of Commerce (2023). In addition to direct regulations, many alternative methods for fighting deforestation may also benefit from concrete estimates of health effects of forest losses, such as conservation contracting and ecosystem payment programs (Harstad, 2020; Jayachandran et al., 2017).

[5] We show this in Figure 5 of the paper. Depending on location and prevailing wind direction, the health costs of trade deforestation varies substantially across space.

2023; Li, 2023; Xing et al., 2023). The most closely related study is Xing et al. (2024), which studies the impact of an urban afforestation program in Beijing, China, on air quality and respiratory emergencies in the city. Our study is the first to identify a direct link between forests and mortality rates, adding a significant dimension to the costs associated with forest loss, and doing so across a wide, inter-city geographical scale.

Our paper also contributes to the literature on trade and mortality by introducing a new channel—deforestation—and quantifying its impact. Existing studies linking trade and mortality focus on two channels: employment and environmental damage. Regarding the former, trade-induced job loss has been shown to increase drug overdoses (Pierce and Schott, 2020), reduce access to employment insurance (Guerrico, 2021), and worsen workers' physical and mental health (Adda and Fawaz, 2020). On the latter, Bombardini and Li (2020) show that Chinese cities specializing in dirty industries experience both pollution and income effects, leading to increases in mortality. In contrast, cities specializing in clean industries only enjoy income effects and experience decreases in mortality. The authors use sectoral emission intensity to classify industries as dirty or clean among manufacturing producers. Gong et al. (2023) examine how foreign demand shocks due to the global economic crisis reduce demand for Chinese products, drive local changes in $PM_{2.5}$ concentrations, and affect downstream mortalities. Unlike their focus on manufacturing producers and industrial emissions, our paper focuses on agricultural trade. Our channel of export-induced natural resource depletion has not yet been studied.

Section 2 continues with institutional knowledge on trade, deforestation, the environment, and human health. Section 3 describes the overall research framework and data sources. Section 4 presents the causal estimation of the effect of export on forest coverage. Section 5 presents the health analysis. Section 6 presents the cost calculation and concludes the paper.

2. Institutional Background

2.1 Agriculture and Deforestation in Brazil

In many developing countries, farmers are driven to deforest land as a means to increase their agricultural output and income. The prevalent method of land conversion involves slash-and-burn techniques, where trees are cut down and the remaining vegetation is burned, enriching

the soil with nutrients from the ash. The opening of new lands for agriculture is frequently facilitated by government subsidies or lax enforcement of environmental regulations, further incentivizing farmers to engage in these practices.

The interplay between agricultural expansion and deforestation is particularly evident in the Brazilian Amazon rainforest, where the cultivation of soybeans and the expansion of cattle ranching are the primary agricultural activities driving deforestation. Soybeans, a major export crop, have seen increased demand globally, particularly from China and the European Union, for use as livestock feed and in various food products. This demand has propelled Brazilian farmers to clear vast tracts of the Amazon to cultivate this lucrative crop. Similarly, cattle ranching, which requires extensive land area for grazing, has expanded dramatically, making Brazil one of the world's largest beef exporters. The profitability of these agricultural ventures, combined with governmental policies that often prioritize economic growth over environmental protection, has led to significant deforestation in the region. This deforestation for agriculture not only has dire consequences for biodiversity and indigenous communities but also impacts global climate change, as the Amazon plays a critical role in carbon sequestration and climate regulation.

2.2 Forests and the Environment

Forests represent a crucial component of natural capital, contributing to the ecosystem and the environment. They function as indispensable habitats for 80% of the global terrestrial species (Aerts and Honnay, 2011), so the health of forest is naturally tied to biodiversity. Besides, forests contribute to water, oxygen, and carbon cycles, and other biogeochemical balances. They also play a pivotal role in the global and local climates. Forests are the world's largest storehouses of carbon. Trees absorb greenhouse gasses, and are primary instruments for carbon storage featured in carbon offset programs like the REDD+. At the local level, forests affect micro-climate via shading and evaporation.

Forests matter a lot for human well-being. Over 1.6 billion people depend on forests for food, fuel, or shelter, and 70 million people including indigenous communities rely on forests for settlement (UN, 2021). Forests also mitigate damages from pollution and natural disasters such as floods. In this paper, we focus on air pollution impacts of forests as short-term consequences of trade-induced deforestation. Unlike other effects, air pollution responds quickly to tree cover

change and could be linked to export changes of interest. Air pollution also leads to immediate mortality responses, enabling us to quantify health externality.

Why do forests mitigate air pollution? First, leaves and bark serve as filters for particles, providing deposition sites for airborne pollutants. Previous studies have documented that forests are efficient pollution collectors and improve air quality (e.g., Betts et al., 2008; Janhall, 2015). Second, leaves can absorb specific hazardous gases like SO_2, NOx, and O_3. They can also interact with pollutants and alter atmospheric compositions. Leaves are reported to biodegrade or transform pollutants into less or nontoxic molecules (Wei et al., 2017). Third, trees indirectly affect the formation and deposition of air pollution by altering the micro-climate. Ozone formation is inhibited by lower temperatures (Coates et al., 2016), and increased rainfall decreases the lifespan of particles.

The link between deforestation, air pollution, and health can span a wide spatial range due to the long-distance transport of pollutants through wind flows. Winds carry upwind pollution into downwind regions, thus wind speed, direction, and other climatic factors determine the residence time and location of pollution. Existing papers have shown the impacts of cross-regional or transboundary air pollution on downwind areas (e.g., Heo, Ito, and Kotamarthi, 2023). Focusing on trees, policymakers have implemented afforestation initiatives to protect downwind populations from air pollution. For example, in China, the Three-North Shelter Forest is designed to restrain sandstorms from the Gobi Desert and improve air quality near the capital region.

3. Research Framework and Data

3.1 Research Framework

Before proceeding, it is worth laying out how various parts of our empirical estimation are eventually going to line up. Our goal is to estimate the external health effects of trade-induced deforestation, recognizing that this effect need not confine locally to where deforestation occurs. The empirical analyses can be thought of as being organized around the following conceptual equation:

$$\text{Health Effects of i} = \sum_r \underbrace{\frac{\partial \text{Forest}_i}{\partial \text{Trade}_i}}_{\text{Section 4}} \cdot \Big(\underbrace{\frac{\partial \text{Health}_r}{\partial \text{Forest}_i} \Big| \overbrace{\text{Wind}_{i \to r}}^{\text{Section 5.1}}}_{\text{Section 5.2}} \Big) \quad (1)$$

Section 4 estimates the causal effect of trade on deforestation. This exercise is entirely local in nature: we want to learn how much of city i's deforestation is due to trade shocks to *that* city. Section 5 estimates the spillover health effects of city i's deforestation on all potential "receiver" cities indexed by r. Section 5.1 first builds an algorithm that identifies these potential receivers. We use high-frequency wind direction information to create a downwind exposure index $\text{Wind}_{i \to r}$ which summarizes the degree to which city i may have an influence on city r's environmental condition due to atmospheric motions. Section 5.2 then builds on this information to estimate the relationship between city i's forest cover and receiver city r's health outcomes – which we show is highly dependent on, and nonlinear with respect to, the downwind exposure index. The total effect of city i's is thus the summation across its effects on various downwind receiver cities.

In Section 6, we will come back to an empirical version of equation (1) which will produce the headline, cost-of-trade-induced-deforestation number.

3.2 Data

Our analysis is done almost entirely using publicly available datasets, many of which are maintained by the Brazilian government. We briefly describe each source.

Exports. Our export data is from the Comex database provided by Brazil's Foreign Trade Secretariat (Secretaria de Comércio Exterior - Comex). The original data is at the municipality by month level and is structured using the Harmonized System at the 4-digit level ("HS-4"), providing detailed categorization of exported goods.

Land Use. Our land use data is from MapBiomas. Based on the Landsat satellite images produced by NASA and USGS, MapBiomas processes the pixel-per-pixel data with machine learning algorithms through the Google Earth Engine platform. It produces annual land use and land cover maps and provides information on the dynamics of different types of land cover for Brazil from 1985 to 2022 on a 30m spatial resolution. Specifically, there are six types of land, namely forest, non-forest natural formation, farming, non-vegetated area, water, and not

observed. We mainly analyze the data on forest land and farming land. There are four types of forest lands, namely forest formation, savanna formation, mangrove, and sandy coastal plain vegetation. There are four types of farming land, namely pasture, agriculture, forest plantation, and mosaic of uses. For agricultural land, the MapBiomas offers information on major crops, such as soybean, sugar cane, rice, cotton, coffee, and citrus. The data used in our paper is the 7.0 version obtained in 2022.

Meteorology. We draw daily temperature and precipitation data from BDMEP (Banco de Dados meteorológicos para Ensino e Pesquisa). During the majority of our study period (2000-2015), BDMEP provided data from 150 weather monitoring stations. The network expanded to include 5,611 stations for the years 2016-2020. For each station-day, we observe daily minimum and maximum temperatures (°C) – from which we calculate average temperature by taking the average of the two – and precipitation (mm). We then compute station by month average temperature and precipitation. To derive city-level weather variables, we match all weather stations within a 300-kilometer radius of each city's centroid and calculate the average temperature and precipitation across these matched stations.

We use wind direction and speed data from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5) reanalysis product. This dataset provides hourly U (horizontal) and V (vertical) components of the wind vectors at a 0.25° grid resolution. We use this data to compute daily prevailing wind directions, which are used in the air flow analysis in Section 5.

Air Quality. Air quality data is sourced from IEMA (the Instituto de Energia e Meio Ambiente). It collects air quality monitoring data from 380 stations in 28 cities from 2015-2022. We observe hourly concentration of six pollutants: $PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO. We aggregate the data to the city-month level.

Mortality. We obtain mortality microdata from the Mortality Information System (Sistema de Informações sobre Mortalidade, or SIM) via the Brazilian National Health System Information Technology Department (Departamento de Informática do Sistema Único de Saúde, DATASUS). The SIM data are derived from administrative death records, which are completed and issued by physicians, and collected from hospitals and medical examiner offices.

From the death record microdata, we observe the decedent's municipality of residence, date of death, and the underlying cause of death coded in ICD-10 classification. Using these information, we aggregate the microdata to obtain death counts at the microregion-by-month level between 2000-2021.[6] We categorize deaths into those due to cardiovascular or respiratory causes ("cardiovascular", ICD10: I00-I99, J00-J99) and all other, non-cardiovascular causes. Separately, we count deaths due to accidents ("external", ICD10: S00-T98, V01/V989), which we will use as a "placebo" outcome in our econometric analysis. Finally, we convert death counts to death rate by dividing microregion level population counts from the Brazilian Institute of Geography and Statistics.

Geography. Because municipality boundaries changed over time over our study period, we aggregate geographic units to the 3,826 Minimum Comparable Areas (henceforth "area") level for temporal consistency (Ehrl, 2017). We drop five municipalities that cannot be identified in the municipality-to-AMC crosswalk.[7] Our primary trade-deforestation analysis is based on AMC-level data. Our environmental data have a coarser resolution than AMCs (e.g., 0.25 degree for wind, 380 air quality stations.) In airflow modeling and the subsequent environmental analysis, we aggregate AMCs into 557 "microregions" as defined by the Brazilian Institute of Geography and Statistics (IBGE) to make computation feasible.

4. Trade, Agricultural Expansion, and Deforestation

4.1 Raw Statistics

We first present a set of summary statistics on the shifts in Brazil's trade structure, agricultural expansion, and deforestation.

Figure 1, panel (a) plots Brazil's national total export value, which grew from about 50 billion USD in 1997 to over 200 billion USD in 2018. This remarkable growth is driven primarily by exports of agricultural/food products – the focus of this paper – and mineral/metal products. In panel (b), we plot changes in the composition of agricultural/food product exports in each year. The graph shows a sharp growth in soybean exports, growing from about 16 percent of total

[6] We drop data between years 1997 and 1999 as causes of death information are missing.
[7] Mojuí dos Campos, Pescaria Brava, Balneário Rincão, Pinito Bandeira, and Paraíso das Águas.

agricultural exports in 1997 to more than 40 percent by 2018. In the Online Appendix, we show this increase is mainly driven by rising demand from China and the EU.

Figure 2, panel (a) shows the geographic distribution of the agricultural export growth. Not surprisingly, growth hotspots concentrate in areas with high crop production. Panel (b) reports corroborative evidence of remarkable agricultural employment growth with a similar geographic pattern.[8] Figure 2, panel (c) shows the declining trends in Brazil's forest cover over the past decades, accompanied by a virtual one-for-one increases in farmland. The map on the right side of the panel highlights areas of deforestation which, once again, shows a strong spatial correspondence with regions experiencing significant agricultural export growth.

We next describe our econometric method to quantify the causal effect of export shocks on deforestation.

4.2 Research Design

The main estimation equation is:

$$\Delta\text{Land Cover}_{i,y} = \beta \cdot \Delta\text{Export}_{i,y} + \lambda \cdot W_{i,y} + \varepsilon_{i,y} \quad (2)$$

where $\Delta\text{Land Cover}_{i,y}$ is the three-year rate of change in a certain type of land cover at area i and year y.[9] For example, when the outcome of interest is deforestation, the land cover variable is defined as $\Delta\text{Forest}_{i,y} = \left(\text{Forest}_{i,y+3} - \text{Forest}_{i,y}\right)/\text{Land}_{i,y}$ where "Forest" denotes forest cover in acreage in the area-year, and "Land" is total area. The key right hand side variable $\Delta\text{Export}_{i,y}$ is the three-year change in agricultural exports per capita. $W_{i,y}$ is a series of control variables, including macro-region fixed effects, year fixed effects, and region-time control variables including income per capita, literacy rate, rural population, urban population, population density, and the presence of genetically-modified soy.[10] The coefficient of interest β therefore captures the

[8] We present maps of mining product exports – the other main sector that saw substantial growth during the study period – in Online Appendix Figure A.4. The geographic pattern is markedly different from that of agricultural exports.

[9] In the appendix, we show that results are robust with one-year and six-year first differences.

[10] We follow Carreira, Costa, and Pessoa (2024) in selecting these control variables. Our estimates are insensitive to the exclusion of these controls.

impact of agricultural export growth on land cover changes. We cluster standard errors at the AMC level. All regressions are weighted with each area's agricultural population in 2000.

Simple OLS estimation of equation (2) may suffer from endogeneity and measurement errors. We build the following shift-share instrumental variable (IV) for the export variable:

$$IV_{i,y} = \left[\sum_j \frac{Export_{i,j,t-k}}{Export_{i,t-k}} \cdot \frac{Export_{j,t+k}^{-i} - Export_{j,t}^{-i}}{\left(Export_{j,t+k}^{-i} + Export_{j,t}^{-i}\right)/2}\right] \cdot \frac{Export_{i,t-k}}{Pop_{i,t-k}} \quad (3)$$

where $Export_{i,j,t-k}/Export_{i,t-k}$ is the share of product j in area i's exports at period $t-k$, which captures an area's reliance on a certain type of agricultural product. The term $Export_{j,t}^{-i}$ is national exports in product j and year t excluding the state in which area i is located. This variable is used as the excluded instrument for the endogenous $\Delta Export_{i,y}$ variable in a two-stage least squares estimation of equation (2). $\frac{Export_{i,t-k}}{Pop_{i,t-k}}$ is initial export per capita of area i in the year $t-k$. The approach in equation (3) predicts an area's agricultural export growth per capita by combining each area's initial agricultural export structure with the product-level national agricultural export growth. It captures the exogenous component of the regional export growth by characterizing the initial exposure of an area to national export growth, which exposes the area to national-level shocks more in some types of goods than in others. Because the export values of many products are zero in early years, it is not possible to calculate export growth for many areas. To address this issue, we construct the modified version of export growth following Davis and Haltiwanger (1992), where the denominator is the average of the start-of-period value and end-of-period value. In our primary specification, we select a lagged period of $k = 4$. In the Online Appendix, we should our results are robust to alternative windows.

The nascent literature on the corrections of the shift-share instrument highlights two alternative frameworks in which the exclusion restriction of shift-share instruments may hold, one based on quasi-randomness of national shocks (Borusyak, Hull, and Jaravel, 2022) and the other on quasi-randomness of local shares (Goldsmith-Pinkham et al., 2020). In the context of this paper, we use the leave-one-out strategy to calculate the national export growth by excluding the agricultural export in the state where the area locates, which excludes the interference of local demand. Therefore, the framework in Borusyak, Hull, and Jaravel (2022) suits the setting of this paper better. To show that the key assumptions in Borusyak, Hull, and Jaravel (2022) are satisfied, we conduct the following tests. First, regarding the key assumption on quasi-random shock

assignment, we show in Figure A5 that agricultural products within the same HS-4-digit product tend to occupy a wide range on the vertical axis, indicating highly varying changes in export growth. Second, the number of products is sufficiently high. Third, regarding the key assumption on product concentration, we calculate the export shares across HS-4 product categories in four different years: 1999, 2004, 2009, and 2014. We show that the agricultural exports are not concentrated in a few products.

4.3 Results

Figure 2 reports the primary estimates. Each node on the tree represents a separate regression. All regressions follow the exact same IV specification as laid out in equations (2)-(3), except for the outcome variables which are denoted by the name of the node. Branches of the tree represent hierarchies of the land use variables defined in the MapBiomas dataset. For example, the "forest" land use category consists of four subcategories: "forest formation", "savanna formation", "mangrove", and "sandy". The numbers on the four sub-nodes thus show the causal impacts of export on each of the four subcategories, which (mechanically) add up to the causal impact on the main, forest node.

The headline "export-deforestation" estimate of this paper is shown on the "ΔForest" node of the figure: each 1,000 BRL increase in export per capita reduces forest cover by 0.174 percentage points. This estimate is statistically significant at the 5 percent level, and the effect size converts to about 5.7 percent of a standard deviations (S.D.) reduction in forest cover per 1 SD change in export. The sub-nodes show that this total deforestation effect is driven mainly by decrease in forest and grass (savanna) formation.

Now come to the lower parts of the graph. We show that export leads to significant increases in farmland coverage by about 0.214 percentage points. This effect is diluted by the fact that, in the land use data, "forest plantation" is counted toward farming activities, which – consistent with our deforestation results above – is *negatively* affected by exports. Consistent with institutional knowledge, the main farming effect is driven by increases in agriculture by 0.334 percentage points.

We can further look at which crop types drive the agriculture effects. The right-hand-side parts of the graph show that export's effect of agriculture is driven primarily by an increase in

soybean farming. This is once again consistent with the well-known fact that soybean demand from China and the EU has substantially increased over the past decades.

5. Downwind Health Effects of Forests

5.1 Air Flow Modeling

To capture downwind environmental and health effects, we build a matrix that summarizes monthly wind flow intensities between all pairs of cities in Brazil. In the interest of space, here we provide an intuitive explanation of the procedure, leaving computational details to the Online Appendix.

The input data are 0.25-degree resolution by hourly wind direction and speed information (i.e., vectors) from the ERA5 product. We first generate a spatial representation of wind *flows* from individual wind *vectors*, as illustrated in Figure 3, panel (a). Beginning from a particular day and city, we construct streamlines by sequentially following the wind's speed and direction on a daily basis. This process maps out the evolving trajectories of the wind field, giving us daily representations of the distribution of wind flows throughout the country.

Our goal is to build an index, $\text{Wind}_{i\to r,d}$, which is a summary of downwind intensity blowing from a sender city i to a receiver city r on day d. Note that, for any pair of cities that are significantly distant from each other, it is meaningless to talk about up/downwind relationship on any given "day" because winds blowing from the sender city may take days to arrive at the receiver city. We therefore track the trajectory of winds "originating" from a sender city and their impacts of downwind cities over multiple days – or "steps" as we refer to them below – using the wind streamline data that we constructed earlier.

We define downwind intensity score using the following formula:

$$\text{Wind}_{i\to r,d,t} = \exp\{-\alpha \cdot \text{rad}_t - \beta \cdot |\theta|_{i\to r,d,t} - \gamma \cdot \text{dist}_{i\to r,d,t}\} \quad (4)$$

Starting from sender city i on day d and at step t, we assume downwind intensity of receiver city r follows an exponential decay as a function of three components (U.S. EPA, 2018; Phillips et al., 2021). The first component is the search radius at the step (rad_t), which captures general decreases of downwind intensity over steps. The initial radius is 300 km, and we increase

the search radius by 20 km at each step to capture both the uncertainty in the streamline computation and the dispersion of air. The second component is the scalar product of the angle between the receiver city and the wind direction originating from the sender's location ($|\theta|_{i\to r,d,t}$), which means we assign higher intensity to receiver cities that sit closer to the exactly-downwind direction of the sender. The third component is simply the distance between the sender and the receiver city ($dist_{i\to r,d,t}$), which captures geographic decay. We assume $Wind_{i\to r,d,t}$ to be zero if $d_{i\to r,d,t} > rad_t$ (i.e., if receiver city lies outside of the search radius at step t) or if $\theta_{i\to r,d,t} > 0.4\pi$ (i.e., if the receiver city is not obviously in the downwind direction from the sender city.[11] Starting from each particular sender city and day, we iterate the procedure for seven steps (i.e., a week).

Perhaps an easier way to think about the procedure above is that, we essentially "move" the location of the sender city with the wind flow and see which receiver cities it touches along the path over the next week period. A visualization of the procedure is shown in Figure 3, panel (b). The red arrow at the center represents the locus of the wind flow starting from the sender city. The growing ball of uncertainty around the arrow shows expanding search radius rad over steps in equation (4). The visualization also explains why both the relative angle and distance variables (θ and $dist$) have starting day and step subscripts (d and t): we track where winds originate and where they move to, and we compute each receiver city's relative angle and distances dynamically.

We aggregate step-wise downwind intensity scores to the day level:

$$Wind_{i\to r,d} = \sum_t Wind_{i\to r,d,t} \quad (5)$$

and in the econometric analysis below, we further average this city pair-daily score to the monthly frequency to make the size of the regression dataset manageable. Because the exact magnitude of these scores are not interpretable, in econometric analysis we use a nonparametric approach, grouping this variable into decile bins for positive values plus a separate bin for when downwind score is zero (i.e., because the receiver city is too far away or it is not in the downwind direction, as discussed earlier).

[11] We use parameter values $\{\alpha, \beta, \gamma\} = \{0.7, 0.5, 0.2\}$. These numbers are empirically determined such that we would obtain a spatially continuous flow coefficient function through the successive steps, and that the directionality of the observed winds are respected through the flow coefficient representation. See Online Appendix for more details.

5.2 Econometric Specification

Our goal is to estimate the relationship between a sender city's quantity of forests and its downwind receiver's environmental and health outcomes. Importantly, we want to leverage month-to-month fluctuations in the downwind score (as illustrated in panel b of Figure 3). Our focal estimation equation is as follows:

$$Y_{r,m,y} = \beta \cdot Forest_{i,y} \times Wind_{i \to r,m,y} + Forest_{i,y} + Wind_{i \to r,m,y} + \alpha_{i,r,m} + \alpha_y + \varepsilon_{i,r,m,y} \quad (6)$$

A unique observation is indexed by sender city (i), receiver city (r), month (m), and year (y). The data is all sender-receiver pairs expanded to year-month balanced panel. The three key variables in this estimation equation are: $Y_{i,m,y}$, which is the receiver city's outcome; $Forest_{i,y}$, which is the sender city's forest acreage (standardized to mean zero and standard deviation one), varying at the yearly rather than monthly level; and $Wind_{i \to r,m,y}$, which is the downwind score reflecting strength of wind blowing from i to r. In practice, we cut the $Wind_{i \to r,m,y}$ variable into 11 groups – 10 decile bins for non-zero downwind scores and another bin for zero wind scores – and include $Wind_{i \to r,m,y}$ as a categorical variable. So conceptually, the β coefficients – one for each of the 11 downwind score bins – capture how the relationship between sender's forest and receiver's outcome vary by downwind intensity.

We include sender-by-receiver-by-month fixed effects ($\alpha_{i,r,m}$), which isolate downwind variation arising within a city pair *and* month of the year to address potential cross-sectional and seasonal confounds. For example, consider the city pair of Sao Paulo and Brasilia. Our specification compares data from a particular January when the wind blowing from Sao Paulo to Brasilia is notably strong with data from other Januaries in different years when the wind strength for the same city pair is comparatively weaker. This nuanced approach enables us to parse out cross-sectional correlations (e.g., the fact that downwind relationships tend to be generally stronger for city pairs where the two cities are in close proximity) and pair-specific seasonality of wind patterns. We also include year fixed effects (α_y) to account for common shocks at the year level. In all regressions, we two-way cluster standard errors at the sender and receiver levels.

5.3 Results

Figure 4 plots the β coefficients from equation (6). Each chart represents results for a separate outcome. Start with panel (a), which shows atmospheric outcomes. To facilitate interpretation and comparison of results, we use standardized outcomes (mean 0, standard deviation 1) in each regression, so that the interpretation of the β coefficients is "the number of standard deviation change in receiver city's atmospheric outcome per 1 standard deviation increase in sender city's forest coverage."

The first two charts of panel (a) suggest there is little/null effect of upwind forests on downwind temperatures and precipitation. The third chart shows, in contrast, forests have a large effect on downwind air pollutants reduction.[12] Two notable patterns are worth pointing out. First, the forest-pollution effect increases monotonically with the downwind intensity score. In the strongest wind decile bin (bin "1st" on the x-axis), a standard deviation increase in upwind forests reduces downwind pollution concentrations by 3 standard deviations.[13] Second, we find precisely estimated zero effect for the no-wind bin (bin "calm" on the x-axis). This serves as a reassuring "placebo" exercise, suggesting that our econometric specification indicates an absence of forests' downstream environmental effect when our wind analysis indeed suggests an absence of a significant downwind relationship between the city pair.

The empirical impacts of forests on the atmospheric conditions is a subject of ongoing research. Here we compare our findings to some existing studies. On temperature effects, most empirical studies we are aware of document the *local* cooling effects of forests through shading and evapotranspiration (e.g., Han et al., 2021), and so it is not surprising we are not finding a transported effect at large-distance scale. On precipitation effects, our null effects seem at odds with recent studies that show positive impacts of forests on downwind rainfalls (Grosset, Papp, and Taylor, 2023). One important difference of our estimation and Grosset, Papp, and Taylor (2023) is the time scale: while we look at monthly frequency impacts, their findings pertain mostly to longer-term changes. In the Online Appendix, we show that, when aggregating our analysis to

[12] Here we use the average z-score across six pollutants reported in the Brazilian data. Individual pollutant regression results are shown in the Online Appendix.
[13] We will discuss plausibility of this effect size in conjunction with the health effects estimates.

the annual level, we do see a substantial positive impact of forests on downwind precipitation.[14] Our pollution results echo Xing et al. (2023) which uses a similar design – though looking at within-city scale – to examine the impact of urban forests on city air quality improvement.

We report health effect estimates in panel (b). Given the findings of forests' impact on air quality, we separately examine mortality due to cardiovascular/respiratory causes – which are expected to be more responsive to pollution changes – and other, non-cardiorespiratory causes. The first chart of panel (b) shows that upwind forests lead to significant reduction in downwind cardiorespiratory mortality rates. Once again, as we saw in the pollution analysis results, the mortality effects are monotonic with respect to downwind intensity score, and we find a precise-zero effect when there are no significant downwind connections between the city pair. The second chart reports that we find little impact of upwind forests on mortality of non-cardiorespiratory causes. Finally, as an additional "placebo" exercise, in the third chart we look specifically at mortality due to external causes ("accidents") and we see no effect.

We can now interpret the effect sizes, and compare them with established estimates in the literature. In particular, there is an abundance of estimates in the economics and medical literature on the causal effects of particulate matter pollution (PM) on mortality. We can therefore check if the ratio between our forest-mortality estimate and our forest-pollution estimate is in line with the effect magnitude from the literature. Our pollution regression in panel (a) suggests that, with the strongest downwind score, each standard deviation increase in forests reduces pollutants by 3 standard deviations, which converts to about 65 ug/m3 reduction in PM_{10}.[15] The mortality regression in panel (b) suggests the corresponding change in cardiorespiratory mortality rate is about -0.37 deaths per 100,000 people. These numbers imply about a PM_{10}-mortality elasticity of 0.99%. In comparison, in one of the most cited quasi-experimental studies in the U.S., the estimated mortality-$PM_{2.5}$ elasticity among the elderly population aged 65 and above is 1.86% (Deryugina et al., 2019).

[14] While forests can contribute to increased local humidity through evapotranspiration, this increased moisture might not be sufficient to alter weather patterns significantly in the downwind area, especially over short distances or periods. Over the long term, extensive forests can modify regional climate patterns. As forests mature, they contribute to higher levels of evapotranspiration. This increased moisture in the air can eventually influence weather patterns, potentially leading to more rainfall in downwind areas.

[15] In the Online Appendix, we also report pollutant-specific regressions where the outcome variable of the analysis is a particular type of air pollutant concentration.

6. Conclusion

We are now ready to draw conclusions by plugging in various pieces of our empirical estimates back into the conceptual equation (1) we laid out in Section 3. The empirical version of the equation is as follows:

$$\text{Excess Deaths due to i} = \sum_{r,m,y} \left(\beta_{\Delta\text{Trade}\rightarrow\Delta\text{Forest}} \cdot \Delta\text{Trade}_i\right) \cdot \left(-\beta^{i\rightarrow r,m,y}_{\text{Forest}\rightarrow\text{CR mortality}} \cdot \text{Population}_{r,y}\right) \quad (7)$$

That is, the total excess deaths due to city i equals the product of trade-induced deforestation in i (the first parenthesis) and excess deaths per unit of deforestation at a receiver city r (the second parenthesis), summed across all receiver cities and time periods. The subscript of the term $\beta^{i\rightarrow r,m,y}_{\text{Forest}\rightarrow\text{CR mortality}}$ means we only consider the impact of forest on downwind cardiovascular mortality – what we found to be responding to upwind forest changes; the superscript indicates we use nonlinear estimates as shown in Figure 4, panel (b) based on the downwind index associated with the city pair $i \rightarrow r$, month m, and year y.

One qualitative conclusion which came directly out of this calculation is that, where trade shocks (and deforestation) happen are very different from where the mortality burden is borne. Figure 5, panel (a) shows a map of trade-induced deforestation based on our estimates. In panel (b), we plot the distribution of total mortality burdens at receiver cities. That is, for each city, we sum up excess deaths due to deforestation from all upwind cities over the study period. The difference of the spatial patterns between the two panels is apparent: because of wind directivity and population distribution, areas with large numbers of excess deaths may situate hundreds if not thousands of miles away from where trade deforestation is happening.

Quantitatively, our estimates suggest a total of 3.6 million ha loss of Brazilian forests due to trade, resulting in over 732,000 excess deaths over the study period. For reference, total deaths in Brazil in the year of 2018 is estimated to be around 1.28 million.

We quantify the mortality impacts in terms of its statistical life value, which is a common way to think about the costs (values) of environmental damage (protection) (e.g., U.S. EPA, 2000). Because direct estimation of the value of a statistical life (VSL) for Brazil is lacking, we use a transfer approach that downscales the U.S.-based VSL estimate of $2.3 million (Ashenfelter and

Greenstone, 2004) by factors of a transfer elasticity of 1.2 (Narain and Sall, 2016) times a Brazil-U.S. income per capita ratio of 7. This yields a VSL estimate of $0.7 million USD. The total statistical life value loss of the 732,000 extra deaths thus amounts to about 513 billion USD. This represents a significant number – about 18 percent of the total agricultural export value of Brazil over the study period.

References

Abman, Ryan, and Clark Lundberg. "Does free trade increase deforestation? The effects of regional trade agreements." Journal of the Association of Environmental and Resource Economists 7.1 (2020): 35-72.

Adda, Jérôme, and Yarine Fawaz. "The health toll of import competition." The Economic Journal 130, no. 630 (2020): 1501-1540.

Aerts, Raf, and Olivier Honnay. "Forest restoration, biodiversity and ecosystem functioning." BMC ecology 11 (2011): 1-10.

Alcalá, Francisco, and Antonio Ciccone. "Trade and productivity." The Quarterly Journal of Economics 119, no. 2 (2004): 613-646.

Antweiler, Werner, Brian R. Copeland, and M. Scott Taylor. "Is free trade good for the environment?." American Economic Review 91, no. 4 (2001): 877-908.

Ashenfelter, Orley, and Michael Greenstone. "Using mandated speed limits to measure the value of a statistical life." Journal of Political Economy 112, no. S1 (2004): S226-S267.

Balboni, Clare, Aaron Berman, Robin Burgess, and Benjamin A. Olken. "The economics of tropical deforestation." Annual Review of Economics 15 (2023): 723-754.

Betts, Richard, Michael Sanderson, and Stephanie Woodward. "Effects of large-scale Amazon forest degradation on climate and air quality through fluxes of carbon dioxide, water, energy, mineral dust and isoprene." Philosophical Transactions of the Royal Society B: Biological Sciences 363.1498 (2008): 1873-1880.

Bombardini, Matilde, and Bingjing Li. "Trade, pollution and mortality in China." Journal of International Economics 125 (2020): 103321.

Bonan, Gordon B. "Forests and climate change: forcings, feedbacks, and the climate benefits of forests." Science 320, no. 5882 (2008): 1444-1449.

Borgschulte, Mark, David Molitor, and Eric Yongchen Zou. "Air pollution and the labor market: Evidence from wildfire smoke." Review of Economics and Statistics (2022): 1-46.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. "Quasi-experimental shift-share research designs." The Review of Economic Studies 89, no. 1 (2022): 181-213.

Carreira, Igor, Francisco Costa, and Joao Paulo Pessoa. "The deforestation effects of trade and agricultural productivity in Brazil." Journal of Development Economics 167 (2024): 103217.

Coates, Jane, Kathleen A. Mar, Narendra Ojha, and Tim M. Butler. "The influence of temperature on ozone production under varying NOx conditions – a modelling study." Atmospheric Chemistry and Physics 16, no. 18 (2016): 11601-11615.

Copeland, Brian R., Joseph S. Shapiro, and M. Scott Taylor. Globalization and the Environment. No. w28797. National Bureau of Economic Research, 2021.

Dasgupta, Partha. The economics of biodiversity: the Dasgupta review. Hm Treasury, 2021.

Davis, Steven J., and John Haltiwanger. "Gross job creation, gross job destruction, and employment reallocation." The Quarterly Journal of Economics 107, no. 3 (1992): 819-863.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. "The mortality and medical costs of air pollution: Evidence from changes in wind direction." American Economic Review 109, no. 12 (2019): 4178-4219.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." Proceedings of the National Academy of Sciences 114, no. 39 (2017): 10384-10389.

Ehrl, Philipp., 2017. Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities. Estudos Econômicos (São Paulo), 47, pp.215-229.

Ferreira, Susana. "Deforestation, property rights, and international trade." Land Economics 80, no. 2 (2004): 174-193.

Frank, Eyal G., and Wolfram Schlenker. "Balancing economic and ecological goals." Science 353, no. 6300 (2016): 651-652.

Frankel, Jeffrey A., and Andrew K. Rose. "Is trade good or bad for the environment? Sorting out the causality." Review of Economics and Statistics 87.1 (2005): 85-91.

Franklin Jr, Sergio L., and Robert S. Pindyck. "Tropical forests, tipping points, and the social cost of deforestation." Ecological Economics 153 (2018): 161-171.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. "Bartik instruments: What, when, why, and how." American Economic Review 110, no. 8 (2020): 2586-2624.

Gong, Yazhen, Shanjun Li, Nicholas J. Sanders, and Guang Shi. "The mortality impact of fine particulate matter in China: Evidence from trade shocks." Journal of Environmental Economics and Management 117 (2023): 102759.

Graff-Zivin, Joshua, and Matthew Neidell. "The impact of pollution on worker productivity." American Economic Review 102, no. 7 (2012): 3652-3673.

Grosset, Florian, Anna Papp, and Charles Taylor. "Rain follows the forest: Land use policy, climate change, and adaptation." Climate Change, and Adaptation (December 1, 2023) (2023).

Grossman, Gene M., and Elhanan Helpman. "Trade, innovation, and growth." The American economic review 80.2 (1990): 86-91.

Guerrico, Sofía Fernández. "The effects of trade-induced worker displacement on health and mortality in Mexico." Journal of health economics 80 (2021): 102538.

Han, Lu, Stephan Heblich, Christopher Timmins, and Yanos Zylberberg. Cool cities: The value of urban trees. No. w32063. National Bureau of Economic Research, 2024.

Harstad, Bård. "Trade and Trees: How Trade Agreements Can Motivate Conservation Instead of Depletion." (2020).

Heo, Seonmin Will, Koichiro Ito, and Rao Kotamarthi. International spillover effects of air pollution: evidence from mortality and health data. No. w30830. National Bureau of Economic Research, 2023.

Jayachandran, Seema, Joost De Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas. "Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation." Science 357, no. 6348 (2017): 267-273.

Jones, Benjamin A., and Andrew L. Goodkind. "Urban afforestation and infant health: Evidence from MillionTreesNYC." Journal of Environmental Economics and Management 95 (2019): 26-44.

Levchenko, Andrei A. "Institutional quality and international trade." The Review of Economic Studies 74.3 (2007): 791-819.

Li, Liqing. "Environmental goods provision and gentrification: Evidence from MillionTreesNYC." Journal of Environmental Economics and Management 120 (2023): 102828.

Managi, Shunsuke, Akira Hibiki, and Tetsuya Tsurumi. "Does trade openness improve environmental quality?." Journal of Environmental Economics and Management 58.3 (2009): 346-363.

Narain, Urvashi, and Chris Sall. "Methodology for valuing the health impacts of air pollution: discussion of challenges and proposed solutions." (2016).

Phillips, Benjamin B., James M. Bullock, Juliet L. Osborne, and Kevin J. Gaston. "Spatial extent of road pollution: A national analysis." Science of the Total Environment 773 (2021): 145589.

Pierce, Justin R., and Peter K. Schott. "Trade liberalization and mortality: evidence from US counties." American Economic Review: Insights 2, no. 1 (2020): 47-63.

UN Department of Economics of Social Affairs. "The Global Forest Goals Report 2021." United Nations Publication (2021).

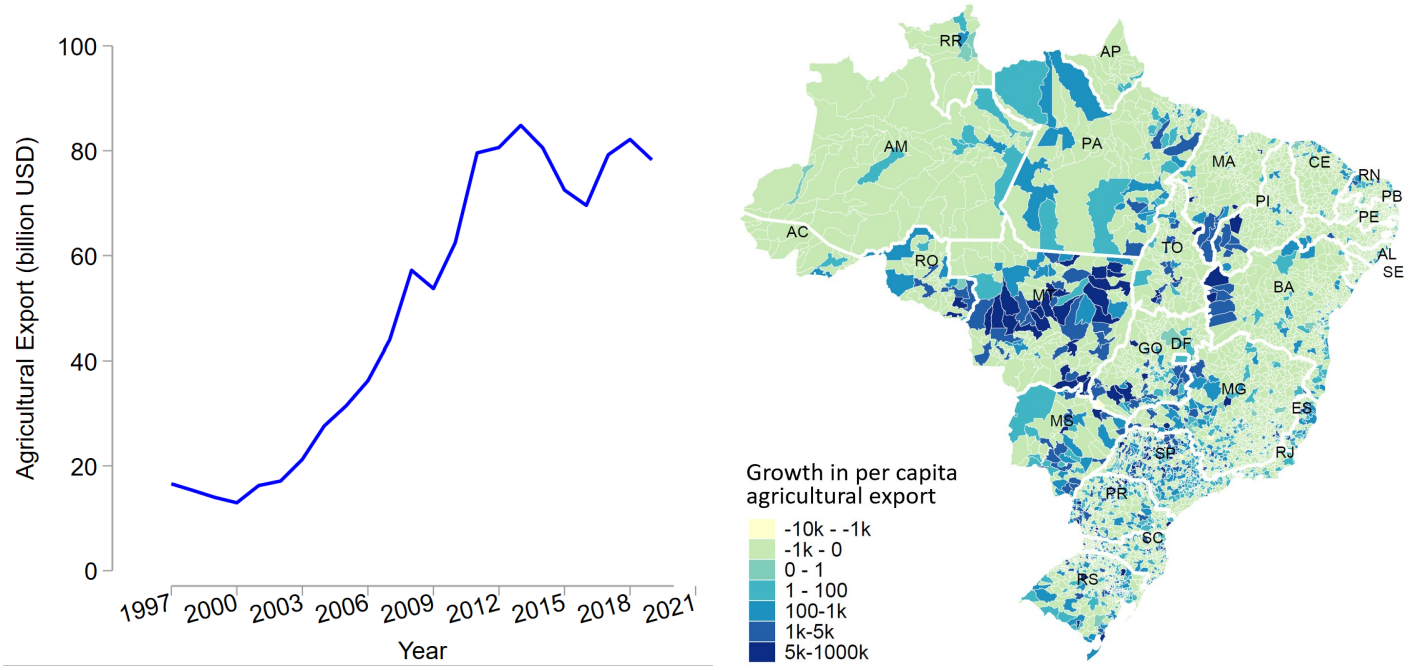
U.S. Environmental Protection Agency. "*Guidelines for Preparing Economic Analyses*." EPA 240-R-00-003 (2000).

Wei, Xiangying, Shiheng Lyu, Ying Yu, Zonghua Wang, Hong Liu, Dongming Pan, and Jianjun Chen. "Phylloremediation of air pollutants: exploiting the potential of plant leaves and leaf-associated microbes." Frontiers in plant science 8 (2017): 270745.

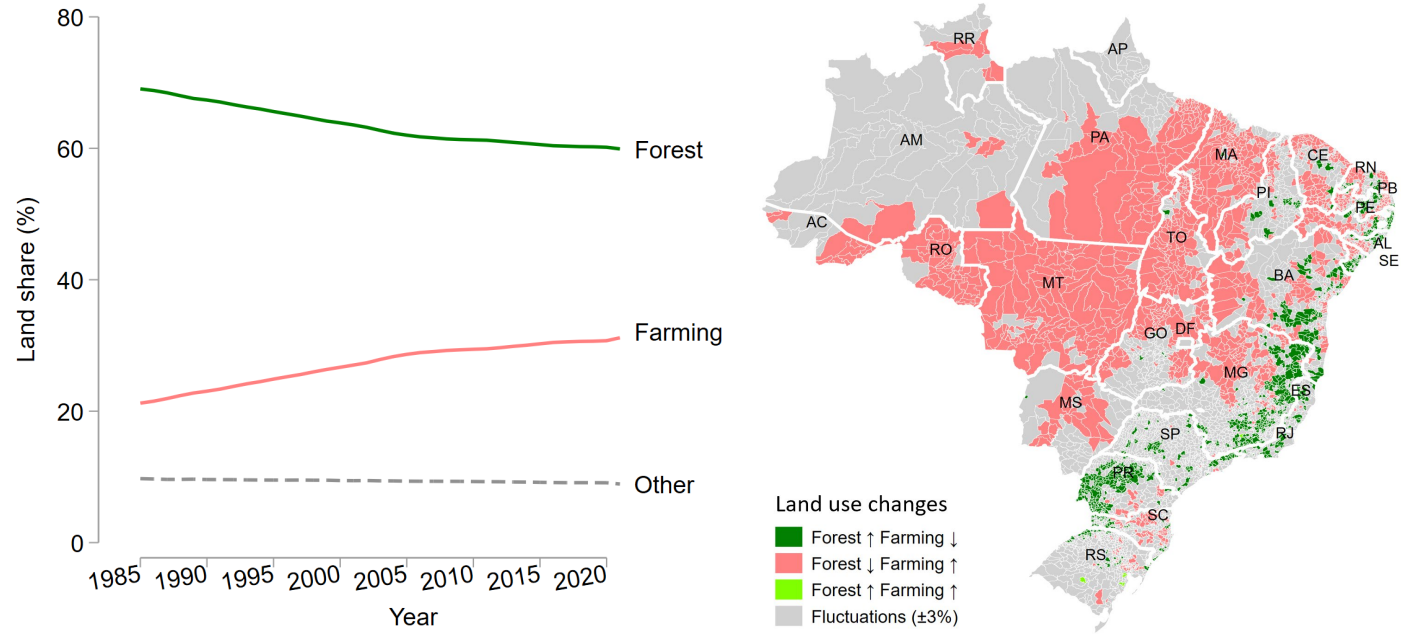
Xing, Jianwei, Zhiren Hu, Fan Xia, Jintao Xu, and Eric Zou. Urban Forests: Environmental Health Values and Risks. No. w31554. National Bureau of Economic Research, 2023.

Figure 1. Trends in Trade, Agricultural Expansion and Deforestation

(a) Agricultural Exports

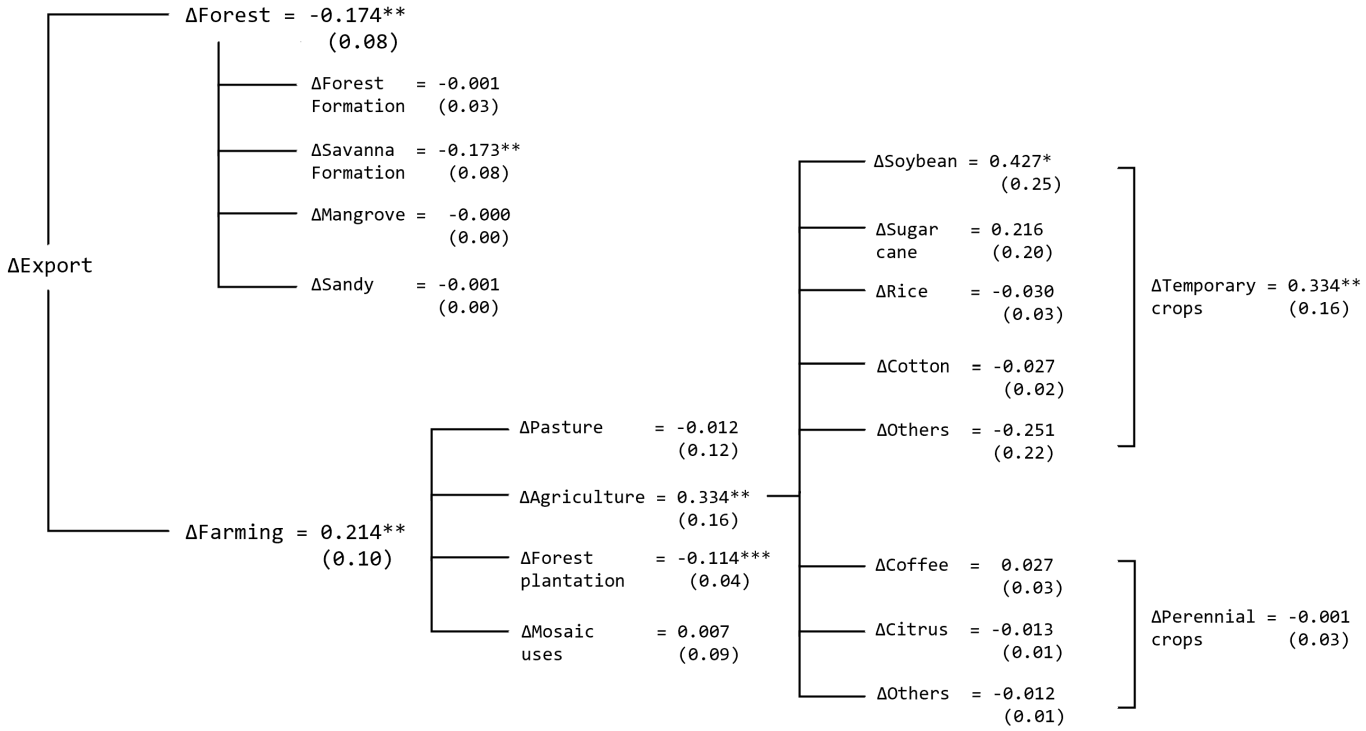


(b) Agricultural Land Expansion and Deforestation



Notes: Panel (a) shows temporal and geographic trends in Brazil's agricultural exports. Panel (b) shows temporal and geographic trends in the substitution between agricultural land and forests. "Fluctuations" correspond to areas with less than 3 percent changes in forest or farmland coverage over the study period.

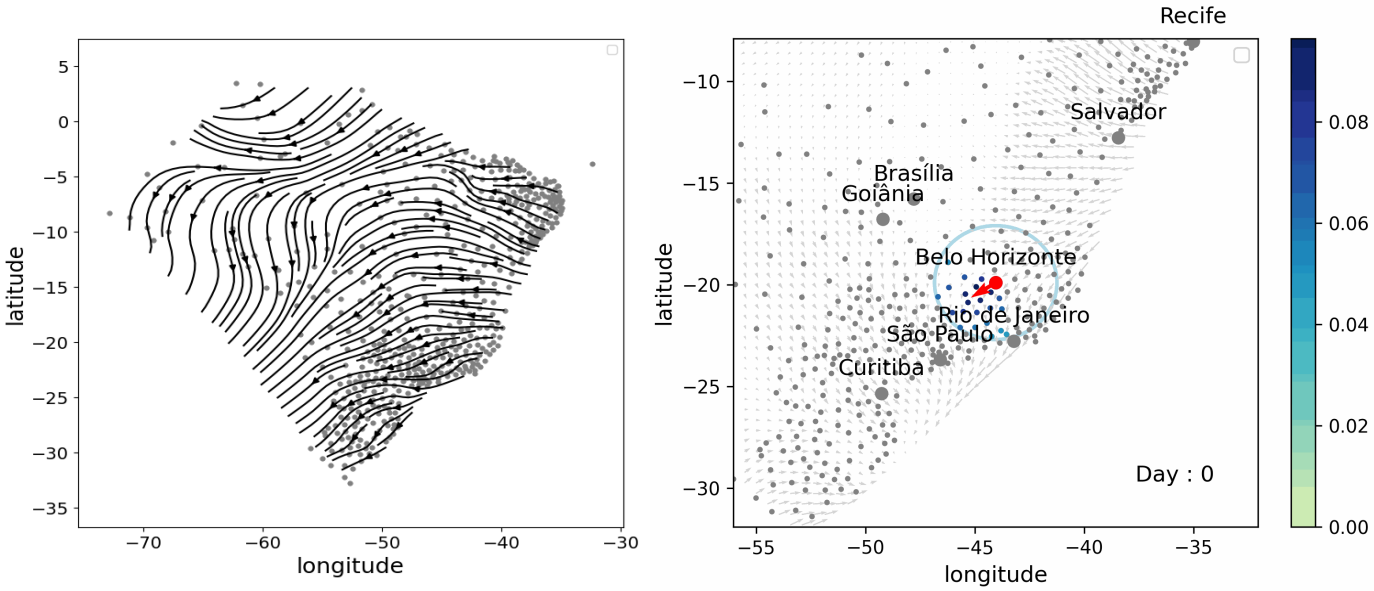
Figure 2. The Effect of Trade on Agricultural Expansion and Deforestation



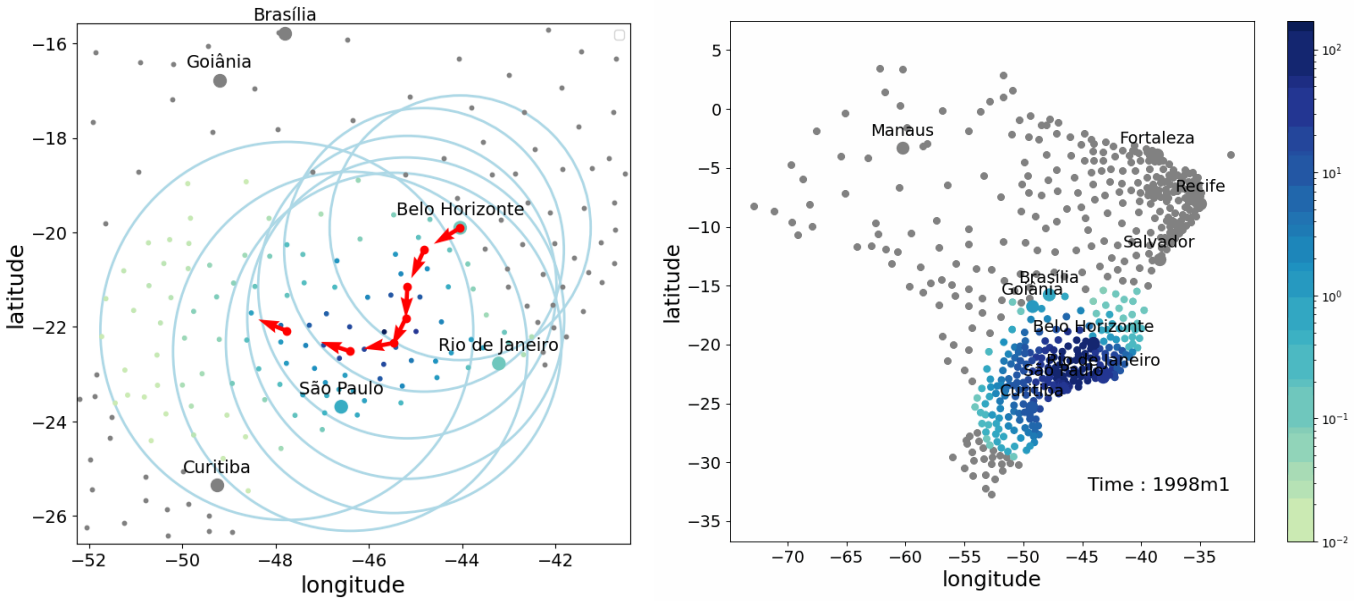
Notes: Each node on the tree represents a separate IV regression following the exact same specification except for the outcome variable, which is denoted by the name of the node. Each regression has 57,189 underlying observations with a first-stage Kleibergen-Paap F-statistic of 8.1. Branches of the tree represent hierarchies of the land use categorizations. Interpretation of the coefficients is change in land cover land use type (in percentage points) per 1,000 BRL increase in export per capita.

Figure 3. Area-of-Effect Modeling

(a) Airflow Modeling



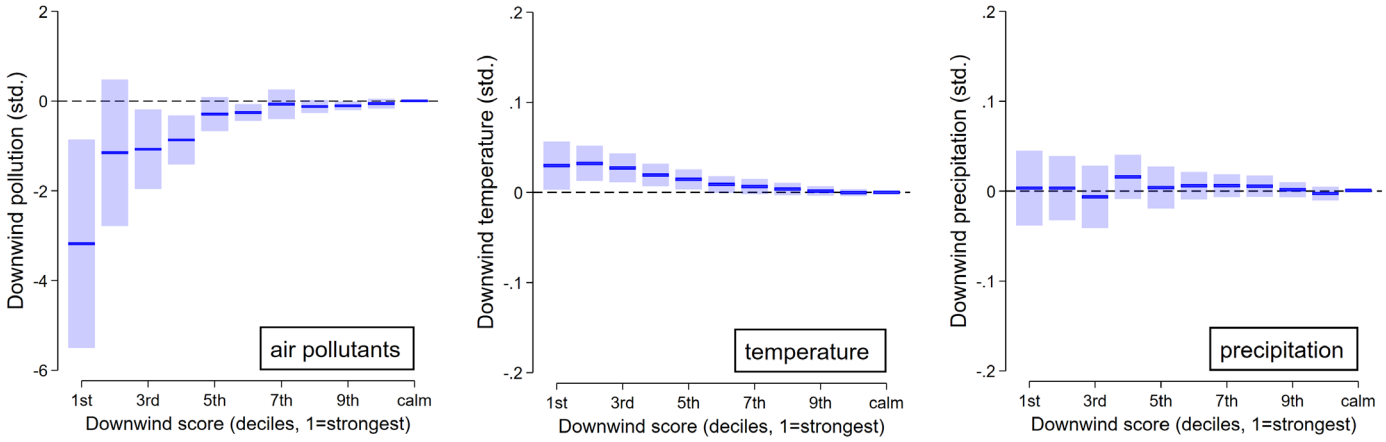
(b) Downwind Score: Belo Horizonte as Example



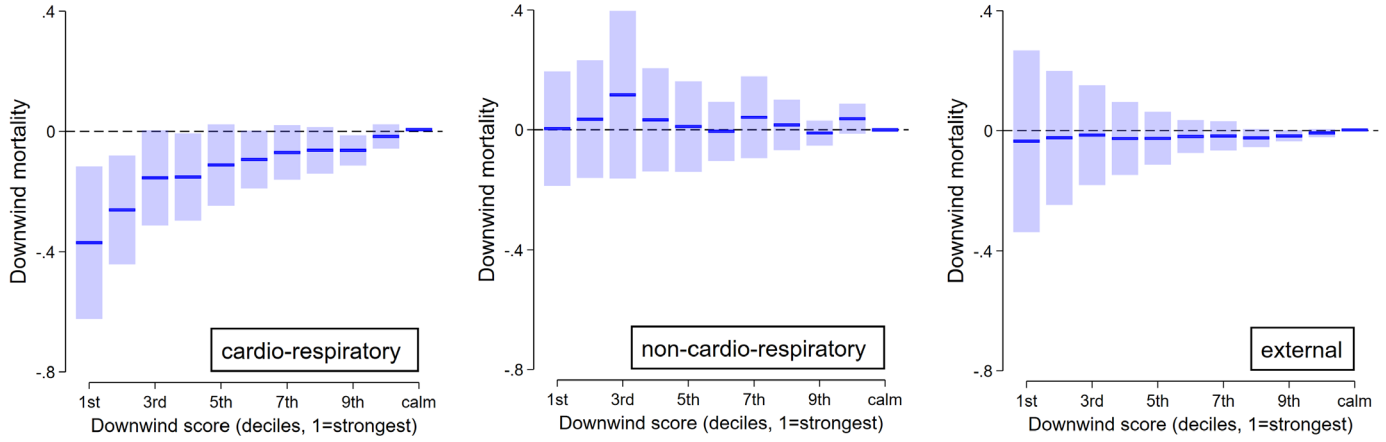
Notes: Figures show dotted map of Brazil where each dot represents centroid of a city. Panel (a) shows a day snapshot of wind flows built from location-specific wind vectors data (left) and a visualization of our wind flow model that tracks downwind cities influenced by winds from a sender city (Belo Horizonte in this example) over the course of seven steps (i.e., days). Panel (b) shows aggregation of step-specific downwind scores to sender-receiver-day level (left) and sender-receiver-month level (right) where the sender city is Belo Horizonte.

Figure 4. The Effect of Forests on Downwind Environment & Health

(a) Atmospheric Outcomes

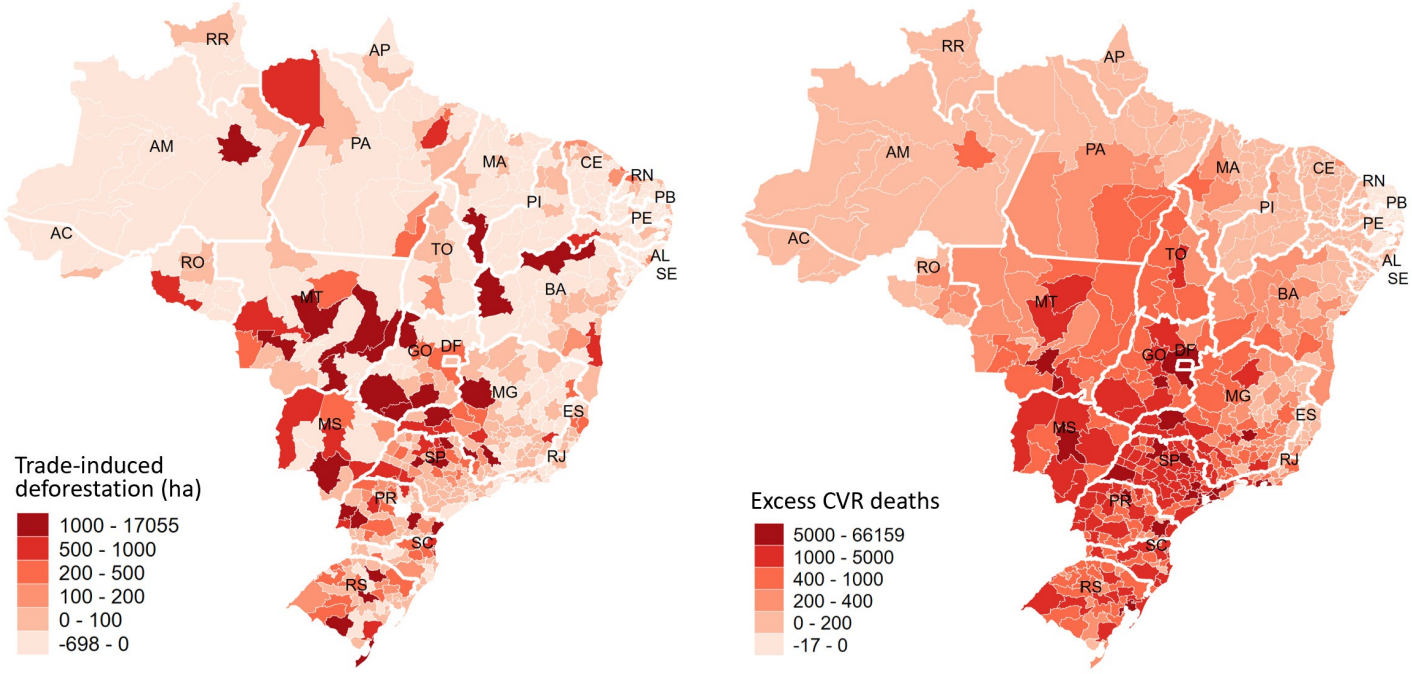


(b) Mortality Outcomes



Notes: Charts show estimates on changes in downwind outcomes per 1 SD increase in upwind forest cover, separately by downwind exposure score bins. Each chart shows a separate regression following the exact same specification except for the outcome variable. Within each chart, horizontal step lines show point estimates, and range bars show 95 percent confidence intervals.

Figure 5. Distribution of Trade-Induced Forest Losses and Excess Deaths

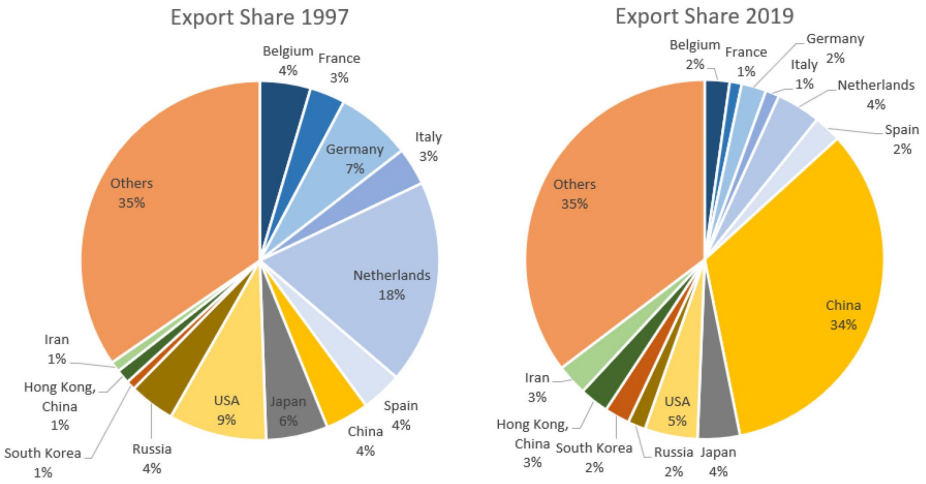


Notes: The map on the left shows the spatial distribution of trade-induced deforestation over the study period. The map on the right shows the spatial distribution of the downstream cardio-respiratory deaths due to trade-induced deforestation.

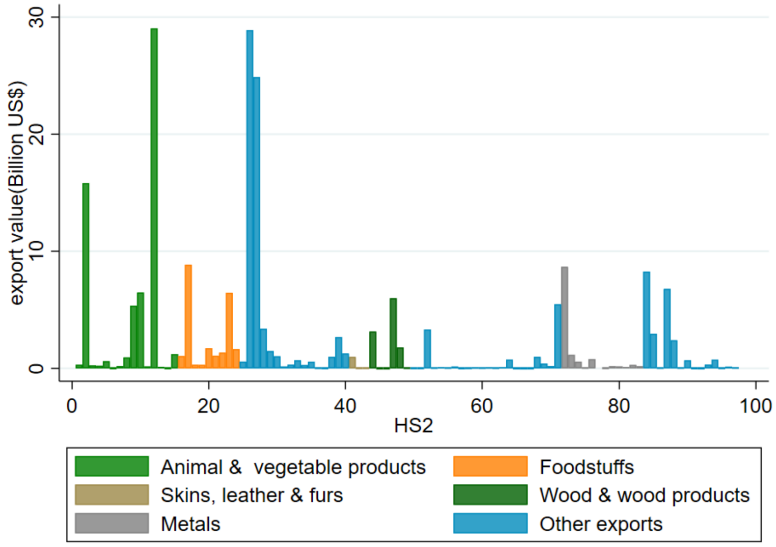
Appendix Figures and Tables

Appendix Figure 1. Brazil's Agricultural Export Structure

(a) Trade Partners

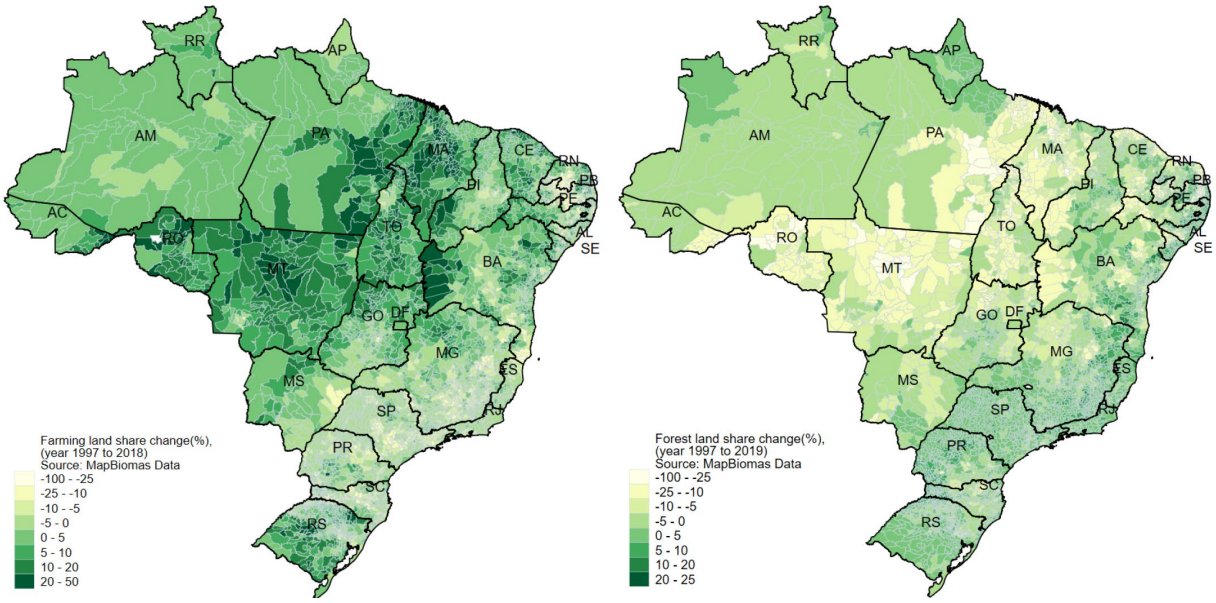


(b) Product Categories



Notes: Panel (a) shows agricultural export share by trade partner in 1997 (left) and in 2019 (right). Panel (b) distribution of export values in billions of USD by HS-2 product category.

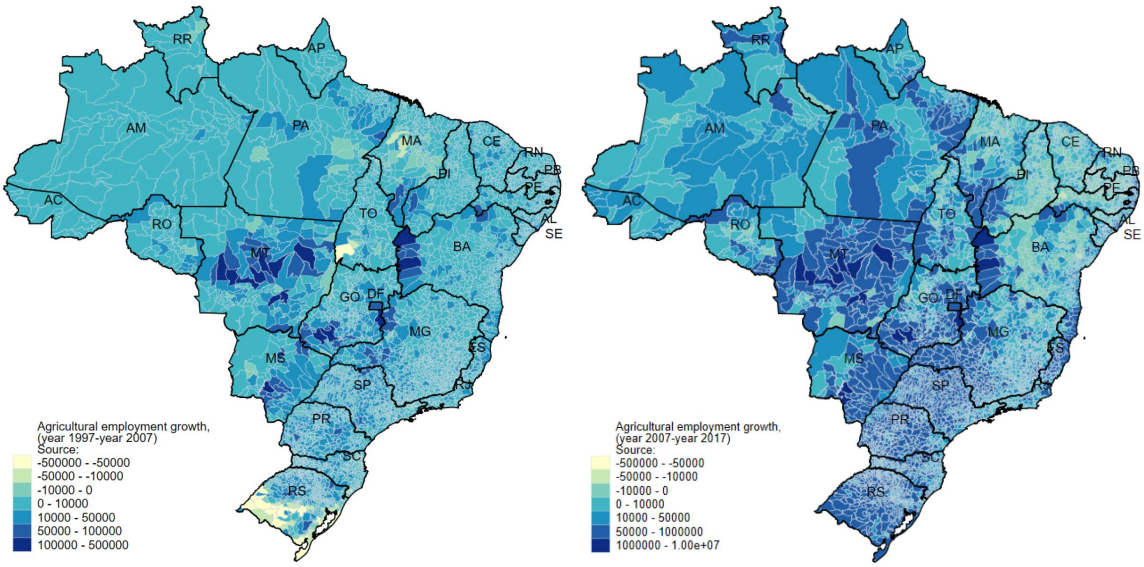
Appendix Figure 2. Geography of Land Use Changes



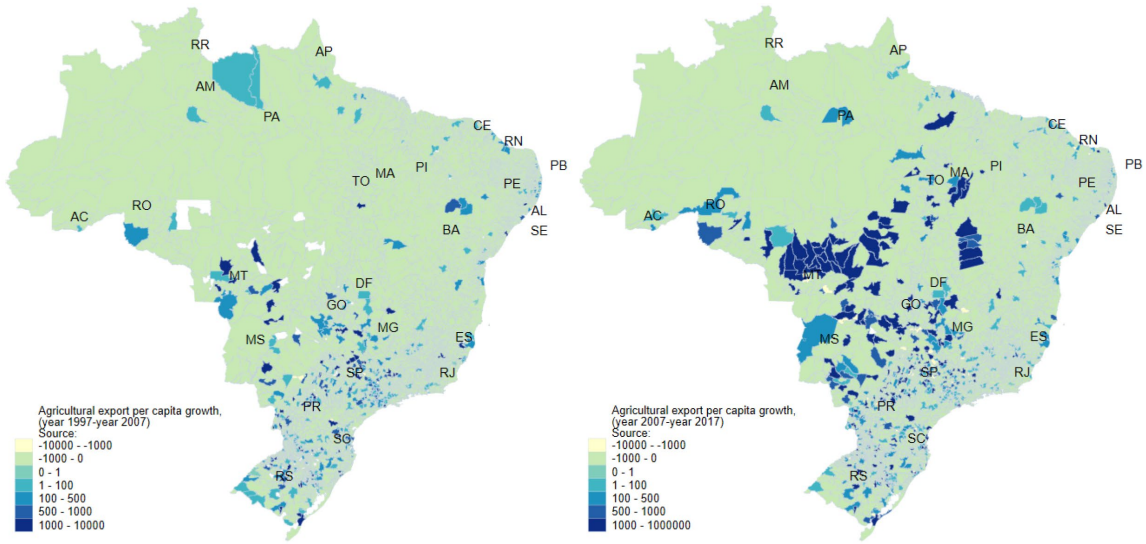
Notes: Maps show 1997-2018 percentage change in land use for farming purposes (left) and forest land (right).

Appendix Figure 3. Geography of Agricultural Growth

(a) Agricultural Employment Growth

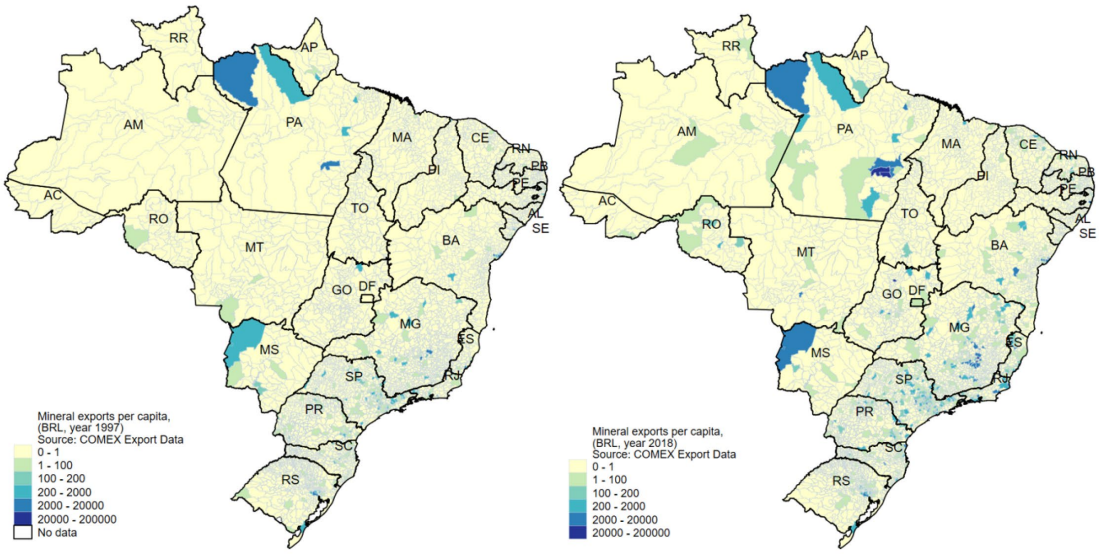


(b) Agricultural Export Growth



Notes: Panel (a) shows growth in agricultural employment over the 1997-2007 period (left) and the 2007-2019 period (right). Panel (b) shows growth in agricultural exports per capita over the 1997-2007 period (left) and the 2007-2019 period (right).

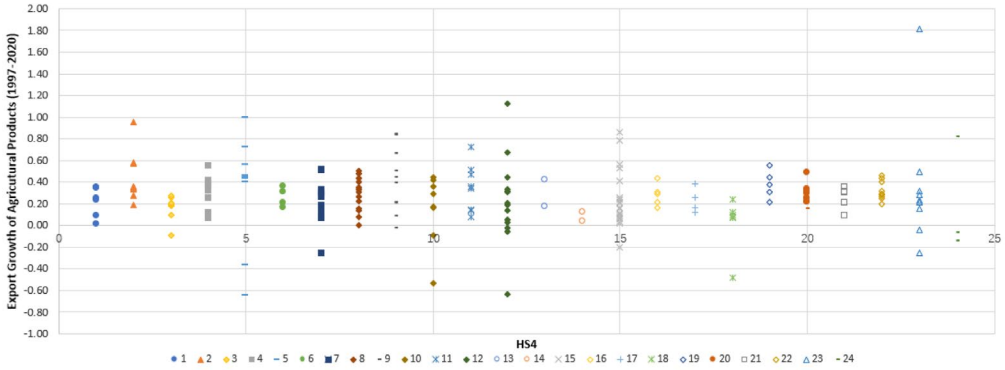
Appendix Figure 4. Geography of Mining Activities



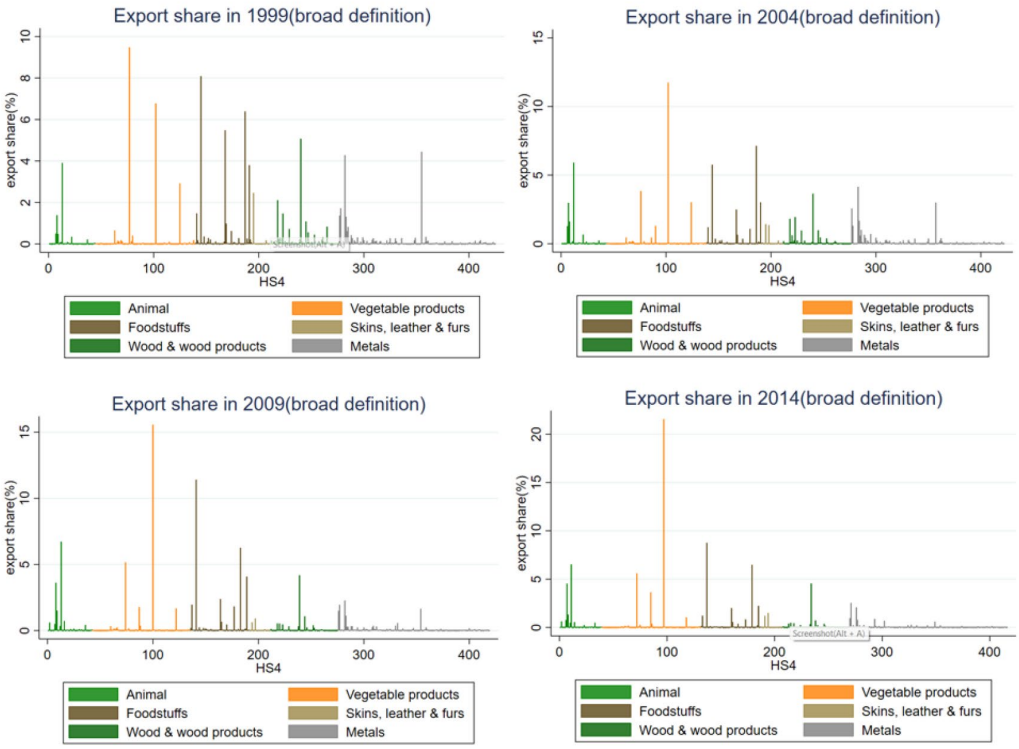
Notes: Maps show mineral exports per capita in year 1997 (left) and in year 2018 (right).

Appendix Figure 5. Product Categories Variability

(a) Export Growth Rates Variation across HS-4 Product Categories

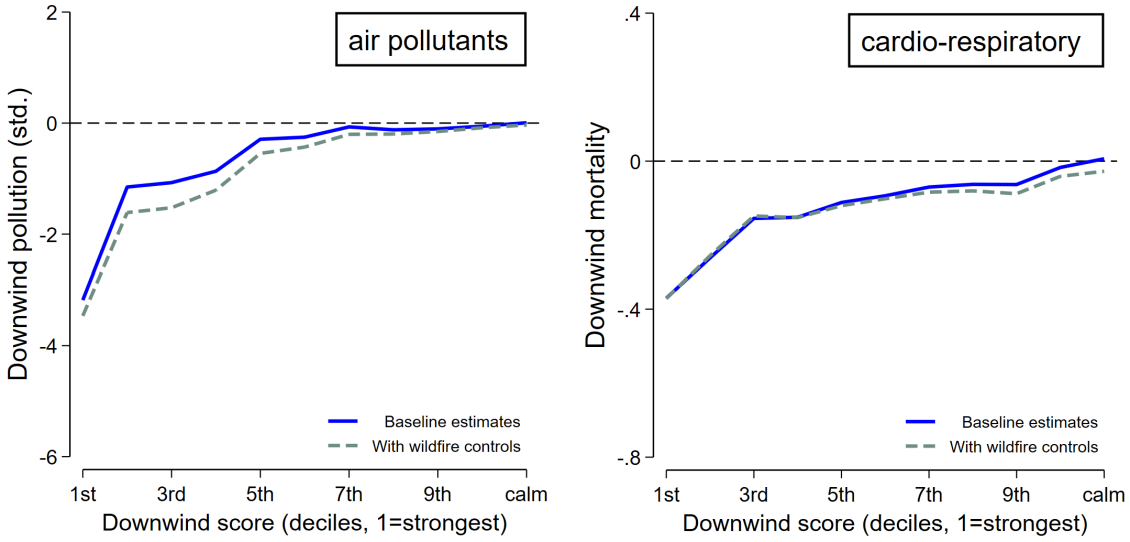


(b) Export Shares across HS-4 Product Categories Over Time



Notes: Panel (a) shows export growth rates variation across HS-4 product categories. Panel (b) shows export shares across HS-4 product categories in four different years: 1999, 2004, 2009, and 2014.

Appendix Figure 6. The Downwind Effect of Forests: Robustness to Wildfire Controls



Notes: Charts show estimates on changes in downwind outcomes per 1 SD increase in upwind forest cover, separately by downwind exposure score bins. Each chart shows two separate regression, one with and the other without controls for upwind wildfire occurrences.

Appendix Table 1. Major Import Countries Selection

Reporter	Rank in 2011	No. of Years with Missing UN Total Import Data	No. of Years with Missing UN Import to Brazil and Missing Comex Export Data	Correlation: UN Import to Brazil vs. Comex Export (above 0.8)	No. of Years with Low Correlation in 2010-2019: UN Import to Brazil vs. Comex Export (below 0.8)	Mismatch between UN Import Data HS4 and HS2 (difference>3%, 23y*24 HS2, <27 mismatches)	Selection
USA	1	0	0	0.989	0	0	Yes
Germany	2	0	0	0.881	0	200	No
Japan	3	0	0	0.981	0	0	Yes
China	4	0	0	0.994	0	0	Yes
United Kingdom	5	0	0	0.947	0	21	Yes
Netherlands	6	0	0	0.944	0	11	Yes
France	7	0	0	0.985	0	2	Yes
Italy	8	0	0	0.990	0	96	No
Belgium	9	2	2	0.859	3	0	No
Russian Federation	10	0	0	0.966	0	0	Yes
Spain	11	0	0	0.997	0	72	No
Canada	12	0	0	0.898	0	0	Yes
Rep. of Korea	13	0	0	0.968	0	0	Yes
Mexico	14	0	0	0.860	2	87	No
China, Hong Kong SAR	15	0	0	0.929	0	0	Yes
Nigeria	16	2	2	0.922	1	3	No
Saudi Arabia	17	1	1	0.992	0	0	No
Poland	18	0	0	0.824	0	0	Yes
Indonesia	19	0	0	0.961	0	0	Yes
Malaysia	20	0	0	0.951	0	0	Yes
Sweden	21	2	0	0.912	0	0	No
India	22	0	0	0.964	0	8	Yes
Egypt	23	1	0	0.907	0	8	No

Notes: This table tabulates Brazil's major export destinations ranked by Brazil's export value, highlighting countries that are selected in constructing the shift-share instruments. Excluded countries are those with either missing UN import or Comex export data, low correlation between UN import and Comex export data, or those with high numbers of mismatches between UN import HS-4 and HS-2 product categories.

Appendix Table 2. Summary Statistics

Variable Name	Obs	Mean	SD	P25	Median	P75
Δ4 Forest land share (%)	57189	-0.13	2.494	-0.621	.0989497	0.638
Δ4 Farming land share (%)	57189	0.05	2.666	-0.852	-.1989841	0.673
Δ4 Pasture land share (%)	57189	-1.10	4.194	-2.828	-.5760103	0.713
Δ4 Agriculture land share (%)	57189	0.98	2.754	-0.000	.0486577	1.151
Δ4 Forest plantation land share (%)	57189	0.26	0.732	0.000	.0011279	0.170
Δ4 Mosaic of uses land share (%)	57189	-0.09	2.884	-1.316	-.0046164	1.178
Δ4 Export per capita (1000 BRL)	57189	0.22	1.327	0.000	0	0.000
Δ4 Export per capita (real 1000 BRL)	57189	0.19	1.562	0.000	0	0.000
Population	57189	48431.26	2.45e5	7177.000	15305	33428.000
Agricultural employee	57189	23721.17	1.82e5	1092.818	3606	13316.364
Income per capita	57189	191.40	114.516	97.690	166.1	254.770
Literacy rate (%)	57189	70.32	16.445	55.300	75.7	83.900
Population density	57189	0.82	3.808	0.137	.2652524	0.510
Rural population	57189	9398.84	16462.777	2642.000	5484	11015.000
Urban population	57189	28374.83	1.99e5	2828.000	6121	15376.000
GE soy	57189	0.03	0.005	0.030	.0339846	0.036

Notes: The variables for land use and export per capita describe the growth over four years within the period from 2001 to 2019; Real export values have been deflated by the Brazilian Consumer Price Index (IPCA) calculated by IBGE and are denominated in 2019 reals; Income per capita, literacy rate, population density, rural population, and urban population are values from 1991, calculated by IBGE.

Appendix Table 3. Trade-Deforestation Regression Estimates: Alternative Time Lags

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. First stage	Export (Δ1y)	Export (Δ2y)	Export (Δ3y)	Export (Δ4y)	Export (Δ5y)	Export (Δ6y)
Shift-share IV	0.21	1.12***	1.78***	0.91***	1.12***	1.04***
	(0.14)	(0.12)	(0.30)	(0.23)	(0.25)	(0.28)
Panel B. Second stage	Forest (Δ1y)	Forest (Δ2y)	Forest (Δ3y)	Forest (Δ4y)	Forest (Δ5y)	Forest (Δ6y)
ΔExport per capita (1000 BRL)	-0.118	-0.038**	-0.043*	-0.174**	-0.207***	-0.192**
	(0.12)	(0.02)	(0.02)	(0.08)	(0.07)	(0.10)
Observations	68,585	68,585	64,814	57,189	49,564	41,939
First-stage F (Kleibergen-Paap)	1.250	81.63	41.01	8.072	16.59	6.492
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Macroregion FE	Yes	Yes	Yes	Yes	Yes	Yes
Control	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The IVs are constructed based on the agricultural imports by top import countries. All monetary values have been deflated by the Brazilian Consumer Price Index (IPCA) calculated by IBGE and are denominated in 2019 reals. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Technical Appendix: Area of Effect Estimation Details

To build a matrix that summarizes monthly wind flow intensities between all pairs of cities in Brazil, we designed a model that evaluates the intensity of the wind between a sender city and a receiver city at a given date. Beginning from a particular day and city, the model constructs streamlines by sequentially following the wind's speed and direction on a daily basis.

The input wind data (wind direction and speed information, i.e., vectors) are at a city-day level. Thus, the first step is to construct for each day a wind vector field with a grid resolution to estimate wind vectors at any point of the Brazilian territory. For that purpose, we divided the largest dimension of Brazil dim in polar coordinates (its width) by $res = 64$. Then, we built a coordinates grid with a resolution of $dim/res \approx 0.6$. The wind vector grid is then computed using linear interpolation from the city-day level wind data on the coordinates grid. Note that the higher the resolution res, the longer the computation so the value of res was chosen to obtain a sufficiently thin grid while keeping the computation time reasonable.

More precisely, we can illustrate how the algorithm works for a sender city i when starting at a given day d. We initialize the step t by $t = 0$ and the position p_t by the position of the sender city:

- At step t , we extract from the input data the wind direction and speed information at day $d + t$, and we build the daily vector grid as described previously. Given that p_t is not exactly at an intersection of the grid, nearest interpolation is used to approximate the wind vector at position p_t. Let's note that wind vector $w_t = (u_t, v_t)$.

- We look for potential receiver cities within a disk of radius rad_t from p_t. For each city r found, we define raw downwind intensity score as:

$$Wind_{i \to r,d,t} = \exp\{-\alpha \cdot rad_t - \beta \cdot |\theta|_{i \to r,d,t} - \gamma \cdot dist_{i \to r,d,t}\}$$

  where α, β, γ are positive parameters.

  The <u>first</u> component is the search radius at step t (rad_t), which captures general decreases of downwind intensity oversteps. We increase the search radius by 0.2 which represents about 21km at Brazil latitudes at each step to capture both the uncertainty in the streamline computation and the dispersion of air. The initial search radius rad_0 has a value of 2.8 which represents about 300km at Brazil latitudes.

  The <u>second</u> component enables to assign higher intensity to receiver cities that sit closer to the exactly-downwind direction of the sender. More precisely, let's note $l_{i \to r}$ the vector between sender city i and receiver city r. The sense of the vector is not important. Let's note v_t the vector defined as $v_t = (v_t, -u_t)/||w_t||$. Thus, it is a normal vector to the wind vector w_t that has a norm of 1. Then, we define

the absolute scalar product $|\theta|_{i\to r,d,t} = v_t \cdot l_{i\to r}$. The higher this term, the closer $l_{i\to r}$ is to be perpendicular to w_t. Since the higher $|\theta|_{i\to r,d,t}$, the lower the score $Wind_{i\to r,d,t}$, and thus the aim of this term is to penalize cities that are less impacted by the wind streamlines because they are not in the exactly-downwind direction from the wind streamline at that step. It is also important to note that v_t is normalized to avoid seeing lower scores when the speed of the wind is higher. However, $l_{i\to r}$ is not normalized to penalize cities that are further from the sender city.

The third component is simply the distance between the sender and the current position p_t ($dist_{i\to r,d,t}$), which captures geographic decay. This component is inspired by Phillips et al. (2021) that also uses an exponential decay with the distance from the emitter to model dispersion. This term also aims to penalize cities further from p_t but gives more flexibility to the formula by decorrelating the decay in terms of angle (second term) and the one in terms of distance.

We assume $Wind_{i\to r,d,t}$ to be zero if $d_{i\to r,d,t} > rad_t$ (i.e., if receiver city lies outside of the search radius at step t) or if $\theta_{i\to r,d,t} > 0.4$ radian (i.e., if the receiver city is not obviously in the downwind direction from the wind streamline at that step). We choose parameter values $\{\alpha, \beta, \gamma\} = \{0.8, 0.49, 0.23\}$. These coefficients are chosen empirically so that the function that attributes wind scores over 7 days is approximately continuous. For that purpose, we used visualisations consisting in heatmaps that simulate the wind scores values not only for cities of interest but for all points of the map for different days and different sender cities. Examples of those heatmaps showing the approximate continuity of the wind score function for the final value of the parameters can be seen on the figure below.

- If $t < 6$, coefficients need to be updated for step $t + 1$. We increase rad_t as described previously by 0.2 to obtain rad_{t+1}. We update $p_t = (x_t, y_t)$ by following the local direction and speed of the wind i.e. using w_t :

$$x_{t+1} = x_t + 24 * 3600 * u_t / dist_m((x_t, y_t), (x_t + 1, y_t))$$

$$y_{t+1} = y_t + 24 * 3600 * v_t / dist_m((x_t, y_t), (x_t + 1, y_t))$$

$$p_{t+1} = (x_{t+1}, y_{t+1})$$

To understand those expressions, we must consider that x and y coordinates are in degrees while the vectors' coordinates u and v are in m/s. The distance (positive or negative) in meters crossed by the wind in 24 hours is of $d_{m,x} = 24 * 3600 * u_t$ along the x-axis and $d_{m,y} = 24 * 3600 * v_t$ along the y-axis. To obtain an approximation of the distance d_p crossed in polar coordinates corresponding to a distance d_m in meters, we use a cross product : if a delta of 1 degree in longitude at the latitude y_t represents

$\text{dist}_m((x_t, y_t), (x_t + 1, y_t))$ meters, then, an approximation of d_p is $d_p \approx 1 * d_m / \text{dist}_m((x_t, y_t), (x_t + 1, y_t))$. Thus,

$$d_{p,x} \approx 1 * d_{m,x} / \text{dist}_m\big((x_t, y_t), (x_t + 1, y_t)\big)$$

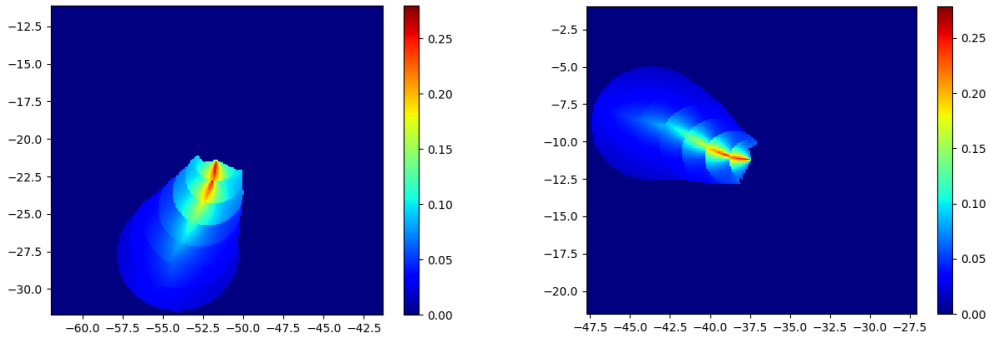
$$d_{p,y} \approx 1 * d_{m,y} / \text{dist}_m((x_t, y_t), (x_t + 1, y_t))$$

hence the expression of $x_{t+1} = x_t + d_{p,x}$ and $y_{t+1} = y_t + d_{p,y}$.

We can then proceed to step $t + 1$.

Starting from each particular sender city and day of the period 1998-2001, we iterate the procedure for seven steps (i.e., a week) so for $t = 0$ to $t = 6$.

Examples of wind indexes heat maps for two emitters



After the computation, we have a set of wind scores $\text{Wind}_{i \to r,d,t}$ that need to be aggregated at a day level, which means that we want to have a single coefficient for a given tuple (sender city i, receiver city r, date of arrival d'). For a given index a given index $\text{Wind}_{i \to r,d,t}$, the date of arrival d' is the sum of the delay in days from the emission at the sender city i.e. step t and of the date of emission $d : d' = t + d$. The downwind intensity score aggregated at a day-level for tuple (sender city i, receiver city r, date of arrival d') is:

$$\text{Wind}_{i \to r,d'} = \sum_{d+t=d'} \text{Wind}_{i \to r,d,t}$$

After that, the second aggregation step is at a month-level to make the size of the regression dataset manageable in the econometric analysis. This aggregation consists in computing average intensity scores for each tuple (sender city i, receiver city r, month of arrival m):

$$\text{Wind}_{i \to r,m} = \text{Average}_{d' \in m} \text{Wind}_{i \to r,d'}$$

At the end, we record every pair (sender city i, receiver city r) among the monthly aggregated intensity scores. For each pair (sender city i, receiver city r), when there is no intensity score found for a given month m of the period of interest (1998-2021), we add the monthly aggregated intensity score $\text{Wind}_{i \to r,m} = 0$. The final matrix

containing monthly intensity scores should therefore present n rows per couple (sender city i, receiver city r) where $n = 24$ is the number of months in the period (1998-2021). Note that not every pair of cities would be in the matrix. Indeed, if wind "originating" from a city i has not reached city r within 7 days, for each day of the period as starting date, then no scores will be associated to couple (sender city i, receiver city r) in the summary matrix.