Section 7 & 8: The Deforestation Effects of Trade and Agricultural Productivity in Brazil

Shuo Yu

2024-03-12

Before we begin today's section, please take 5 minutes to fill out the midterm discussion section evaluation form. Your feedback is valuable and helps improve the course experience. Thank you!

Today, we are going to explore the impact of agricultural productivity and trade on deforestation in Brazil by replicating results from Carreira et al. (2024).

Introduction to the Paper

Objective: Quantify the relative impacts of two drivers of deforestation in Brazil (2000–2017):

- **Agricultural technology**: Adoption of genetically engineered (GE) soy seeds.
- **International trade**: Rising Chinese demand for Brazilian commodities.

Key Question: Do productivity gains from GE soy seeds or export demand from China play a larger role in driving deforestation in Brazil?

Motivations

1. **Why Study Deforestation in Brazil?**

- Brazil is home to the Amazon rainforest, one of the world's most important carbon sinks, and plays a critical role in global climate regulation.
- However, agriculture is the main driver of deforestation especially in tropical countries. Deforestation contributes to climate change (CO_{2} emissions) and biodiversity loss.

- Understanding how trade and technological changes affect deforestation is crucial for policy design to balance economic growth and environmental conservation.

2. **The Role of Agricultural Trade and Productivity**

- Agricultural Productivity
  - The adoption of Genetically Engineered (GE) soy dramatically increased yields in Brazil after the early 2000s.
  - Theoretical Ambiguity:
    * Higher productivity means more soy can be grown on less land, potentially reducing pressure on forests.
    * However, if higher productivity increases profitability, farmers may be incentivized to expand soy plantations into forest areas. [Jevons Paradox: As resource use becomes more efficient, its demand increases, leading to an overall rise in resource consumption. This conjecture holds if there are no bidding constraints in factor or product markets limiting agricultural expansion.]
- Trade
  - Brazil is a major exporter of soy, supplying China and other global markets.
  - Theoretical Ambiguity:
    * Trade can yield productivity gains by offering access to cheaper inputs and stimulating innovation.
    * Improvements in agricultural productivity can alter a country's comparative advantage, fostering production specialization.
  - Previous studies have found mixed evidence.

```r
rm(list = ls())

library(haven)  # For reading .dta files
library(tidyverse)  # For data manipulation
library(fixest) # For fixed-effects regression
library(modelsummary) # For table formatting

setwd("~/Dropbox/161/Sections/Section7&8")
```

Data

Variables and Data Sources

- Land use in Brazil: MapBiomas
- Exposure to new agricultural technologies: Food and Agriculture Organization's project Global-Agroecological Zones (FAO-GAEZ). Exposure to new agricultural technologies is measured by the increase in potential soy yields due to modern inputs like GE soy seeds and fertilizers. It is calculated as the difference between high-input (modern) and low-input (traditional) potential yields in each municipality. This difference reflects how much productivity could improve with advanced farming methods.
- Exposure to Chinese demand: BACI-CEPII bilateral trade data and the 1996 Agricultural Census. The exposure to Chinese demand is measured by how much a municipality in Brazil was affected by the increase in China's import demand for agricultural commodities between 2000 and 2017. It is calculated in two steps:

    1. Measuring Chinese demand growth: The study first estimates how much China's demand for each agricultural product grew compared to the rest of the world (excluding Brazil), using international trade data.
    2. Allocating demand to municipalities: This growth in demand is then distributed across Brazilian municipalities based on their production composition in 1995. This means that municipalities that were already producing more of the goods highly demanded by China were considered more exposed to the demand shock.

    This measure captures how much a municipality's agricultural sector was influenced by China's increasing demand for Brazilian exports, particularly for soy and other commodities.
- Other control variables: 1991 Demographic Census.

```
data <- read_dta("data.dta")

vars_to_check <- c("iv2k_china", "A_soy_d", "intst")
data <- data %>%
  drop_na(all_of(vars_to_check))
```

Data Visualization

Cropland

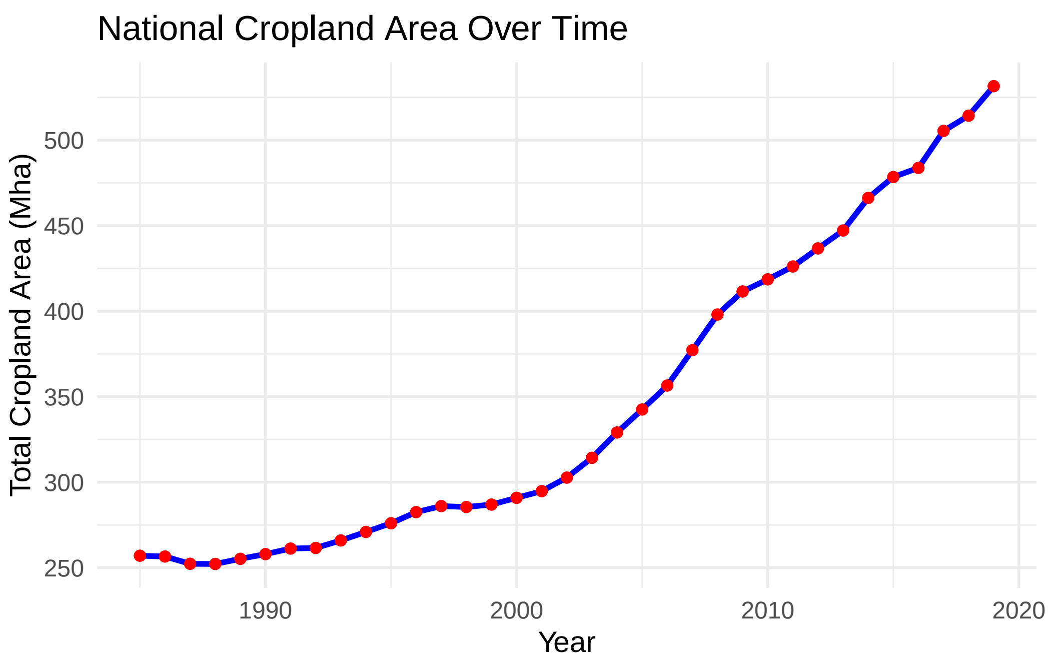
```r
# Step 1: Sum cropland area across all counties at the national level
national_cropland <- data %>%
  summarise(across(matches("^area_cropland_.*_mapbio$"), \(x) sum(x, na.rm = TRUE)))

# Step 2: Convert to Long Format
national_cropland_long <- national_cropland %>%
  pivot_longer(cols = everything(),
               names_to = "year",
               values_to = "total_cropland")

national_cropland_long <- national_cropland_long %>%
  mutate(year = as.numeric(gsub("area_cropland_|_mapbio", "", year))) # Extract numeric year


# Step 3: Plot the Time Trend
ggplot(national_cropland_long, aes(x = year, y = total_cropland)) +
  geom_line(color = "blue", size = 1) +  # Line plot
  geom_point(color = "red") +  # Points for each year
  labs(title = "National Cropland Area Over Time",
       x = "Year",
       y = "Total Cropland Area (Mha)") +
  theme_minimal()
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```



Practice: Natural Forest

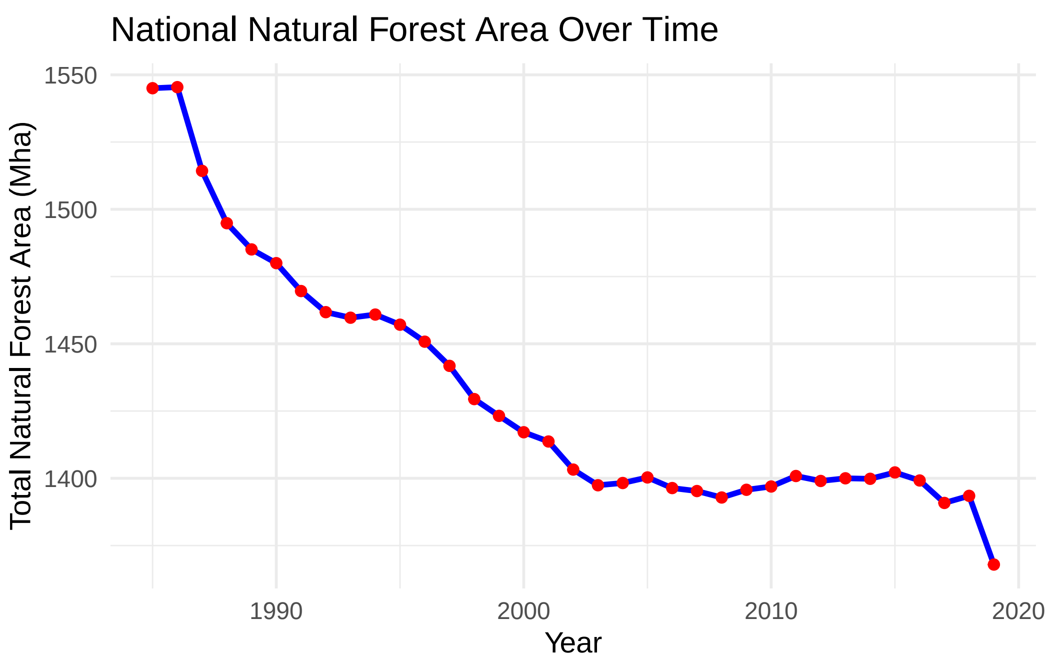
```
# Step 1: Sum forest area across all counties at the national level
national_natural_forest <- data %>%
  summarise(across(matches("^area_natural_forest_.*_mapbio$"), \(x) sum(x, na.rm = TRUE)))

# Step 2: Convert to Long Format
national_natural_forest_long <- national_natural_forest %>%
  pivot_longer(cols = everything(),
               names_to = "year",
               values_to = "total_forest")

national_natural_forest_long <- national_natural_forest_long %>%
  mutate(year = as.numeric(gsub("area_natural_forest_|_mapbio", "", year))) # Extract numeri


# Step 3: Plot the Time Trend
ggplot(national_natural_forest_long, aes(x = year, y = total_forest)) +
  geom_line(color = "blue", size = 1) +  # Line plot
  geom_point(color = "red") +  # Points for each year
  labs(title = "National Natural Forest Area Over Time",
       x = "Year",
```

```
    y = "Total Natural Forest Area (Mha)") +
  theme_minimal()
```



Regressions

Empirical Specification

The following equation estimates the impact of Chinese demand exposure and GE soy productivity gains on land-use changes from 2000 to 2017:

$$\Delta y_i = \alpha \hat{X}_i + \beta A_i + \delta(\hat{X}_i \times A_i) + W_i' \gamma + \eta_s + \epsilon_i$$

where:

- Δy_i = Change in land use (Forest, Cropland, or Pastureland) from 2000-2017.
- $\hat{X}_i$ = Exposure to Chinese demand.
- A_i = Exposure to GE soy productivity gains.
- $\hat{X}_i \times A_i$ = Interaction term between Chinese demand and GE soy exposure.
- W_i = Control variables (e.g., socioeconomic characteristics, geographic factors).
- η_s = state fixed effect.
- ϵ_i = Error term.

A long-difference regression compares changes in a variable over a long period of time, rather than using yearly or short-term variations.

- By regressing land use changes at the municipality level, the long-difference regression account for any factors specific to each municipality that influence land use.
- Land-use changes (deforestation, cropland expansion, pasture reduction) happen gradually, not in short-term fluctuations. A long difference (change over many years) helps capture long-term structural shifts, rather than short-term noise.
- If we estimate the effect of Chinese demand on land use using yearly data, we risk capturing a feedback loop: Deforestation or cropland expansion in one year might increase Brazil's soybean exports the next year. This could lead to China increasing its trade demand, making it appear as though Chinese demand is driving land use when the reverse is also true. A long-difference approach (2000-2017) measures the total land-use change over nearly two decades, rather than year-to-year fluctuations. Since trade policies and market adjustments happen gradually, any feedback effect from land-use changes on trade demand is less likely to affect the long-run results. Chinese demand for soybeans is primarily driven by China's own economic growth, consumption needs, and trade policies (especially the event of China's accession to the World Trade Organization (WTO) in 2001), which are exteranl factors not directly by Brazil's land-use decisions. While land-use changes in Brazil could affect trade in the short run, they are unlikely to drive China's long-term demand trends.

Panel A in Table 2: Impact on Deforestation

```
# Regression models
model1 <- feols(d2area_forest_2017_mapbio ~ iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model2 <- feols(d2area_forest_2017_mapbio ~ A_soy_d +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model3 <- feols(d2area_forest_2017_mapbio ~ A_soy_d + iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)
```

	(1)	(2)	(3)	(4)
Chinese Demand	−0.009*		−0.002	−0.056**
	(0.005)		(0.005)	(0.024)
GE Soy		−0.718***	−0.705**	−0.745***
		(0.247)	(0.264)	(0.260)
GE Soy × Chinese Demand				1.772**
				(0.723)

* p <0.1, ** p <0.05, *** p <0.01

```r
model4 <- feols(d2area_forest_2017_mapbio ~ A_soy_d + iv2k_china + intst +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)
```

```r
# Present the results from 4 models in one table
models <- list(
  "(1)" = model1,
  "(2)" = model2,
  "(3)" = model3,
  "(4)" = model4
)

# Define custom significance stars
custom_stars <- c('*' = 0.1, '**' = 0.05, '***' = 0.01)

# Define custom variable labels
labels <- c(
  "iv2k_china" = "Chinese Demand",
  "A_soy_d" = "GE Soy",
  "intst" = "GE Soy × Chinese Demand"
)

# Display in console
modelsummary(models, stars = custom_stars,
             coef_map = labels,
             gof_omit = ".*",
             fmt = 3)
```

Practice

Please replicate panel B and C in Table 2.

```r
# Cropland
# Regression models
model1 <- feols(d2area_cropland_2017_mapbio ~ iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model2 <- feols(d2area_cropland_2017_mapbio ~ A_soy_d +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model3 <- feols(d2area_cropland_2017_mapbio ~ A_soy_d + iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model4 <- feols(d2area_cropland_2017_mapbio ~ A_soy_d + iv2k_china + intst +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

# Present the results from 4 models in one table
models <- list(
  "(1)" = model1,
  "(2)" = model2,
  "(3)" = model3,
  "(4)" = model4
)

# Define custom significance stars
custom_stars <- c('*' = 0.1, '**' = 0.05, '***' = 0.01)

# Define custom variable labels
labels <- c(
  "iv2k_china" = "Chinese Demand",
  "A_soy_d" = "GE Soy",
  "intst" = "GE Soy × Chinese Demand"
)
```

	(1)	(2)	(3)	(4)
Chinese Demand	−0.028		−0.068**	0.141*
	(0.027)		(0.027)	(0.069)
GE Soy		3.623***	4.050***	4.200***
		(1.048)	(1.200)	(1.223)
GE Soy × Chinese Demand				−6.828***
				(1.932)

* p <0.1, ** p <0.05, *** p <0.01

```r
# Display in console
modelsummary(models, stars = custom_stars,
             coef_map = labels,
             gof_omit = ".*",
             fmt = 3)
```

```r
# Pasture
# Regression models
model1 <- feols(d2area_pasture_2017_mapbio ~ iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model2 <- feols(d2area_pasture_2017_mapbio ~ A_soy_d +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model3 <- feols(d2area_pasture_2017_mapbio ~ A_soy_d + iv2k_china +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

model4 <- feols(d2area_pasture_2017_mapbio ~ A_soy_d + iv2k_china + intst +
                  al_share + l_income_percapita_91 + literacy_rate_91 +
                  l_pop_density_91 + rural_pop_91 |
                  state, data = data)

# Present the results from 4 models in one table
```

	(1)	(2)	(3)	(4)
Chinese Demand	0.040		0.071**	−0.091
	(0.024)		(0.027)	(0.083)
GE Soy		−2.742**	−3.188**	−3.305**
		(1.170)	(1.327)	(1.355)
GE Soy × Chinese Demand				5.290**
				(2.490)

* p <0.1, ** p <0.05, *** p <0.01

```r
models <- list(
  "(1)" = model1,
  "(2)" = model2,
  "(3)" = model3,
  "(4)" = model4
)

# Define custom significance stars
custom_stars <- c('*' = 0.1, '**' = 0.05, '***' = 0.01)

# Define custom variable labels
labels <- c(
  "iv2k_china" = "Chinese Demand",
  "A_soy_d" = "GE Soy",
  "intst" = "GE Soy × Chinese Demand"
)

# Display in console
modelsummary(models, stars = custom_stars,
             coef_map = labels,
             gof_omit = ".*",
             fmt = 3)
```

Questions

1. Considering Chinese demand alone (Panel A, column 1), what is the effect of Chinese demand exposure on deforestation?

Municipalities more exposed to rising Chinese demand saw faster deforestation than less exposed areas. This suggests that export-driven land-use change contributed to forest loss.

2. Considering GE soy alone, how does exposure to GE soy productivity gains affect deforestation?

Municipalities with higher potential productivity gains from GE soy also experienced higher deforestation rates. The increase in productivity likely incentivized forest clearing for soy expansion.

3. Considering both Chinese demand and GE soy together, what happens to the effect of Chinese demand on deforestation when we also account for GE soy exposure? If we only looked at Chinese demand, what incorrect conclusion might we draw about deforestation?

Chinese demand effect disappears (not significant anymore) when GE soy exposure is included. GE soy's effect remains significant and unchanged. If we only looked at Chinese demand, we might wrongly conclude it caused deforestation. Instead, GE soy appears to be the key driver of forest loss.

4. Based on estimates in Panel A, column 3, how much more forest cover did a municipality at the 75th percentile of GE soy exposure (GE soy exposure=0.025) lose compared to one at the 25th percentile (GE soy exposure=0.011)? The study gives the example of Comodoro, Mato Grosso (21,520 km^2). How much additional forest area did it lose due to GE soy exposure?

The estimates in Panel A, column 3, suggest that a municipality at the 75th percentile of exposure to the GE soy technology (= 0.025) lost about 1 [=100×(−0.705)×(0.025–0.011)] percentage point of forest cover more than a municipality at the 25th percentile (= 0.011). To put that into perspective, the municipality of Comodoro in the state of Mato Grosso is close to the 75th percentile of the shock. With an area of approximately 21,520 km2, our estimates suggest that this single municipality lost about 212.6 km2 more forest area than a municipality not so exposed to the soy technology shock.

5. Based on estimates in Panel B and C, column 3, how much more cropland and pastureland did a municipality at the 75th percentile of GE soy exposure (GE soy exposure=0.025) gain (lose) compared to one at the 25th percentile (GE soy exposure=0.011)?

The estimates in Panel B and C, column 3, suggest that a municipality at the 75th percentile of exposure to the GE soy technology (= 0.025) gained about 5.7 [=100×(4.050)×(0.025–0.011)] percentage point of cropland and lost 4.5 [=100×(-3.19)×(0.025–0.011)] percentage points of pastureland more than a municipality at the 25th percentile (= 0.011). The main driver of deforestation in municipalities greatly exposed to GE soy seeds is the expansion in cropland over forest and pasture.

6. Based on estimates in Panel B and C, column 3, how much more cropland and less pastureland did a municipality at the 75th percentile of exposure to demand from China (Chinese demand exposure=0.009) lose (gain) compared to one at the 25th percentile (Chinese demand exposure=0.001)?

The estimates in Panel B, column 3, suggest that a municipality at the 75th percentile of exposure to the demand from China (0.009) gained about 5.7 [$=100\times(0.071)\times(0.009–0.001)$] percentage point of pastureland and lost 5.4 [$=100\times(-0.068)\times(0.009–0.001)$] percentage points of pastureland more than a municipality at the 25th percentile (0.001). Despite no relative change in deforestation, municipalities greatly exposed to demand from China experienced some substitution of pasture for cropland.

7. Considering the interaction between GE soy and Chinese demand, how does the interaction between the two shocks affect deforestation rates? Why do municipalities exposed to both shocks experience less deforestation than those exposed to GE soy alone?

The interaction term shows that exposure to both shocks led to less deforestation than areas only exposed to GE soy. Possible Explanation: Municipalities benefiting from both trade and technology shocks may have transitioned land use differently (more efficient soy production without requiring as much new land).

Our findings suggest different conclusions when analyzing the local effects of trade demand in isolation versus in conjunction with agricultural productivity gains. Specifically, we find evidence that, while Chinese demand did not seem to exert a substantial differential direct impact on deforestation, it mitigated the deforestation effects arising from GE soy seeds.

(Other evidence in the paper shows that regions significantly affected by both shocks did not exhibit a differential statistically significant increase in the share of land allocated to soy production compared to areas unaffected by both high trade and technology shocks. Thus, the attenuation effect on deforestation does not result from reduced soy expansion but rather from a marked decline in the cultivation of other temporary cash crops.)

8. Based on estimates in Panel A, column 4, how much less forest cover did a municipality with maximum Chinese demand exposure (Chinese demand exposure=1.273) and GE soy exposure at the 75th percentile (0.025) lose compared to one at the 25th percentile of GE soy exposure (GE soy exposure=0.011)? Consider again the exaple of Comodoro, Mato Grosso (21,520 km^2), how much less forest area did it lose due to higher Chinese demand exposure?

The estimates in Panel A, column 4, suggest that a municipality at the 75th percentile of exposure to the GE soy technology (= 0.025) gain about 2 [$=100\times(-0.745+1.772\times1.273)\times(0.025–0.011)$] percentage point of forest cover more than a municipality at the 25th percentile (= 0.011) if it has a maximum level of Chinese demand exposure. To put that into perspective, the municipality of Comodoro in the state of Mato Grosso is close to the 75th percentile of the shock. With an area of approximately 21,520 km2, our estimates suggest that this single municipality lost about 455 km2 less forest area than a municipality not so exposed to the soy technology shock.

References

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. link