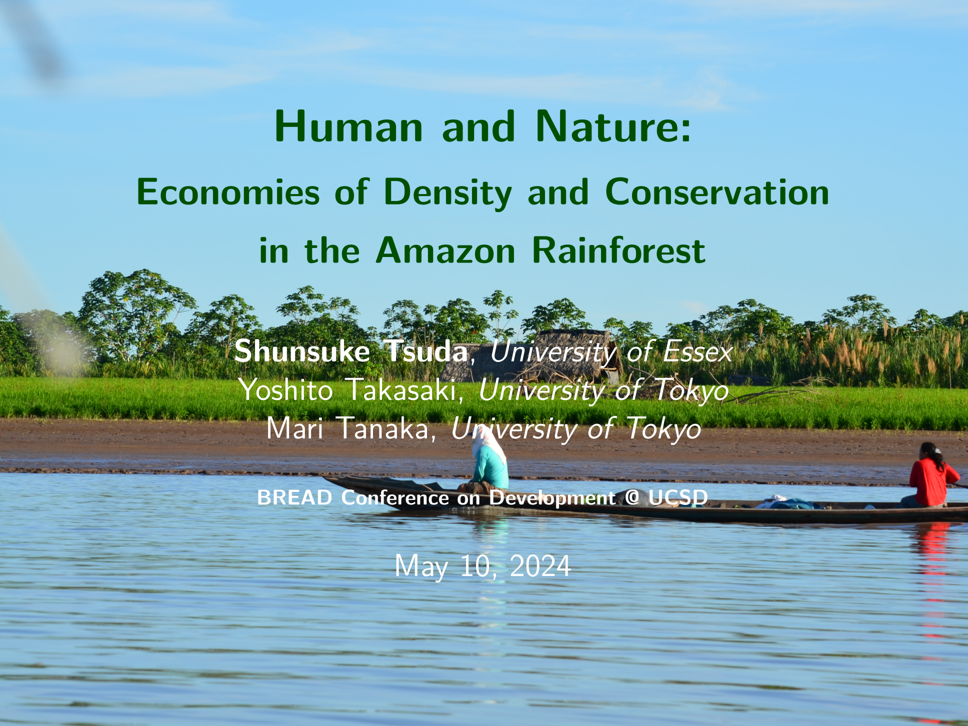


Human and Nature:
Economies of Density and Conservation in the Amazon Rainforest

Shunsuke Tsuda, *University of Essex*
Yoshito Takasaki, *University of Tokyo*
Mari Tanaka, *University of Tokyo*

BREAD Conference on Development @ UCSD

May 10, 2024

Human-Nature Interactions across Space in Rainforests

Tropical forests:

- Home to much of the world's bio-diversity and natural resources
- Growing concerns about the role of small-scale farmers in conservation

Human-Nature Interactions across Space in Rainforests



Tropical forests:

- Home to much of the world's bio-diversity and natural resources
- Growing concerns about the role of small-scale farmers in conservation

Policymaker's problems:

- Trade-off between rainforest conservation vs. local populations' welfare:

e.g. Cost of forest clearing $\uparrow \Rightarrow$ Agriculture income $\downarrow \Rightarrow$ Welfare $\downarrow$

Human-Nature Interactions across Space in Rainforests

Tropical forests:

- Home to much of the world's bio-diversity and natural resources
- Growing concerns about the role of small-scale farmers in conservation

Policymaker's problems:

- Trade-off between rainforest conservation vs. local populations' welfare:

e.g. Cost of forest clearing $\uparrow \Rightarrow$ Agriculture income $\downarrow \Rightarrow$ Welfare $\downarrow$

- Human adaptation may undermine conservation policy goals:

e.g. Sectoral reallocation: Cost of forest clearing $\uparrow \Rightarrow$ Fishing/hunting $\uparrow \Rightarrow$ Bio-diversity $\downarrow$
Spatial reallocation: Protected areas $\Rightarrow$ Resource depletion in other locations $\uparrow$

Human-Nature Interactions across Space in Rainforests

Tropical forests:

- Home to much of the world's bio-diversity and natural resources
- Growing concerns about the role of small-scale farmers in conservation

Policymaker's problems:

- Trade-off between rainforest conservation vs. local populations' welfare:

e.g. Cost of forest clearing $\uparrow \Rightarrow$ Agriculture income $\downarrow \Rightarrow$ Welfare $\downarrow$

- Human adaptation may undermine conservation policy goals:

e.g. Sectoral reallocation: Cost of forest clearing $\uparrow \Rightarrow$ Fishing/hunting $\uparrow \Rightarrow$ Bio-diversity $\downarrow$
Spatial reallocation: Protected areas $\Rightarrow$ Resource depletion in other locations $\uparrow$

Research question: Are there policies that improve both local populations' welfare and ecological conservation?

This Paper

Rural Agglomeration Economies and the Conservation

1. Spatial GE model with density externalities in multiple rural sectors (agriculture, forest clearing, and natural resource extraction):
   - **Agglomeration**: Population density $\uparrow \Rightarrow$ Productivity $\uparrow$
   - **Congestion**: Population density $\uparrow \Rightarrow$ Productivity $\downarrow$

This Paper

Rural Agglomeration Economies and the Conservation

1. Spatial GE model with density externalities in multiple rural sectors (agriculture, forest clearing, and natural resource extraction):
    - **Agglomeration**: Population density $\uparrow \Rightarrow$ Productivity $\uparrow$
    - **Congestion**: Population density $\uparrow \Rightarrow$ Productivity $\downarrow$
2. Estimation:
    - Geo-referenced data from river basins in the Peruvian Amazon
    - Exploits variation in the structure of river networks

This Paper

Rural Agglomeration Economies and the Conservation

1. Spatial GE model with density externalities in multiple rural sectors (agriculture, forest clearing, and natural resource extraction):
   - **Agglomeration**: Population density $\uparrow \Rightarrow$ Productivity $\uparrow$
   - **Congestion**: Population density $\uparrow \Rightarrow$ Productivity $\downarrow$
2. Estimation:
   - Geo-referenced data from river basins in the Peruvian Amazon
   - Exploits variation in the structure of river networks
3. Conterfactuals:
   - Protection policies
   - River transport infrastructure

Preview of Findings: Estimation

The density externalities:

- **Agglomeration in agriculture > Congestion in access to land**

  Concentration $\Rightarrow$ Productivity $\uparrow$ & Deforestation *per farmer* $\downarrow$

Preview of Findings: Estimation

The density externalities:

- **Agglomeration in agriculture > Congestion in access to land**

  Concentration $\Rightarrow$ Productivity $\uparrow$ & Deforestation *per farmer* $\downarrow$

- Without the agglomeration $\Rightarrow$ Welfare $\downarrow$ 10% & Deforestation $\uparrow$ 30%

Preview of Findings: Estimation

The density externalities:

- **Agglomeration in agriculture > Congestion in access to land**

  Concentration $\Rightarrow$ Productivity $\uparrow$ & Deforestation *per farmer* $\downarrow$
- Without the agglomeration $\Rightarrow$ Welfare $\downarrow$ 10% & Deforestation $\uparrow$ 30%
- **Congestion with spatial spillovers in natural resource extraction**

Preview of Findings: Estimation

The density externalities:

- **Agglomeration in agriculture > Congestion in access to land**

  Concentration $\Rightarrow$ Productivity $\uparrow$ & Deforestation *per farmer* $\downarrow$
- Without the agglomeration $\Rightarrow$ Welfare $\downarrow$ 10% & Deforestation $\uparrow$ 30%
- **Congestion with spatial spillovers in natural resource extraction**

Mechanisms behind the agglomeration:

- Economies of scale in transport technology
- Economies of scale in agricultural intensification

Preview of Findings: Counterfactuals

- Combining ***well-targeted*** place-based protection policies and transport infrastructure improves both human & ecological well-being:
    - Human welfare $\uparrow$
    - Deforestation $\downarrow$
    - Natural resource depletion $\downarrow$

Preview of Findings: Counterfactuals

- Combining ***well-targeted*** place-based protection policies and transport infrastructure improves both human & ecological well-being:
  - Human welfare $\uparrow$
  - Deforestation $\downarrow$
  - Natural resource depletion $\downarrow$
- Spatial targeting matters:
  - The *direction* of the natural resource depletion impacts depends on *where* protected areas are set
  - The *direction* of the deforestation impacts depends on *where* transport infrastructure is improved

Preview of Findings: Counterfactuals

- Combining ***well-targeted*** place-based protection policies and transport infrastructure improves both human & ecological well-being:
    - Human welfare $\uparrow$
    - Deforestation $\downarrow$
    - Natural resource depletion $\downarrow$
- Spatial targeting matters:
    - The *direction* of the natural resource depletion impacts depends on *where* protected areas are set
    - The *direction* of the deforestation impacts depends on *where* transport infrastructure is improved

⋆ The rural agglomeration economies are key to determining these impacts

Literature

1. Trade-offs between economic development and environmental goals

- **Agriculture and deforestation**: Abman et al. (2023); Angelsen (1999JDE; 2010PNAS); Carreira et al. (2024JDE); Foster et al. (2002); Szerman et al. (2022)
- **Policies for combating deforestation**: Alix-Garcia et al. (2015AEJ); Araujo et al. (2024); Asher et al. (2020EJ); Assunçao et al. (2023RES); Hsiao (2022); Jayachandran (2023ERL); Madhok (2024); Sousa-Rodrigues (2019RES)

Contributions

- Conservation in GE
- Agricultural productivity ↑ & Deforestation ↓
- Policies

Literature

1. Trade-offs between economic development and environmental goals

- **Agriculture and deforestation**: Abman et al. (2023); Angelsen (1999JDE; 2010PNAS); Carreira et al. (2024JDE); Foster et al. (2002); Szerman et al. (2022)
- **Policies for combating deforestation**: Alix-Garcia et al. (2015AEJ); Araujo et al. (2024); Asher et al. (2020EJ); Assunçao et al. (2023RES); Hsiao (2022); Jayachandran (2023ERL); Madhok (2024); Sousa-Rodrigues (2019RES)

Contributions

- Conservation in GE
- Agricultural productivity ↑ & Deforestation ↓
- Policies

2. Economic geography and quantitative spatial/urban models

- **Agglomeration economies**: Ahlfeldt et al. (2015ECMA); Duranton & Puga (2020JEP)
- **Environment/Deforestation in spatial models**: Balboni (2020); Conte (2024); Costinot et al. (2016JPE); Cruz & Rossi-Hansberg (2021); Farrokhi et al. (2024); Gollin & Wolfersberger (2023); Jedwab et al. (2023); Salazar & Mariante (2024)

Contributions

- Agglomeration externality in rural areas
- Endogenize both deforestation and common-pool resource extraction
- External validity

Data & Facts

The Peruvian Amazon

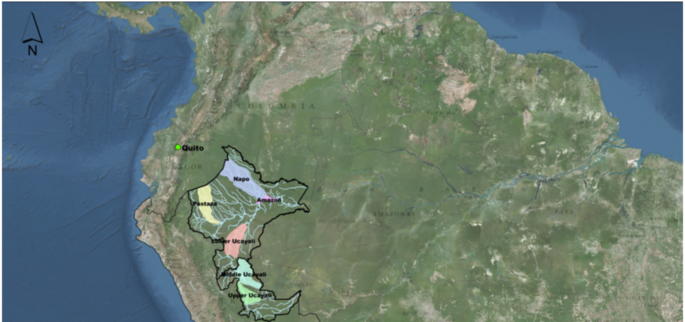
❶ Traditional ways of life in remote areas without modern technology and large-scale external investments (in contrast to the Brazilian Amazon)

⟹ Attribute resource extractions to small-scale farmers and hunter-gatherers and focus on externalities that they cause

- Primary livelihoods: agriculture (shifting cultivation), fishing, hunting, forest products
- Small-scale deforestation recently increased in the Amazon (Kalamandeen et al. 2018)

❷ River networks almost solely constitute the transportation routes

⟹ Identify key structural parameters by exploiting exogenous river shapes



Main Data Sources

1. **Peruvian Amazon Rural Livelihoods and Poverty (PARLAP)** project:
   - **Community census (CC, 2012-2014)** from rural communities ($n = 919$) in the four major river basins

     → Sectoral populations, prices, transport modes, and many others

   - Household survey (HS, 2014-2016) from selected communities

Main Data Sources

1. **Peruvian Amazon Rural Livelihoods and Poverty (PARLAP)** project:
   - **Community census (CC, 2012-2014)** from rural communities ($n = 919$) in the four major river basins

     → Sectoral populations, prices, transport modes, and many others
   - Household survey (HS, 2014-2016) from selected communities
2. **Satellite images**:
   - Forest cover measures by remote sensing experts
   - Grid cell-level (1km × 1km) and community-level

Main Data Sources

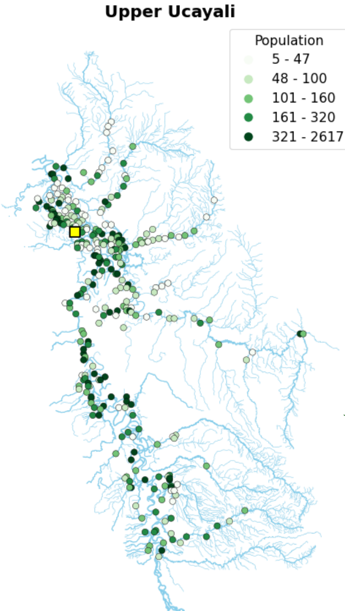
1. **Peruvian Amazon Rural Livelihoods and Poverty (PARLAP)** project:
   - **Community census (CC, 2012-2014)** from rural communities ($n = 919$) in the four major river basins

     $\rightarrow$ Sectoral populations, prices, transport modes, and many others
   - Household survey (HS, 2014-2016) from selected communities
2. **Satellite images**:
   - Forest cover measures by remote sensing experts
   - Grid cell-level (1km $\times$ 1km) and community-level
3. **National censuses** by National Institute of Statistics and Informatics (INEI):
   - Peru Population and Housing Census (2007, 2017)

     $\rightarrow$ Complement population information (esp. urban populations)
   - Peruvian Agricultural Census (2012)

     $\rightarrow$ Technology use by all producers

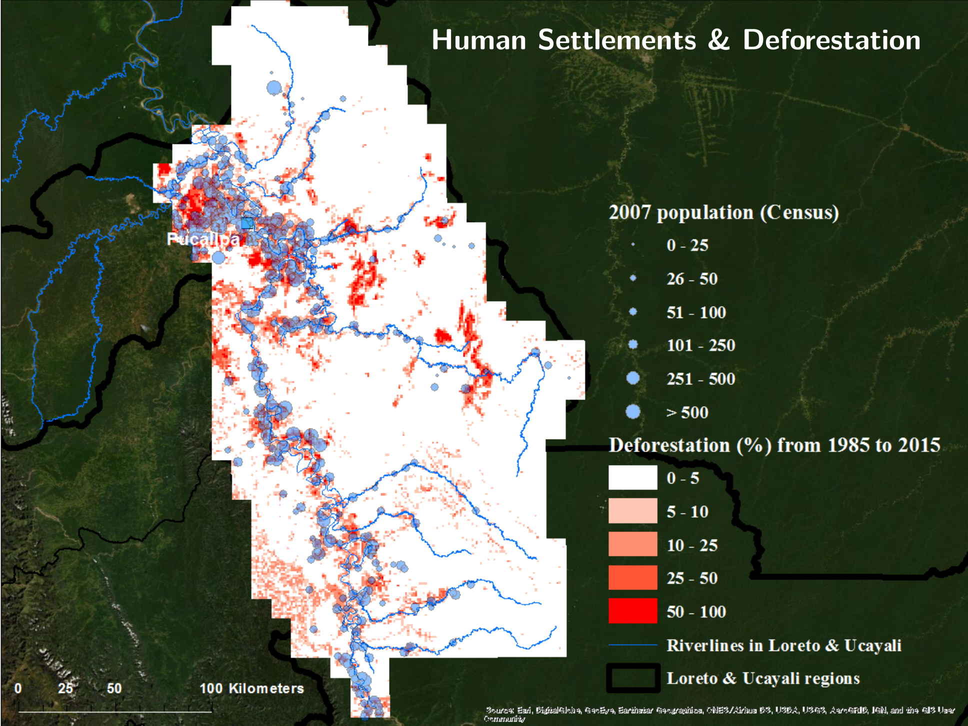
Spatial Concentration and Dispersion of Communities and Populations



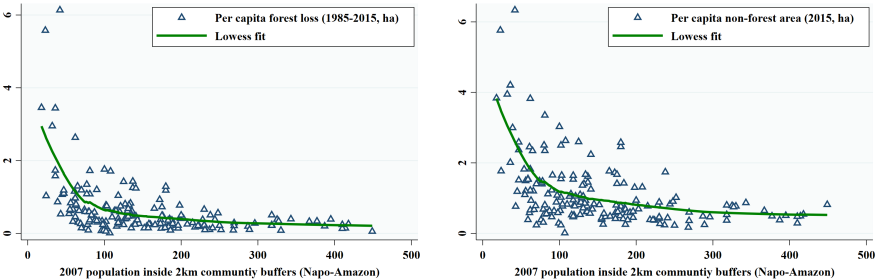
- A river basin
- The legend is based on quantiles (throughout the presentation)

  ≈ 80% of the rural communities have populations smaller than 320

⋆ Mostly indigenous and folk populations
Colonist settlements ≈ 1.4% of communities


Human Settlements & Deforestation
Pucallpa
2007 population (Census)
0 - 25
26 - 50
51 - 100
101 - 250
251 - 500
> 500
Deforestation (%) from 1985 to 2015
0 - 5
5 - 10
10 - 25
25 - 50
50 - 100
Riverlines in Loreto & Ucayali
Loreto & Ucayali regions
0
25
50
100 Kilometers
Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

Negative and Convex Relationship between Population & *Per Capita* Land Footprint



To simplify, take this as structural, then it implies:

- Congestion force in forest clearing (without land market)
- A mean-preserving reduction in the variance of the settlement size can reduce total deforestation

Dynamics?

Stylized Facts: Summary

1. Spatial concentration and dispersion of populations and communities

$\Rightarrow$ Suggestive of *strong agglomeration and congestion forces of economic activities*

2. Negative & convex relationship b/w population & per capita land footprint

$\Rightarrow$ Suggestive of:

- *Congestion force in forest clearing*
- *Reducing total deforestation by spatial reallocation of populations*

Stylized Facts: Summary

1. Spatial concentration and dispersion of populations and communities

⇒ Suggestive of *strong agglomeration and congestion forces of economic activities*

2. Negative & convex relationship b/w population & per capita land footprint

⇒ Suggestive of:

- *Congestion force in forest clearing*
- *Reducing total deforestation by spatial reallocation of populations*

3A. Agriculture is widely observed in both concentrated and dispersed areas

3B. Natural resource extraction is observed more in communities with lower population densities in the surrounding area Details

⇒ Suggests that *the spatial extent over which density externalities operate is distinct across sectors*

Model

Spatial Model of Rainforest Communities

Spatial general equilibrium in a river basin with:

- Trade across multiple **rural** locations and one urban center
- Mobile population within the river basin

3 Sectors:

- Agriculture (Ag): produced in rural locations
- Natural resource extraction (Nr): produced in rural locations
- Urban good (M): produced in the urban center

Spatial Model of Rainforest Communities

Spatial general equilibrium in a river basin with:

- Trade across multiple **rural** locations and one urban center
- Mobile population within the river basin

3 Sectors:

- Agriculture (Ag): produced in rural locations
- Natural resource extraction (Nr): produced in rural locations
- Urban good (M): produced in the urban center

The balance b/w concentration and dispersion forces determines the equilibrium:

- Concentration forces:
  - **Agglomeration in agricultural production**
  - Higher market access
  - Proximity to an urban center
- Dispersion forces:
  - **Congestion in land access by clearing forests**
  - **Congestion in natural resource extraction**

Agriculture with Agglomeration & Congestion Externalities

Land access by forest clearing: $$L_o(j) = \underbrace{A_{o,L} N_{o,Ag}^{-\mu_L}}_{\text{productivity}} \cdot N_{o,L}(j)$$

Final output: $$Q_{o,Ag}(j) = \underbrace{z_{o,Ag}(j) N_{o,Ag}^{\mu_{Ag}}}_{\text{productivity}} \cdot N_{o,C}(j)^{\gamma} L_o(j)^{(1-\gamma)}$$

- $N_{o,L}$ & $N_{o,C}$: Employment for forest clearing & cropping in community o
- $N_{o,Ag} = N_{o,L} + N_{o,C}$: Total employment for agriculture

Agriculture with Agglomeration & Congestion Externalities

Land access by forest clearing: $$L_o(j) = \underbrace{A_{o,L} N_{o,Ag}^{-\mu_L}}_{\text{productivity}} \cdot N_{o,L}(j)$$

Final output: $$Q_{o,Ag}(j) = \underbrace{z_{o,Ag}(j) N_{o,Ag}^{\mu_{Ag}}}_{\text{productivity}} \cdot N_{o,C}(j)^{\gamma} L_o(j)^{(1-\gamma)}$$

- $N_{o,L}$ & $N_{o,C}$: Employment for forest clearing & cropping in community o
- $N_{o,Ag} = N_{o,L} + N_{o,C}$: Total employment for agriculture

- $A_{o,L}$: productivity fundamentals for forest clearing
- $z_{o,Ag}(j)$: Fréchet shock of the productivity for cropping variety j (θ: comparative advantage, $A_{o,Ag}$: absolute advantage)

Agriculture with Agglomeration & Congestion Externalities

$$\textbf{Land access by forest clearing}: \quad L_o(j) = \underbrace{A_{o,L} N_{o,Ag}^{-\mu_L}}_{\text{productivity}} \cdot N_{o,L}(j)$$

$$\textbf{Final output}: \quad Q_{o,Ag}(j) = \underbrace{z_{o,Ag}(j) N_{o,Ag}^{\mu_{Ag}}}_{\text{productivity}} \cdot N_{o,C}(j)^{\gamma} L_o(j)^{(1-\gamma)}$$

- μ_{Ag}: Parameter governing **agglomeration forces in agricultural production and marketing** (e.g. facilitating collective investments)

Agriculture with Agglomeration & Congestion Externalities

Land access by forest clearing: $L_o(j) = \underbrace{A_{o,L} N_{o,Ag}^{-\mu_L}}_{\text{productivity}} \cdot N_{o,L}(j)$

Final output: $Q_{o,Ag}(j) = \underbrace{z_{o,Ag}(j) N_{o,Ag}^{\mu_{Ag}}}_{\text{productivity}} \cdot N_{o,C}(j)^{\gamma} L_o(j)^{(1-\gamma)}$

- μ_{Ag}: Parameter governing **agglomeration forces in agricultural production and marketing** (e.g. facilitating collective investments)
- μ_L: Parameter governing **congestion forces in forest clearing**

Congestion Forces in Forest Clearing

- Farmers clear forests to obtain land only nearby their residential locations along the river (mean/median of land footprint depths = 1 km/0.85km)
  - High monitoring cost with weak property rights
  - High cost of carrying products from inland to the riverside

Congestion Forces in Forest Clearing

- Farmers clear forests to obtain land only nearby their residential locations along the river (mean/median of land footprint depths = 1 km/0.85km)
  - High monitoring cost with weak property rights
  - High cost of carrying products from inland to the riverside
- Costs of monitoring and negotiating land allocations to farmers increase with the agricultural population

$\Rightarrow$ Deforested areas cannot increase proportionally to the increase in the community population size

Congestion Forces in Forest Clearing

- Farmers clear forests to obtain land only nearby their residential locations along the river (mean/median of land footprint depths = 1 km/0.85km)
  - High monitoring cost with weak property rights
  - High cost of carrying products from inland to the riverside
- Costs of monitoring and negotiating land allocations to farmers increase with the agricultural population

⇒ Deforested areas cannot increase proportionally to the increase in the community population size

- These costs may be small among very small communities: consistent with the convexity

Natural Resources with Congestion Externality across Space

$$Q_o^{Nr}(j) = \underbrace{z_{o,Nr}(j)\Big[\sum_{d\in\mathcal{R}} D_{od}^{-\nu} N_{d,Nr}\Big]^{-\mu_{Nr}}}_{\text{productivity}} \cdot N_{o,Nr}(j)$$

- People travel for natural resource extraction
- $N_{o,Nr}$: Total employment for natural resource extraction in o

Natural Resources with Congestion Externality across Space

$$Q_o^{Nr}(j) \quad = \quad \underbrace{z_{o,Nr}(j)\Big[\sum_{d\in\mathcal{R}} D_{od}^{-\nu} N_{d,Nr}\Big]^{-\mu_{Nr}}}_{\text{productivity}} \cdot N_{o,Nr}(j)$$

- People travel for natural resource extraction
- $N_{o,Nr}$: Total employment for natural resource extraction in o
- μ_{Nr}: Parameter governing **congestion forces with spatial spillovers from surrounding populations**
- D_{od}: River-equivalent distance between cells o & d in the shortest path
- ν: Parameter governing **spatial decay in accessing natural resources**

($z_{o,Nr}(j)$: Fréchet shock of variety j productivity)

Across-sector externality?

Estimation

Estimating the Model in a Sequential Procedure

Parameter	Description	Estimation strategy
δ_K	Elasticity of trade cost $(K = Ag, Nr, M)$	Commodity prices from the CC
λ_{up}, λ_{land}	Relative costs in terms of downstream river	Travel time and transport costs survey
σ	Within-sector elasticity of substitution	Expenditure information from ENAHO
$\bar{\sigma}$	Across-sector elasticity of substitution	Expenditure information from ENAHO
γ	Labor share in agricultural production	From the literature
θ	Trade elasticity	From the literature
μ_L	**Congestion in forest clearing**	Linear IV using the community-level data
μ_{Ag}	**Agglomeration in agricultural production**	Model inversion and linear IV
μ_{Nr}	**Congestion in natural resource extraction**	Model inversion and non-linear GMM
ν	**Spatial decay in natural resource access**	Model inversion and non-linear GMM
$\{A_{o,K}\}$	Absolute advantages $(K = Ag, Nr)$	Calibration
A_M	Absolute advantages in the urban sector	Calibration

Estimating the Model in a Sequential Procedure

Parameter	Description	Estimation strategy
δ_K	Elasticity of trade cost $(K = Ag, Nr, M)$	Commodity prices from the CC
λ_{up}, λ_{land}	Relative costs in terms of downstream river	Travel time and transport costs survey
σ	Within-sector elasticity of substitution	Expenditure information from ENAHO
$\bar{\sigma}$	Across-sector elasticity of substitution	Expenditure information from ENAHO
γ	Labor share in agricultural production	From the literature
θ	Trade elasticity	From the literature
μ_L	**Congestion in forest clearing**	Linear IV using the community-level data
μ_{Ag}	**Agglomeration in agricultural production**	Model inversion and linear IV
μ_{Nr}	**Congestion in natural resource extraction**	Model inversion and non-linear GMM
ν	**Spatial decay in natural resource access**	Model inversion and non-linear GMM
$\{A_{o,K}\}$	Absolute advantages $(K = Ag, Nr)$	Calibration
A_M	Absolute advantages in the urban sector	Calibration

1. Obtain some parameters without solving the model ▸ Details

Estimating the Model in a Sequential Procedure

Parameter	Description	Estimation strategy
δ_K	Elasticity of trade cost $(K = Ag, Nr, M)$	Commodity prices from the CC
λ_{up}, λ_{land}	Relative costs in terms of downstream river	Travel time and transport costs survey
σ	Within-sector elasticity of substitution	Expenditure information from ENAHO
$\bar{\sigma}$	Across-sector elasticity of substitution	Expenditure information from ENAHO
γ	Labor share in agricultural production	From the literature
θ	Trade elasticity	From the literature
μ_L	**Congestion in forest clearing**	Linear IV using the community-level data
μ_{Ag}	**Agglomeration in agricultural production**	Model inversion and linear IV
μ_{Nr}	**Congestion in natural resource extraction**	Model inversion and non-linear GMM
ν	**Spatial decay in natural resource access**	Model inversion and non-linear GMM
$\{A_{o,K}\}$	Absolute advantages $(K = Ag, Nr)$	Calibration
A_M	Absolute advantages in the urban sector	Calibration

1. Obtain some parameters without solving the model ▸ Details
2. Given the parameters obtained in the previous step, invert the model to recover wages and productivities that rationalize the observable endogenous variables as a spatial equilibrium ▸ Details

Estimating the Model in a Sequential Procedure

Parameter	Description	Estimation strategy
δ_K	Elasticity of trade cost $(K = Ag, Nr, M)$	Commodity prices from the CC
λ_{up}, λ_{land}	Relative costs in terms of downstream river	Travel time and transport costs survey
σ	Within-sector elasticity of substitution	Expenditure information from ENAHO
$\bar{\sigma}$	Across-sector elasticity of substitution	Expenditure information from ENAHO
γ	Labor share in agricultural production	From the literature
θ	Trade elasticity	From the literature
μ_L	**Congestion in forest clearing**	Linear IV using the community-level data
μ_{Ag}	**Agglomeration in agricultural production**	Model inversion and linear IV
μ_{Nr}	**Congestion in natural resource extraction**	Model inversion and non-linear GMM
ν	**Spatial decay in natural resource access**	Model inversion and non-linear GMM
$\{A_{o,K}\}$	Absolute advantages $(K = Ag, Nr)$	Calibration
A_M	Absolute advantages in the urban sector	Calibration

1. Obtain some parameters without solving the model ▸ Details
2. Given the parameters obtained in the previous step, invert the model to recover wages and productivities that rationalize the observable endogenous variables as a spatial equilibrium ▸ Details
3. Employ GMM to estimate density externality parameters from inverted productivities obtained in the previous step

Step 3. Density Externalities in Agriculture

- Inverted productivity composites of agriculture:

$$\tilde{A}_{o,Ag} \equiv \underbrace{A_{o,Ag} A_{o,L}^{(1-\gamma)\theta} \kappa_1^{\theta}}_{\text{fundamentals}} \cdot \underbrace{N_{o,Ag}^{\tilde{\mu}_{Ag}\theta}}_{\text{externalities}} \quad \text{where} \quad \tilde{\mu}_{Ag} \equiv \mu_{Ag} - (1-\gamma)\mu_L$$

- Procedure:
  1. Estimate $\tilde{\mu}_{Ag}$
  2. Estimate μ_L
  3. Back out μ_{Ag}

Step 3. Agglomeration Externality in Agriculture via Linear IV

- Inverted productivity composites of agriculture:

$$\tilde{A}_{o,Ag} \equiv \underbrace{A_{o,Ag} A_{o,L}^{(1-\gamma)\theta} \kappa_1^{\theta}}_{\text{fundamentals}} \cdot \underbrace{N_{o,Ag}^{\tilde{\mu}_{Ag}\theta}}_{\text{externalities}} \quad \text{where } \tilde{\mu}_{Ag} \equiv \mu_{Ag} - (1-\gamma)\mu_L$$

- Empirical specification:

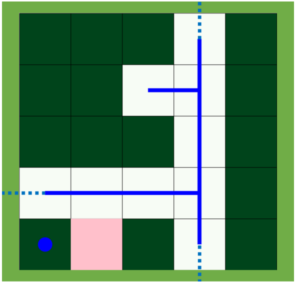
$$\ln \tilde{A}_{o,Ag} = \tilde{\mu}_{Ag}\theta \ln N_{o,Ag} + X_o'\beta + \phi_B + \epsilon_{o,Ag}$$

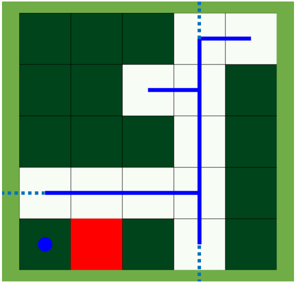
- "River Network Access" as an IV for $\ln N_{o,Ag}$:

$RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$ where RC: grid cells with rivers

Intuition of Identifying the Density Externalities

$RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$: River Network Access ($RC$: white cells)

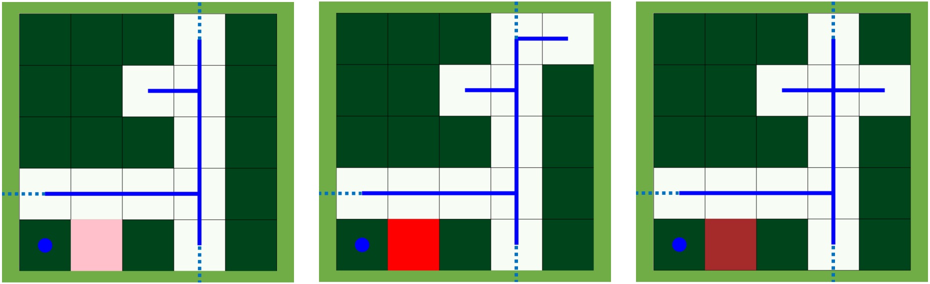




RNA at the pink cell $<$ RNA at the red cell

Intuition of Identifying the Density Externalities

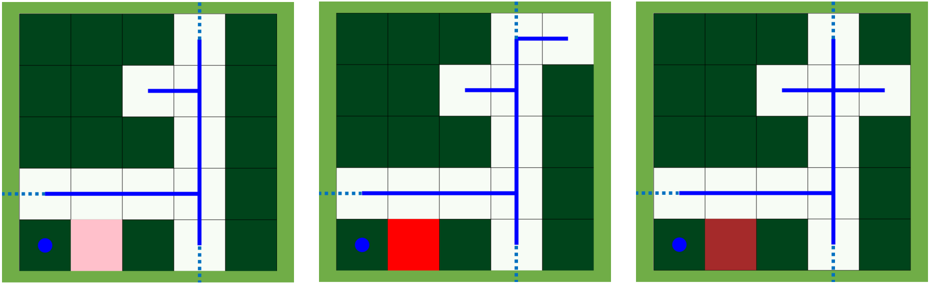
$RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$: River Network Access ($RC$: white cells)



RNA **at the pink cell** $<$ RNA **at the red cell** $<$ RNA **at the brown cell**

Intuition of Identifying the Density Externalities

$RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$: River Network Access ($RC$: white cells)

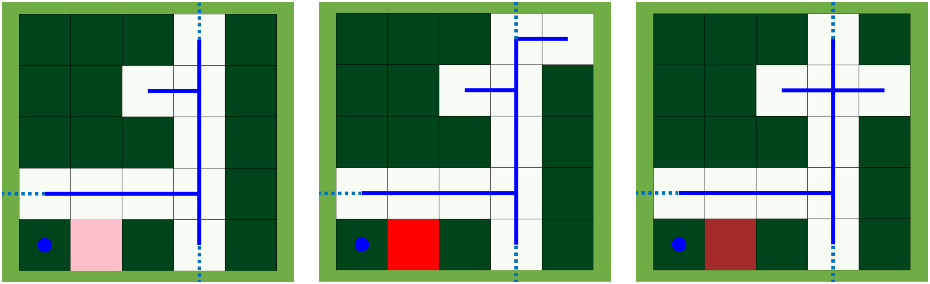


RNA **at the pink cell** $<$ RNA **at the red cell** $<$ RNA **at the brown cell**

- **Independence**: Given the same observable agricultural conditions (river proximity, water areas, soil conditions, etc), unobservable productivity fundamentals are uncorrelated with the variation in RNA *that stems from exogenous river shapes in locations far away from the own location*

Intuition of Identifying the Density Externalities

$RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$: River Network Access ($RC$: white cells)



RNA **at the pink cell** $<$ RNA **at the red cell** $<$ RNA **at the brown cell**

- **Independence**: Given the same observable agricultural conditions (river proximity, water areas, soil conditions, etc), unobservable productivity fundamentals are uncorrelated with the variation in RNA *that stems from exogenous river shapes in locations far away from the own location*
- **Exclusion**: RNA (as a market potential shifter) affects productivity only through its effect on employment and thus through externalities that arise

Step 3. Agglomeration Externality in Agriculture via Linear IV

- Inverted productivity composites of agriculture:

$$\tilde{A}_{o,Ag} \equiv \underbrace{A_{o,Ag} A_{o,L}^{(1-\gamma)\theta} \kappa_1^{\theta}}_{\text{fundamentals}} \cdot \underbrace{N_{o,Ag}^{\tilde{\mu}_{Ag}\theta}}_{\text{externalities}} \quad \text{where } \tilde{\mu}_{Ag} \equiv \mu_{Ag} - (1-\gamma)\mu_L$$

- Empirical specification:

$$\ln \tilde{A}_{o,Ag} = \tilde{\mu}_{Ag}\theta \ln N_{o,Ag} + X_o'\beta + \phi_B + \epsilon_{o,Ag}$$

- "River Network Access" as an IV for $\ln N_{o,Ag}$:

  $RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$ where RC: grid cells with rivers

- Identifying assumption:

  After controlling for ▸ geographic characteristics of the own location, productivity fundamentals are uncorrelated with accessibility to other locations

  ▸ Historical IV ▸ First stage ▸ Randomness

Step 3. Agglomeration Externality in Agriculture via Linear IV

- Inverted productivity composites of agriculture:

$$\tilde{A}_{o,Ag} \equiv \underbrace{A_{o,Ag} A_{o,L}^{(1-\gamma)\theta} \kappa_1^{\theta}}_{\text{fundamentals}} \cdot \underbrace{N_{o,Ag}^{\tilde{\mu}_{Ag}\theta}}_{\text{externalities}} \quad \text{where } \tilde{\mu}_{Ag} \equiv \mu_{Ag} - (1-\gamma)\mu_L$$

- Empirical specification:

$$\ln \tilde{A}_{o,Ag} = \tilde{\mu}_{Ag}\theta \ln N_{o,Ag} + X_o'\beta + \phi_B + \epsilon_{o,Ag}$$

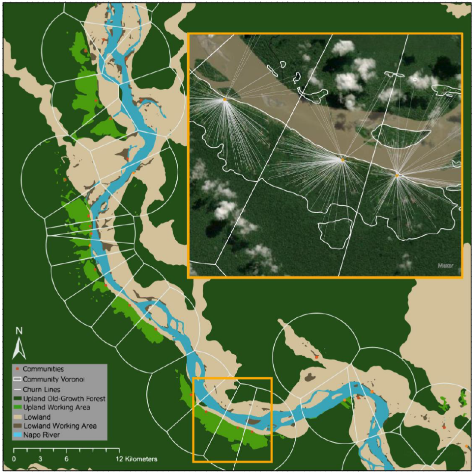
- "River Network Access" as an IV for $\ln N_{o,Ag}$:

  $RNA_o \equiv \sum_{d \in RC} (\tilde{\tau}_{od})^{-\theta}$ where RC: grid cells with rivers

- Identifying assumption:

  After controlling for ▸ geographic characteristics of the own location, productivity fundamentals are uncorrelated with accessibility to other locations
  ▸ Historical IV ▸ First stage ▸ Randomness

- A similar strategy for estimating μ_L due to common residual factors using community-level land footprint obtained from satellite images



Voronoi Polygons and Land Footprint around the Census Communities

Notes: To proxy community boundaries for agricultural land use, we partition land in the study area into voronoi polygons. Within each community voronoi polygon, we detect all patches of agricultural fields and secondary forests through satellite images. We then sum them up to calculate the land footprint of each community. See Coomes et al. (2021) for more details.

Empirical specification: $\ln \frac{L_o}{N_{o,Ag}} = -\mu_L \ln N_{o,Ag} + X_o'\beta + \phi_B + \epsilon_{o,L}$

Step 3. Congestion in Natural Resource Extraction with Spatial Spillovers via Non-Linear GMM

- Inverted productivity composites of natural resource extraction:

$$\tilde{A}_{o,Nr} \equiv \underbrace{A_{o,Nr}}_{\text{fundamentals}} \cdot \underbrace{[\sum_d D_{od}^{-\nu} N_{d,Nr}]^{-\mu_{Nr}\theta}}_{\text{externalities}}$$

- Moment conditions:

$$\mathbb{E}[\epsilon_{o,Nr} \ln RNA_o] = 0 \quad \text{and} \quad \mathbb{E}[\epsilon_{o,Nr} \ln(\sum_{d|D_{o,d} \leq x} RNA_d)] = 0, \ x \in \mathcal{X}$$

  - $\epsilon_{o,Nr}$: the residual variation in $\ln A_{o,Nr}$ (productivity fundamentals)
  - $\mathcal{X} = \{2, 5, 10, 25, 50, 75, 100\,(\text{km})\}$
  - Similar to the identification strategy by Ahlfeldt et al. (2015)

- Estimate ν & μ_{Nr} by the two-step nonlinear GMM

Across-sector externality?

Step 3. Density Externalities in Agriculture and Natural Resource Extraction

Parameter	Point estimate	Standard errror	Description
(A) Agriculture			
$\tilde{\mu}_{Ag}$	0.064	0.010	$= \mu_{Ag} - (1-\gamma)\mu_L$
	J test p-value $= 0.648$		
μ_L	0.522	0.094	Congestion in forest clearing
μ_{Ag}	0.273		Agglomeration in agricultural production
(B) Natural resource extraction			
μ_{Nr}	0.335	0.042	Congestion in natural resource extraction
ν	0.593	0.075	Spatial decay of congestion externality
	J test p-value $= 0.821$		

Notes: Estimates of density externalities in agriculture (panel A) are based on the linear specification using $\ln RNA_o$ and the initial community existence in 1940 as instruments. Estimates of parameters governing congestion externality in natural resource extraction (panel B) are based on the non-linear GMM using $\ln RNA_o$ and $\{\ln \sum_{d|D_{o,d} \leq x} RNA_d\}$ for $x \in \mathcal{X} = \{2, 5, 10, 25, 50, 75, 100\}$ as instruments.

- ⋆ **Agglomeration in agriculture > Congestion in access to land**

  Population ↑ ⇒ Productivity ↑ & Deforestation *per farmer* ↓

▸ Robustness to different IVs and samples ▸ Robustness to different controls

Quantitative Effects of the Agglomeration Externality

Counterfactual Outcomes by $\mu_{Ag} \to 0$ ▸ Maps

Basin	Welfare	Deforestation	Natural resource depletion
Napo	-11.2%	+55.5%	-0.6%
Pastaza	-8.9%	+24.1%	-0.7%
LowerUcayali	-12.7%	+21.4%	-2.8%
UpperUcayali	-7.1%	+17.7%	-1.2%

- Spatial reallocation of agriculture from concentration into dispersion
- Sectoral reallocation: Ag employment $\uparrow$ to feed the population given that consumption demands across Ag & Nr goods are complementary

Quantitative Effects of the Agglomeration Externality

Counterfactual Outcomes by $\mu_{Ag} \to 0$ ▸ Maps

Basin	Welfare	Deforestation	Natural resource depletion
Napo	-11.2%	+55.5%	-0.6%
Pastaza	-8.9%	+24.1%	-0.7%
LowerUcayali	-12.7%	+21.4%	-2.8%
UpperUcayali	-7.1%	+17.7%	-1.2%

- Spatial reallocation of agriculture from concentration into dispersion
- Sectoral reallocation: Ag employment $\uparrow$ to feed the population given that consumption demands across Ag & Nr goods are complementary
- $\mu_{Ag} \uparrow \Rightarrow$ Large effects on improving welfare and reducing deforestation but at the expense of natural resource depletion
- To improve both human and ecological well-being, need other policies that distribute the agglomeration benefits well

Mechanisms behind the Agglomeration Externality in Agriculture

1. Economies of scale in transport technology ▸ Details
   - Endogenous transport modes (different types of boats available)
   - Endogenous transaction costs
   - Trade costs decrease in origin populations
   - Isomorphic to the original model

2. Economies of scale in agricultural intensification ▸ Details
   - Direct inputs into land and crops (insecticides, herbicides, fungicides)
   - Complementary equipment (sprayers)
   - Crop processing technology to facilitate marketing (grain mill)

Counterfactuals

Overview

1. Combining ***well-targeted*** place-based protection policies and transport infrastructure simultaneously achieves:
    - Local populations' welfare ↑
    - Deforestation ↓
    - Natural resource depletion ↓

   (while any single policy cannot)

Overview

1. Combining ***well-targeted*** place-based protection policies and transport infrastructure simultaneously achieves:
   - Local populations' welfare $\uparrow$
   - Deforestation $\downarrow$
   - Natural resource depletion $\downarrow$

   (while any single policy cannot)

2. The *direction* of any environmental impact depends on *where* the place-based policy is implemented

Overview

1. Combining ***well-targeted*** place-based protection policies and transport infrastructure simultaneously achieves:
   - Local populations' welfare ↑
   - Deforestation ↓
   - Natural resource depletion ↓

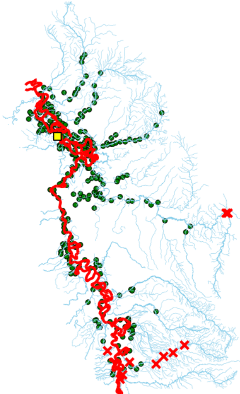
   (while any single policy cannot)

2. The *direction* of any environmental impact depends on *where* the place-based policy is implemented

Counterfactuals	**Welfare**	**Deforestation**	**Natural resource depletion**
(A) Protection policies			
i. Protecting the rural frontier	−	−	−
ii. Targeting the smallest communities	−	−	+
(B) River Transport infrastructure Details			
i. Connecting hinterlands to the center	+	−	?
ii. Concentrating in the center	+	+	?
(A) i. + (B) i.	**+**	**−**	**−**

(A) Protecting the rural frontier &
(B) Transport infrastructure that connects hinterlands to the center

River basin	Welfare	Deforestation	Natural resource depletion
Napo	+1.6%	-6.7%	-0.6%
Pastaza	+1.0%	-4.7%	-0.3%
LowerUcayali	+2.1%	-1.0%	-2.4%
UpperUcayali	+1.0%	-3.1%	-0.5%



The trade-offs are relaxed by:

(A) ×: Treated areas for resettlement

i.e. Shrink the outer edge within a basin where human settlement is advancing into undeveloped areas

&

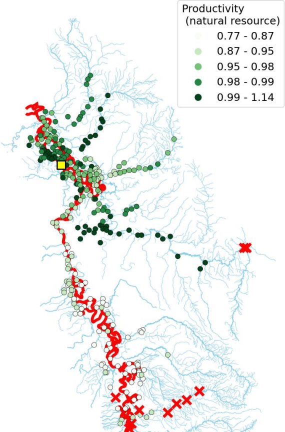
(B) **Red river lines**:

- No asymmetric transport costs
- -20% of the downstream-river-equivalent distance

(Robust across all four basins)

(A) Protecting the rural frontier &
(B) Transport infrastructure that connects hinterlands to the center

River basin	Welfare	Deforestation	**Natural resource depletion**
Napo	+1.6%	-6.7%	-0.6%
Pastaza	+1.0%	-4.7%	-0.3%
LowerUcayali	+2.1%	-1.0%	-2.4%
UpperUcayali	+1.0%	-3.1%	-0.5%



Intuition:

(A) More compact basin for human settlements

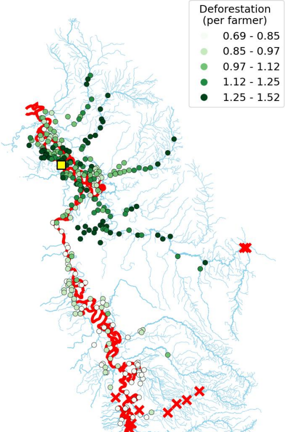
- Surrounding population density ↑ in most of populated areas
- Congestion externality with spatial spillovers ↑ & Productivity ↓ in most areas

⇒ Overall natural resource depletion ↓

Legend: values in the counterfactual relative to those in the benchmark equilibrium

(A) Protecting the rural frontier &
(B) Transport infrastructure that connects hinterlands to the center

River basin	**Welfare**	**Deforestation**	Natural resource depletion
Napo	+1.6%	-6.7%	-0.6%
Pastaza	+1.0%	-4.7%	-0.3%
LowerUcayali	+2.1%	-1.0%	-2.4%
UpperUcayali	+1.0%	-3.1%	-0.5%



Intuition:

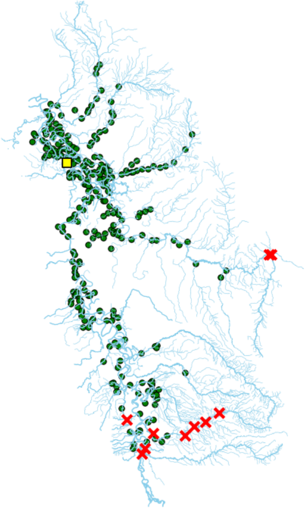
(B) Integrated between the center and hinterlands

- Total deforestation ↓ given the convex structure of congestion forces in access to land
- Reallocate farmers from the central to remote areas
- Settlement size variance ↓ & Agricultural productivity in remote areas ↑

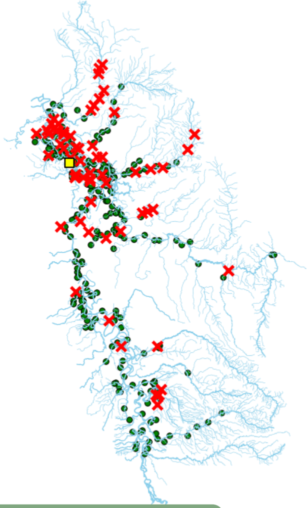
Agglomeration benefits spread more evenly across the basin with more medium-sized settlements

Spatial Targeting Matters: Protection Policies

(A) Protecting the rural frontier



(B) Small communities not allowed



×: Protected areas for resettlement

Community formation patterns

Spatial Targeting Matters: Protection Policies

Basin	Welfare	**Deforestation**	**Natural resource depletion**
(A) Protecting the rural frontier			
Napo	-0.2%	-5.2%	-0.3%
UpperUcayali	-0.2%	-2.0%	-0.8%
(B) Not allowing for small communities			
Napo	-0.3%	-13.1%	+0.2%
UpperUcayali	-0.1%	-7.3%	+0.5%

- Both policies treat 2.5% of the total rural population in the basin

(B) $\Rightarrow$ **natural resource depletion** $\uparrow$

- Protected areas are dispersed
- Congestion from surrounding populations not much affected

Spatial Targeting Matters: Protection Policies

Basin	Welfare	**Deforestation**	**Natural resource depletion**
(A) Protecting the rural frontier			
Napo	-0.2%	-5.2%	-0.3%
UpperUcayali	-0.2%	-2.0%	-0.8%
(B) Not allowing for small communities			
Napo	-0.3%	-13.1%	+0.2%
UpperUcayali	-0.1%	-7.3%	+0.5%

- Both policies treat 2.5% of the total rural population in the basin

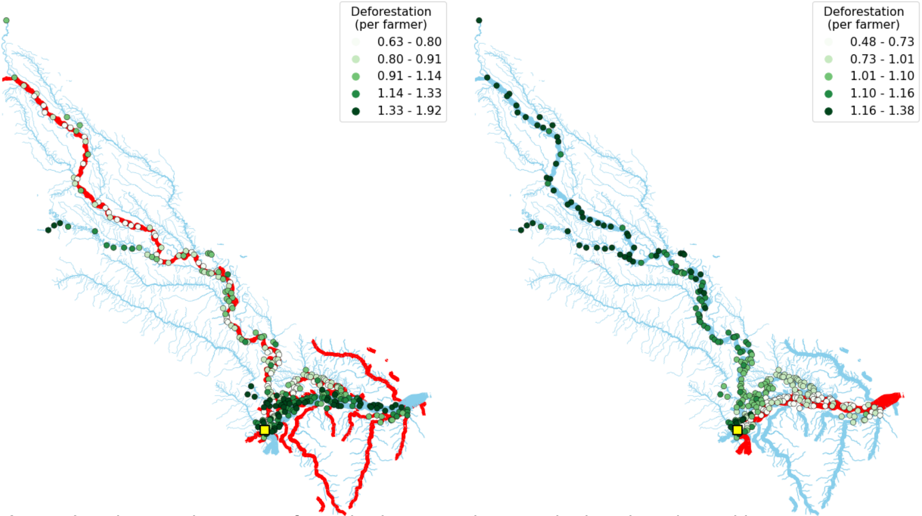
(B) $\Rightarrow$ **natural resource depletion ↑**

  - Protected areas are dispersed
  - Congestion from surrounding populations not much affected

(B) $\Rightarrow$ **deforestation ↓ more** by concentrating the land footprint in much fewer spots (rather than having many small communities)

★ Policymaker's trade-off between mitigating different environmental costs

Spatial Targeting Matters: River Transport Infrastructure



Legend: values in the counterfactual relative to those in the benchmark equilibrium
Red lines: transport infrastructure improvement

Spatial Targeting Matters: River Transport Infrastructure

Basin	Welfare	Deforestation	Natural resource depletion
(A) Connecting hinterlands to the center			
Napo	+1.8%	**-1.1%**	-0.3%
(B) Concentrating improvements in the center			
Napo	+1.0%	**+6.0%**	-0.4%

Improving the infrastructure only in densely populated areas $\Rightarrow$ Deforestation $\uparrow$

- More concentration in the central area of the basin
- Smaller communities with lower agricultural productivity and higher deforestation *per farmer* in the hinterland

Conclusions

RQ. *How can we design policies that improve both tropical forest conservation and local populations' welfare?*

- This question turns on the direction and magnitude of density externalities in agricultural production, forest clearing, and natural resource extraction
- Combination of a protection policy and transport infrastructure investments
- Spatial targeting matters for both
- Policies that spread the agglomeration benefits more evenly across space and consolidate congestion forces into a more compact space are desirable to the environment

Appendix

Static vs. Dynamic Variations in Forest Cover

- Shifting cultivation with the swidden-fallow cycle:
  1. Clear primary (old-growth) forests, burn the vegetation to obtain land plots, and plant crops
  2. When plots become no longer productive, plots are left in fallow and the secondary forest regrows
  3. After several years of fallow, farmers clear such secondary forests again ($\rightarrow$ 1.)

Static vs. Dynamic Variations in Forest Cover

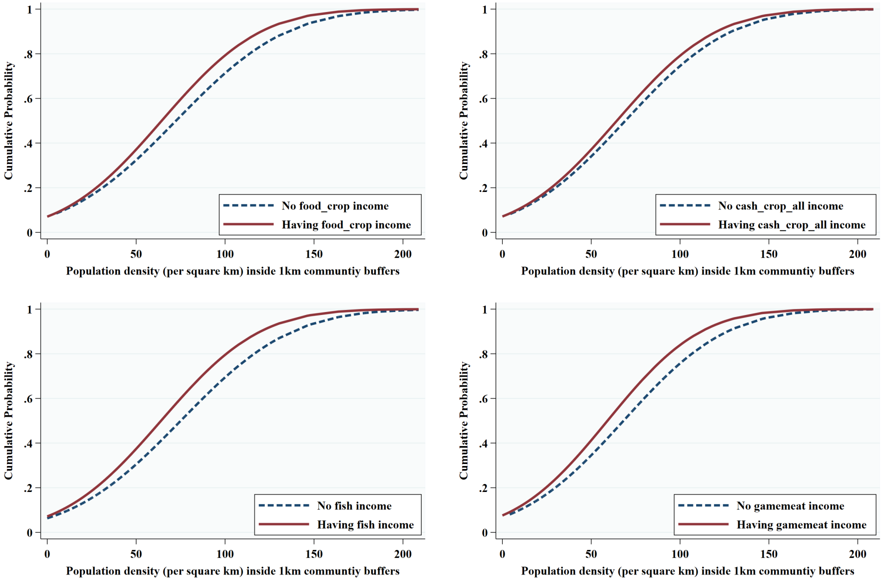
- Shifting cultivation with the swidden-fallow cycle:
  1. Clear primary (old-growth) forests, burn the vegetation to obtain land plots, and plant crops
  2. When plots become no longer productive, plots are left in fallow and the secondary forest regrows
  3. After several years of fallow, farmers clear such secondary forests again ($\rightarrow$ 1.)
- Although the deforested locations around the community are moving over time due to this cycle, at any given moment the stock of forest fallow and the total deforested area around the community remain relatively constant (Coomes et al. 2021)

$\Rightarrow$ The cross-sectional relationship between the settlement size and deforestation has a more significant variation

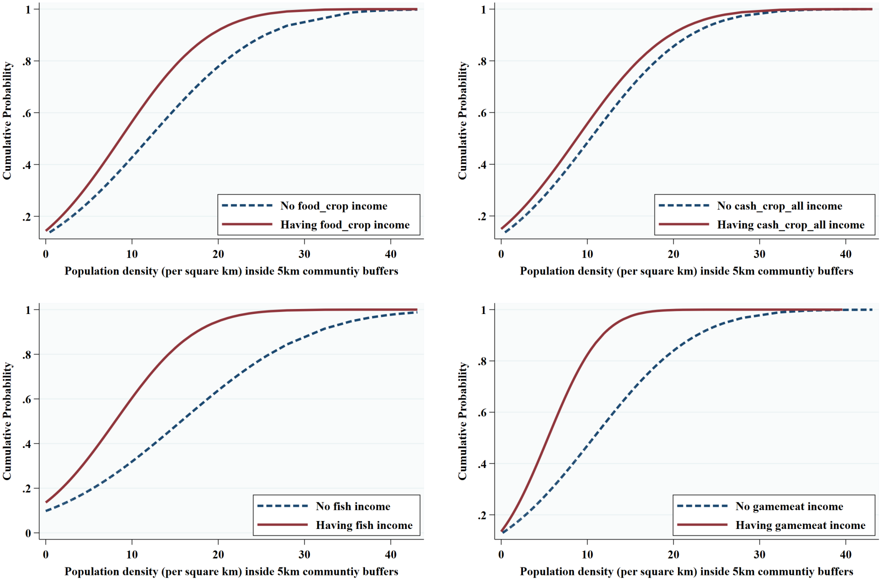
$\Rightarrow$ A static theoretical model

Back

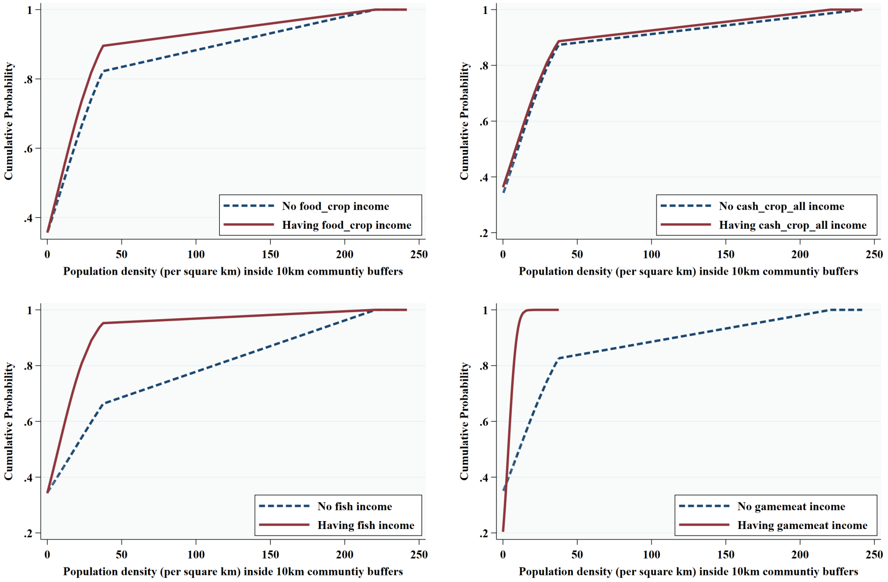
Population Density ($1km$ buffer) & Activity/Occupation: Agricultural Sector vs. Wild Animal Extractions



Population Density (*5km* buffer) & Activity/Occupation: Agricultural Sector vs. Wild Animal Extractions



Population Density (*10km* buffer) & Activity/Occupation: Agricultural Sector vs. Wild Animal Extractions



Spatial Distribution of Agriculture and Natural Resource Extraction

- Population density in a smaller buffer of a community becomes close to the population density of that community itself
- A smaller population density in a large buffer of a community implies more sparseness of communities and less competition over natural resources
- **Spatial distributions of natural resource extractions become more contrastable from agricultural activities as the buffer size enlarges** ($x = 5km, 10km$ compared to $x = 1km, 2km$)

Back

Step 1. Obtaining Parameters without Solving the Model

Parameter	Description	Estimation strategy	Value
δ_{Ag}	Elasticity of trade cost	Commodity prices from the CC	0.178
δ_{Nr}	Elasticity of trade cost	Commodity prices from the CC	0.137
δ_{M}	Elasticity of trade cost	Commodity prices from the CC	0.098
λ_{up}	Relative upstream-river travel cost	Travel time and transport costs survey	1.282
λ_{land}	Relative land travel cost	Travel time and transport costs survey	36.767
σ	Within-sector elasticity of substitution	Expenditure information from ENAHO	2.401
$\bar{\sigma}$	Across-sector elasticity of substitution	Expenditure information from ENAHO	0.752
γ	Labor share in agricultural production	From the literature	0.6
θ	Trade elasticity	From the literature	7.8

Back

Step 2. Model Inversion

- Observable data: sectoral populations ($\{N_{o,Ag}\}$, $\{N_{o,Nr}\}$, $\{N_{u,M}\}$)
- Use the $2|\mathcal{R}| + 1 + |\mathcal{I}|$ equations from the spatial equilibrium conditions (sectoral labor market clearing + utility equalization across space)

  with the observables

  to solve for $2|\mathcal{R}| + 1 + |\mathcal{I}|$ unknowns (productivity composites + wages): $\{\tilde{A}_{o,Ag}\}$, $\{\tilde{A}_{o,Nr}\}$, $A_{u,M}$, $\{w_o\}$

  $\rightarrow$ Use the inverted productivity composites as data in the next step

Back

Geographic Controls

- River cell dummy
- Distance to the river point and its square
- Interaction between the above two
- River confluences
- Elevation
- Flood experience
- Geology measures
- Water (main and non-main) areas
- Distance to the urban center

Back

Historical IV

IV: Community existence in its current location by 1940

- The primary reason for early settlement was the opportunity to extract natural resource products
- The Amazon Rubber Boom:
  - Began the late 19th century, but collapsed around 1940
  - Significantly impacted initial settlements (Barham et al. 1996; Coomes 1995)

⇒ The locations of communities established before 1940 were likely to be determined primarily by natural resource endowments, not by advantages in agricultural productivity

Back

Table: River Networks, Initial Communities, and Current Populations

	$\log(N_{o,Ag})$			Community existence (1940)	
	(1)	(2)	(3)	(4)	(5)
$\log(RNA_o)$	0.758***		0.711***	-0.0145	0.0699
	(0.223)		(0.218)	(0.0254)	(0.0726)
Community existence (1940)		0.740***	0.730***		
		(0.0983)	(0.0980)		
Basin FE	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	No	Yes
Mean (Dep. var.)	4.322	4.322	4.322	0.194	0.194
SD (Dep. var.)	1.192	1.192	1.192	0.395	0.395
R^2	0.154	0.195	0.206	0.094	0.117
Observations	893	893	893	904	899

Notes: Robust standard errors in parentheses. The sample includes 1 square km grid cells that have positive populations. Geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluences, flood vulnerability, geology measures, and open water access measures.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

(A) River Network Access	(1) Water share: non-main channel	(2) River confluence: 1st×2nd or 2nd×3rd	(3) River confluence: 3rd×4th	(4) Flood vulnerability	(5) Pleistocene soil share	(6) Tertiary soil share
$\log(RNA_o)$	0.00751	0.0470	-0.0743	-0.217	-0.0444	-0.0498
	(0.0161)	(0.0470)	(0.0618)	(0.307)	(0.0362)	(0.0474)
Mean (Dep. var.)	0.030	0.077	0.083	1.606	0.021	0.211
SD (Dep. var.)	0.087	0.266	0.277	1.606	0.115	0.344
R^2	0.068	0.095	0.137	0.130	0.057	0.735
Observations	899	899	899	899	899	899
(B) Early human settlements	(1) Water share: main channel	(2) Water share: non-main channel	(3) Flood vulnerability	(4) Floodplain soil share	(5) Pleistocene soil share	(6) Tertiary soil share
Community existence (1940)	0.0263	-0.00352	0.218	0.00191	0.00777	-0.0108
	(0.0193)	(0.00698)	(0.142)	(0.0268)	(0.0122)	(0.0257)
Mean (Dep. var.)	0.109	0.030	1.606	0.584	0.021	0.211
SD (Dep. var.)	0.203	0.087	1.606	0.359	0.115	0.344
R^2	0.162	0.038	0.116	0.250	0.030	0.243
Observations	899	899	899	899	899	899
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. The sample includes 1 square km grid cells that have positive populations. In panel (A), geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, water share of main channel rivers, and floodplain soil share. In panel (B), geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, and river confluences.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

▸ Back

Across-Sector Externality?

- There may exist the across-sector externality——the effect of clearing forests for agriculture on the productivity of natural resource extraction
- We are not incorporating it. 3 comments:

1. The spatial extent of these sectors' activities is distinct:
   - Deforestation for agricultural land is distributed along the rivers:
     mean, median, max (land footprint depths) = 1 km, 0.85km, 5.5 km
   - Natural resources are also extracted in deep inland areas away from the river
2. Natural resource endowments are not significantly correlated with the community-level land footprint Table
3. This model choice does not affect the inversion problem, but affects outcomes in counterfactual policy simulations:
   - We investigate policies that reduce total deforestation in a river basin
   - We can interpret these policies' welfare effects as lower bounds

Back (Model) Back (Estimation)

Table: Natural Resource Endowments, Calibrated Productivity, and Community Land Footprint

(A)	Number of species found around a community				
	(1)	(2)	(3)	(4)	(5)
	Total	Fish	Timber	NTFP	Game
$\log(\tilde{A}_{o,Nr})$ (calibrated)	0.206***	0.0220	0.386***	0.0488**	0.380***
	(0.0306)	(0.0334)	(0.0407)	(0.0204)	(0.0437)
Mean (Dep. var.)	2.025	3.161	1.788	0.552	1.958
SD (Dep. var.)	1.145	1.163	1.676	0.893	1.636
R^2	0.059	0.150	0.126	0.349	0.213
Observations	909	909	909	909	909
(B)	Number of species found around a community				
	(1)	(2)	(3)	(4)	(5)
	Total	Fish	Timber	NTFP	Game
log (land footprint)	0.0171	-0.0653*	-0.0209	0.0126	0.0752
	(0.0381)	(0.0383)	(0.0533)	(0.0245)	(0.0478)
Mean (Dep. var.)	2.021	3.147	1.796	0.555	1.956
SD (Dep. var.)	1.147	1.183	1.677	0.894	1.637
R^2	0.014	0.163	0.059	0.336	0.146
Observations	906	906	906	906	906
(C)	Number of species found around a community				
	(1)	(2)	(3)	(4)	(5)
	Total	Fish	Timber	NTFP	Game
log (depth of land footprint)	0.0458	-0.0358	-0.135*	-0.0248	0.0911
	(0.0587)	(0.0541)	(0.0771)	(0.0342)	(0.0728)
Mean (Dep. var.)	2.070	3.168	1.873	0.550	1.964
SD (Dep. var.)	1.141	1.159	1.676	0.885	1.626
R^2	0.026	0.170	0.059	0.364	0.146
Observations	811	811	811	811	811
Basin FE	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	No

Notes: Robust standard errors in parentheses. The unit of anaylysis is a community in the PARLAP Community Census (CC) in 2014. In panel (B), the land footprint represents the community-level land footprint within a voronoi polygon around the settlement, detected in satellite images. In panel (C), the land footprint depth represents the distance from the river to the furthest inland point in the community-level land footprint.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table: Agglomeration Externality in Agriculture

	The calibrated value of $\log(\tilde{A}_{o,Ag})$							
	All locations				$N_o < 1000$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\log(N_{o,Ag})$	0.676***	0.440**	0.514***	0.501***	0.735***	0.384**	0.509***	0.464***
	(0.0207)	(0.171)	(0.0809)	(0.0790)	(0.0196)	(0.169)	(0.124)	(0.109)
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV: RNA	No	Yes	No	Yes	No	Yes	No	Yes
IV: Historical	No	No	Yes	Yes	No	No	Yes	Yes
Mean (Dep. var.)	-0.096	-0.096	-0.096	-0.096	-0.172	-0.172	-0.172	-0.172
SD (Dep. var.)	4.578	4.578	4.578	4.578	4.614	4.614	4.614	4.614
First stage *F*-stat		11.502	56.653	31.005		15.298	35.632	22.822
Hansen's *J* test *p*-value				0.648				0.472
Observations	893	893	893	893	852	852	852	852

Notes: Robust standard errors in parentheses. The sample includes 1 square km grid cells that have positive populations. We use $\log(RNA_o)$ (IV: RNA) and the initial community existence in 1940 (IV: Historical) as instruments for $\log(N_{o,Ag})$. Geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluences, flood vulnerability, geology measures, and open water access measures.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

Table: Agglomeration Externality in Agriculture

	The calibrated value of $\log(\tilde{A}_{o,Ag})$						
	IV						OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\log(N_{o,Ag})$	0.434***	0.519***	0.519***	0.521***	0.509***	0.501***	0.676***
	(0.0920)	(0.0789)	(0.0788)	(0.0789)	(0.0789)	(0.0790)	(0.0207)
log (Elevation)		2.341***	2.354***	2.324***	2.360***	2.397***	2.252***
		(0.171)	(0.175)	(0.179)	(0.176)	(0.177)	(0.176)
River confluence (1st×2nd or 2nd×3rd)			0.0155	0.0186	0.0206	0.0309	0.0180
			(0.0958)	(0.0964)	(0.0982)	(0.0995)	(0.0969)
River confluence (3rd×4th)			-0.0356	-0.0339	-0.0246	-0.0266	0.0173
			(0.0724)	(0.0723)	(0.0730)	(0.0733)	(0.0618)
Flood vulnerability (1-4)				-0.0115	-0.00947	-0.0123	-0.0154
				(0.0136)	(0.0137)	(0.0137)	(0.0130)
Water share: non-main channel					0.0806	0.123	-0.00146
					(0.238)	(0.238)	(0.203)
Water share: main channel					0.161	0.185	0.189
					(0.122)	(0.121)	(0.120)
Floodplain soil share						0.127**	0.126**
						(0.0625)	(0.0575)
Pleistocene soil share						0.175	0.333
						(0.222)	(0.227)
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean (Dep. var.)	-0.094	-0.096	-0.096	-0.096	-0.096	-0.096	-0.096
SD (Dep. var.)	4.576	4.578	4.578	4.578	4.578	4.578	4.578
First stage *F*-stat	28.030	29.419	29.974	29.634	30.770	31.005	
Observations	894	893	893	893	893	893	893

Notes: Robust standard errors in parentheses. The sample includes 1 square km grid cells that have positive populations. We use $\log(RNA_o)$ and the initial community existence in 1940 as instruments for $\log(N_{o,Ag})$. Other controls include distance to the urban center, distance to the river, squared distance to the river, and interaction terms of these two variables with a river cell dummy.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

Table: Density Externality in Forest Clearing

	log (per capita land footprint)					
	All locations		$N_o < 1000$		$N_o < 500$	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
$\log(N_{o,Ag})$	-0.650***	-0.522***	-0.654***	-0.552***	-0.674***	-0.545***
	(0.0307)	(0.0940)	(0.0323)	(0.109)	(0.0346)	(0.123)
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean (Dep. var.)	0.929	0.929	0.956	0.956	0.981	0.981
SD (Dep. var.)	1.231	1.231	1.218	1.218	1.223	1.223
First stage *F*-stat		34.198		28.141		23.709
Hansen's *J* test *p*-value		0.987		0.896		0.969
Observations	895	895	878	878	847	847

Notes: Robust standard errors in parentheses. The unit of anaylysis is a community in the PARLAP Community Census (CC) in 2014. We use $\log(RNA_o)$ and the initial community existence in 1940 as instruments for $\log(N_{o,Ag})$. Geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluence, flood vulnerability, geology measures, and open water access measures for a grid cell where each census community belongs.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

Table: Congestion Externality in Natural Resource Extraction with Spatial Spillovers

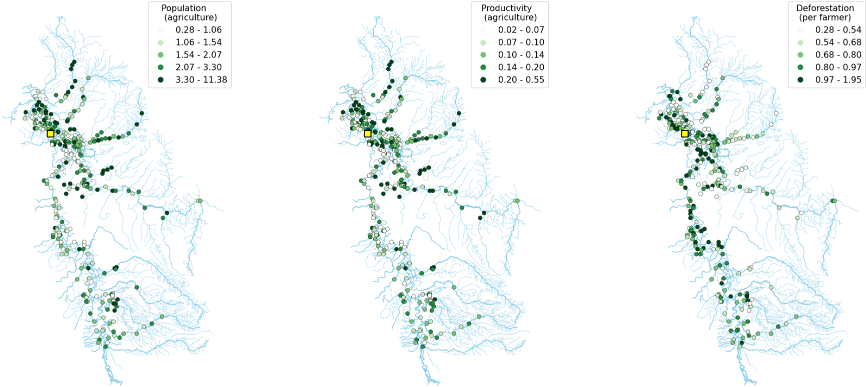
	The calibrated value of $\log(\tilde{A}_{o,Nr})$									
	IV									OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\log(N_{o,Nr})$	-2.127**	-1.385	-0.879	-0.581	0.0960	0.278	0.573***	0.634***	0.606***	0.950***
	(1.075)	(0.939)	(0.688)	(0.558)	(0.322)	(0.280)	(0.208)	(0.179)	(0.184)	(0.0516)
$\log(\sum_{d\|D_{o,d}\leq 2km} N_{d,Nr})$		-0.573*	-0.0343	0.0235	-0.0745	-0.0705	-0.0648	-0.0745	-0.0856	-0.0663
		(0.331)	(0.282)	(0.236)	(0.143)	(0.122)	(0.0883)	(0.0822)	(0.0838)	(0.0611)
$\log(\sum_{d\|D_{o,d}\leq 5km} N_{d,Nr})$			-0.596***	-0.286	-0.138	-0.130	-0.120*	-0.111*	-0.112*	-0.132***
			(0.189)	(0.183)	(0.106)	(0.0888)	(0.0637)	(0.0597)	(0.0613)	(0.0425)
$\log(\sum_{d\|D_{o,d}\leq 10km} N_{d,Nr})$				-0.337**	0.0345	0.0364	0.0579	0.0425	0.0322	0.0140
				(0.141)	(0.107)	(0.0885)	(0.0625)	(0.0596)	(0.0639)	(0.0378)
$\log(\sum_{d\|D_{o,d}\leq 25km} N_{d,Nr})$					-0.470***	-0.357***	-0.327***	-0.294***	-0.285***	-0.165***
					(0.0918)	(0.0837)	(0.0584)	(0.0560)	(0.0571)	(0.0283)
$\log(\sum_{d\|D_{o,d}\leq 50km} N_{d,Nr})$						-0.195***	-0.0318	-0.0548	-0.0480	-0.0619**
						(0.0610)	(0.0581)	(0.0526)	(0.0548)	(0.0242)
$\log(\sum_{d\|D_{o,d}\leq 75km} N_{d,Nr})$							-0.280***	-0.0758	-0.0407	-0.0989***
							(0.0779)	(0.125)	(0.142)	(0.0352)
$\log(\sum_{d\|D_{o,d}\leq 100km} N_{d,Nr})$								-0.258*	-0.439*	-0.263***
								(0.141)	(0.231)	(0.0498)
$\log(\sum_{d\|D_{o,d}\leq 150km} N_{d,Nr})$									0.187	0.0970*
									(0.171)	(0.0567)
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean (Dep. Var.)	0.337	0.337	0.337	0.337	0.337	0.337	0.337	0.337	0.337	0.337
SD (Dep. Var.)	2.862	2.862	2.862	2.862	2.862	2.862	2.862	2.862	2.862	2.862
Observations	894	894	894	894	894	894	894	894	894	894

Notes: Robust standard errors in parentheses. The sample includes 1 square km grid cells that have positive populations. We use $\ln RNA_o$ and $\{\ln \sum_{d|D_{o,d}\leq x} RNA_d\}$ for $x \in \mathcal{X}$ as instruments when endogenous variables include $\log(N_{o,Nr})$ and $\{\ln \sum_{d|D_{o,d}\leq x} N_{d,Nr}\}$ for $x \in \mathcal{X}$. Geographical controls include a dummy of high river orders (4 and 5), distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluence, flood vulnerability, geology measures, and open water access measures.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

Spatial Distribution of Agriculture and Deforestation without the Agglomeration Externality



Back

Economies of Scale in Transport Technology

- Consider an alternative model without the agglomeration externality in the production function but with **endogenous trade costs**:

$$\tilde{\tau}_{od,Ag} \quad = \quad N_{o,Ag}^{-\mu_{Ag}} \tau_{od,Ag}$$

- This model is **isomorphic to the original model**
- The trade cost can be decreasing in the origin population possibly because:
  - Large commercial boats (‘lancha’) are more likely to stop by
  - Collective investment in motor boats (‘rapido’)
  - The average transport cost charged is decreasing in the amount of products traded

Transport Modes in the Peruvian Amazon

Canue

Lancha

Peque-peque (most widely available)

Rapido (express motor boat)

Table: Community Population and Availability of Transport Modes

	Availability of Transport Modes in a Community							
	Lancha		Colectivo		Rapido		Peque-peque	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
$\log(N_{o,Ag})$	0.0469***	0.144***	0.0478***	0.0280	0.0522***	0.0566*	-0.00528	0.00418
	(0.0111)	(0.0430)	(0.0115)	(0.0383)	(0.0108)	(0.0292)	(0.00576)	(0.0156)
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean (Dep. var.)	0.492	0.492	0.386	0.386	0.110	0.110	0.972	0.972
SD (Dep. var.)	0.500	0.500	0.487	0.487	0.314	0.314	0.164	0.164
First stage *F*-stat		24.84462		24.84462		24.84462		24.84462
Observations	906	906	906	906	906	906	906	906

Notes: Robust standard errors in parentheses. The unit of anaylysis is a community in the PARLAP Community Census (CC) in 2014. We use $\log(RNA_o)$ and the initial community existence in 1940 as instruments for $\log(N_{o,Ag})$. Geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluence, flood vulnerability, geology measures, and open water access measures for a grid cell where each census community belongs.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

- The data supports the fact that ‘peque-peque’ is most widely available
- Significant scale effects on the availability of ‘lancha’ and ‘rapido’
- Consistent results for the frequency of transport modes available and other proxies of transaction costs as well

Back

Economies of Scale in Agricultural Intensification

- Test this using producer-level information from the Agricultural Census
- Modern technologies are limited: each of 24 listed modern technologies (except for boat) is used by <10% of agricultural producers
- Significant scale effects on:
    - direct inputs into land and crops (fertilizers, insecticides, herbicides, fungicides)
    - complementary equipment (sprayers)
    - crop processing technology to facilitate marketing (grain mill)

Table: Community Population and Modern Technology Use

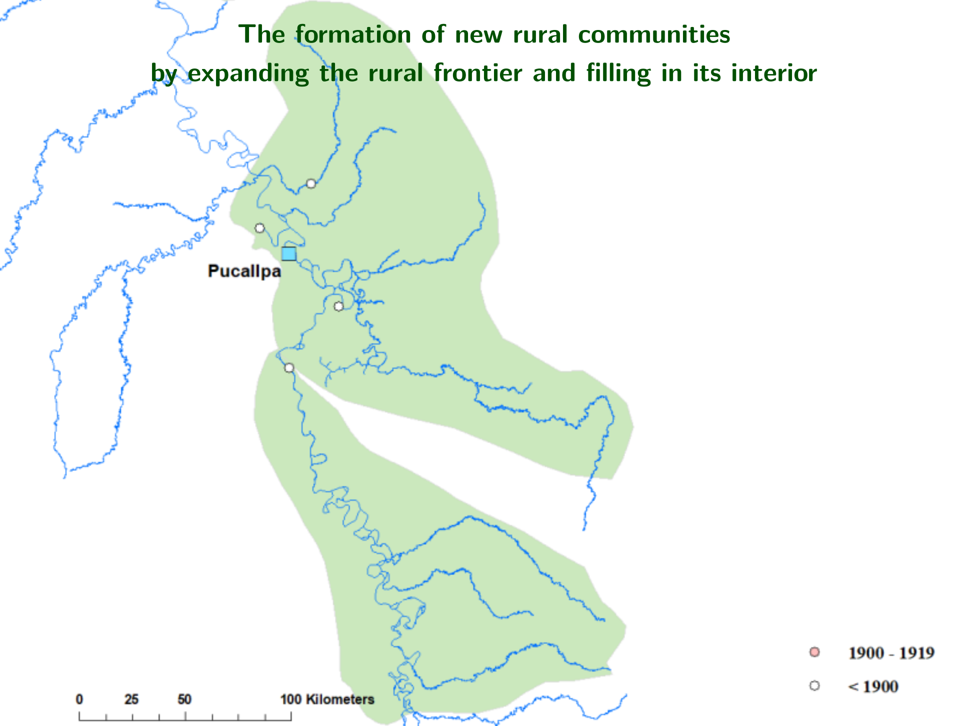
(A) Basic infrastructure	(1)	(2)	(3)	(4)	(5)	(6)
	Irrigation	Certified seed	Crops have been certified organic	Electricity for agricultural work	Animals for agricultural work	Tractors for agricultural work
$\log(N_{o,Ag})$	-0.00329*	-0.000857	0.0000692	-0.000688	0.00315	0.000476
	(0.00180)	(0.00430)	(0.000584)	(0.000863)	(0.00206)	(0.000811)
Mean (Dep. var.)	0.013	0.064	0.001	0.003	0.010	0.002
SD (Dep. var.)	0.112	0.245	0.037	0.054	0.098	0.044
First stage F-stat	1649.082	1649.082	1649.082	1649.082	1649.082	1649.082
Observations	25827	25827	25827	25827	25827	25827
(B) Inputs into land and crops	(1)	(2)	(3)	(4)	(5)	(6)
	Guano/manure/ compost	Chemical fertilizers	Insecticides	Herbicides	Fungicides	Biologic control
$\log(N_{o,Ag})$	0.000807	0.00265**	0.0228***	0.0314***	0.0118***	-0.00239
	(0.00111)	(0.00115)	(0.00353)	(0.00371)	(0.00219)	(0.00239)
Mean (Dep. var.)	0.005	0.004	0.040	0.051	0.012	0.020
SD (Dep. var.)	0.069	0.063	0.197	0.221	0.111	0.140
First stage F-stat	1649.082	1649.082	1649.082	1649.082	1649.082	1649.082
Observations	25827	25827	25827	25827	25827	25827
(C) Animal, electrical, or mechanical energy	(1)	(2)	(3)	(4)	(5)	(6)
	Iron plow of animal traction	Wooden plow of animal traction	Harvester	Foot plow	Motorized sprayer	Manual sprayer
$\log(N_{o,Ag})$	-0.000796	-0.000223	-0.000229	-0.000806	0.00197**	0.0214***
	(0.000523)	(0.000311)	(0.000282)	(0.000556)	(0.000815)	(0.00401)
Mean (Dep. var.)	0.001	0.000	0.001	0.001	0.002	0.062
SD (Dep. var.)	0.035	0.022	0.025	0.035	0.043	0.241
First stage F-stat	1649.082	1649.082	1649.082	1649.082	1649.082	1649.082
Observations	25827	25827	25827	25827	25827	25827
(D) Electrical or mechanical energy	(1)	(2)	(3)	(4)	(5)	(6)
	Grain mill	Grass chopper	Thresher	Electric generator	Wheel tractor	Boat/canue/ speedboat
$\log(N_{o,Ag})$	0.00696***	0.000462	-0.00102	-0.0103***	0.000932	-0.0187**
	(0.00194)	(0.000448)	(0.000674)	(0.00323)	(0.000590)	(0.00746)
Mean (Dep. var.)	0.013	0.001	0.004	0.036	0.001	0.618
SD (Dep. var.)	0.111	0.025	0.061	0.186	0.030	0.486
First stage F-stat	1649.082	1649.082	1649.082	1649.082	1649.082	1649.082
Observations	25827	25827	25827	25827	25827	25827
Basin FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. The unit of anaylysis is a household in the 2012 Peruvian Agricultural Census. We use $\log(RNA_o)$ and the initial community existence in 1940 as instruments for $\log(N_{o,Ag})$. Geographical controls include a dummy of high river orders (4 and 5), distance to the urban center, distance to the river, squared distance to the river, interaction terms of these two variables with a river cell dummy, elevation, river confluence, flood vulnerability, geology measures, and open water access measures for a grid cell where each census community belongs.

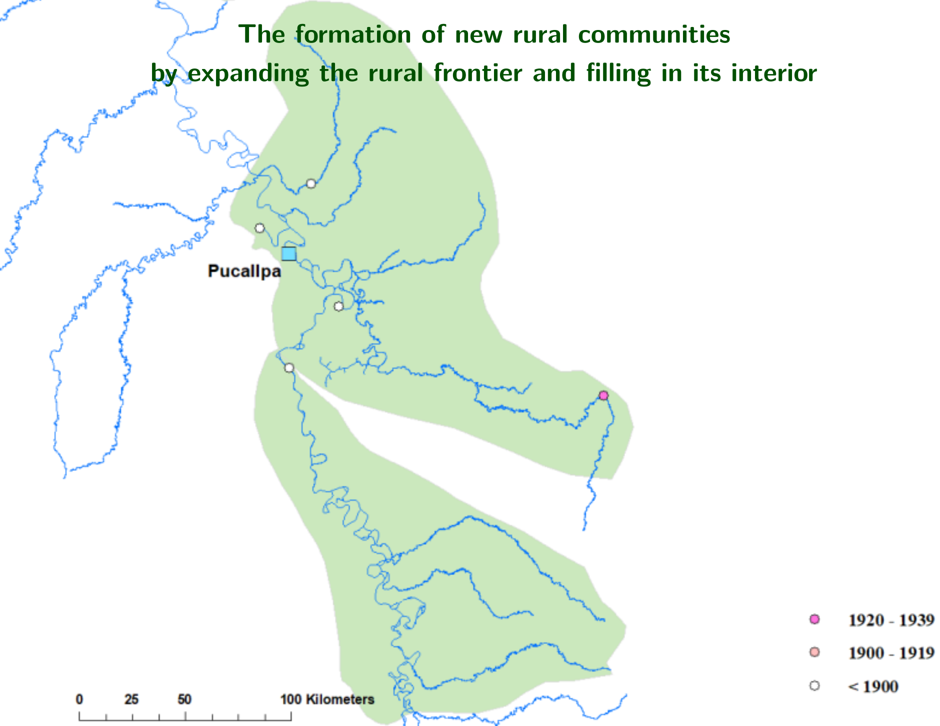
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Back

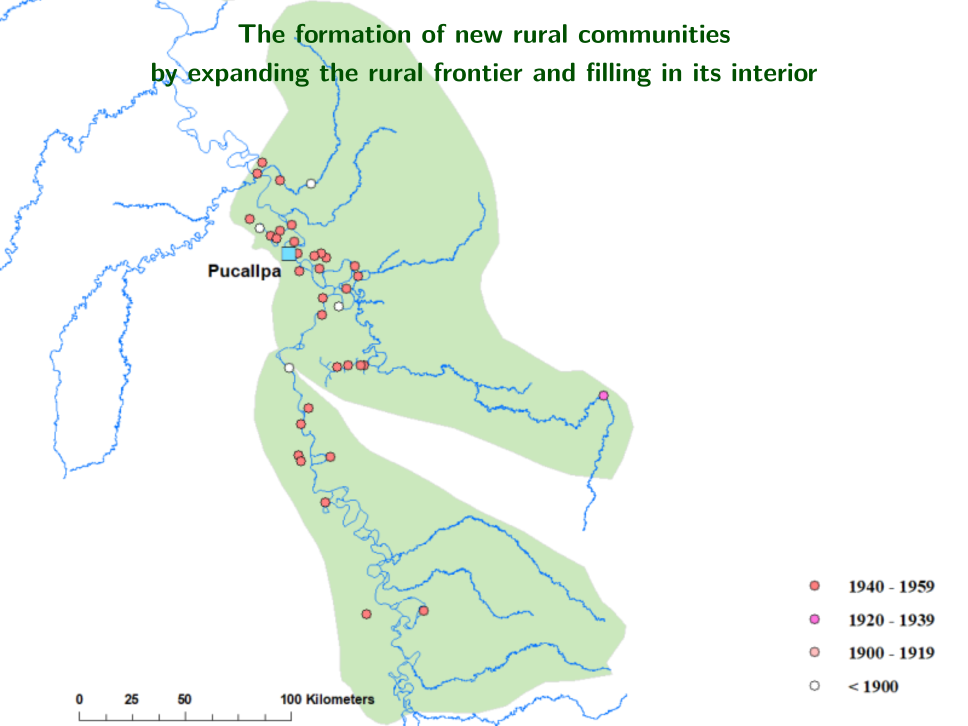
The formation of new rural communities by expanding the rural frontier and filling in its interior



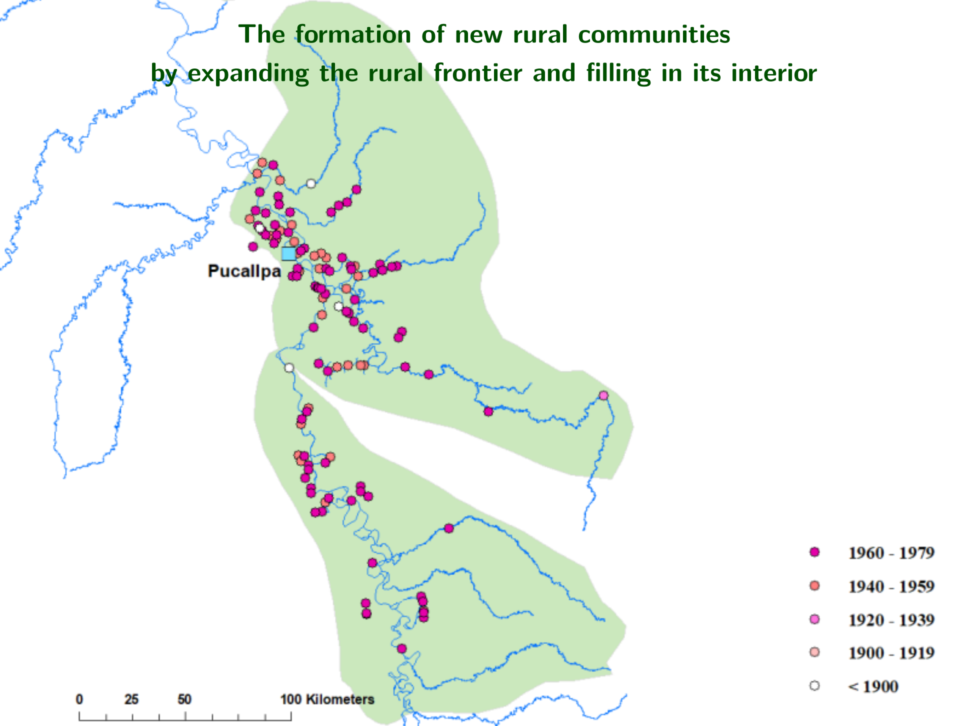
The formation of new rural communities by expanding the rural frontier and filling in its interior



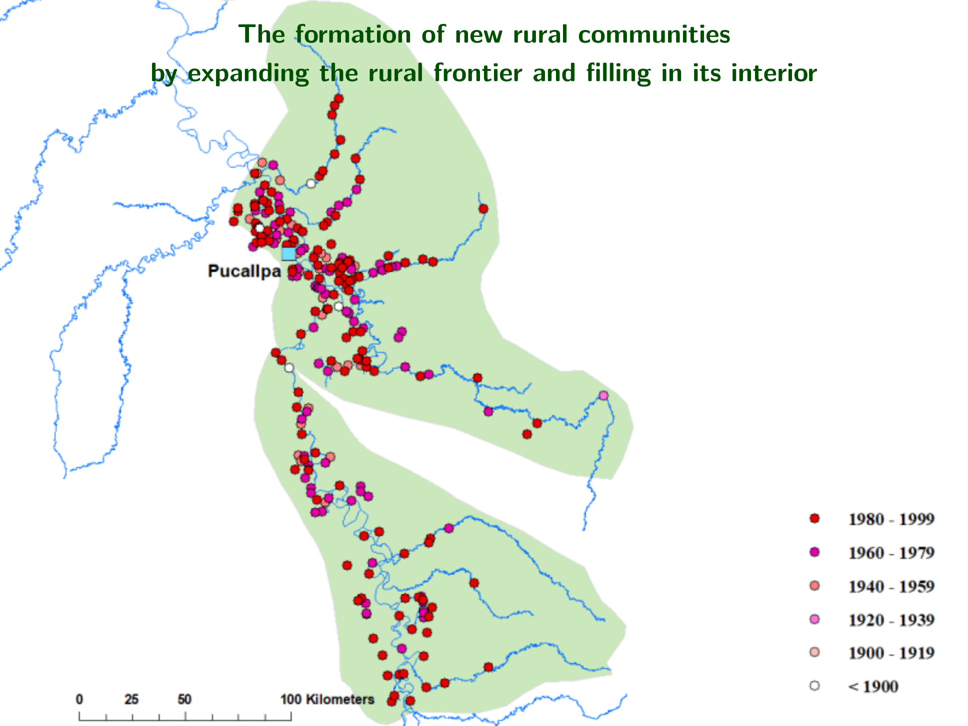
The formation of new rural communities by expanding the rural frontier and filling in its interior



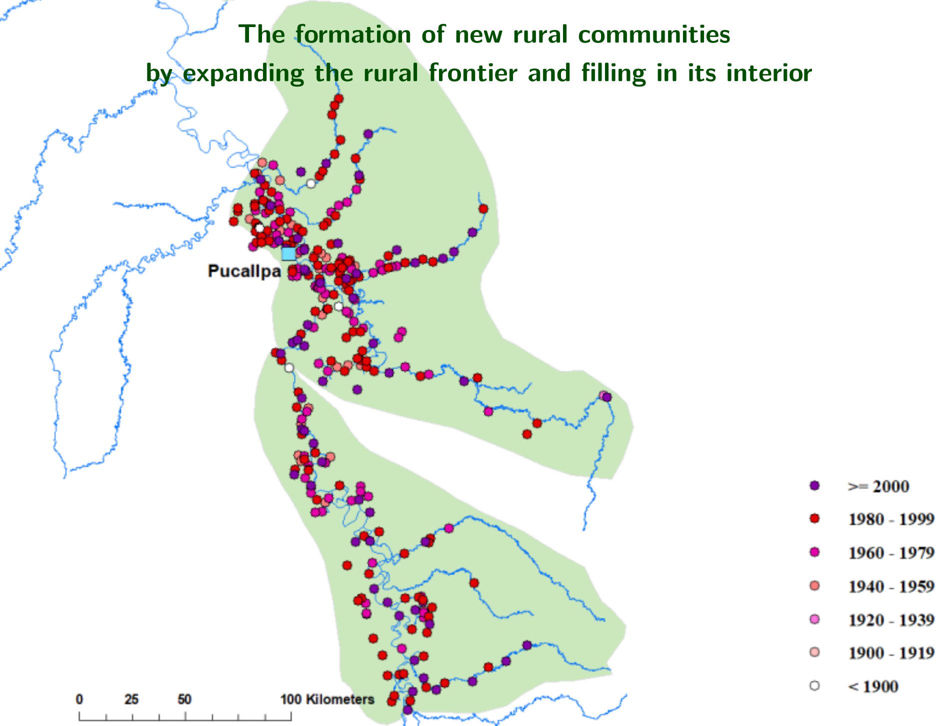
The formation of new rural communities by expanding the rural frontier and filling in its interior



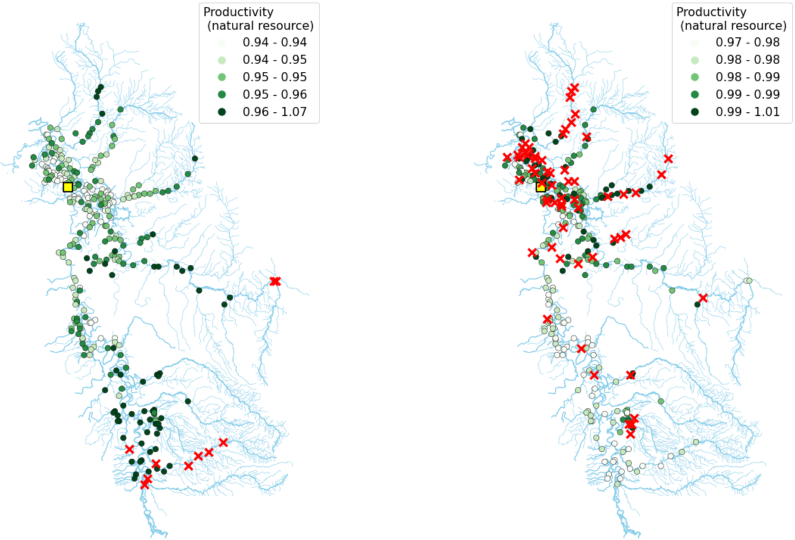
The formation of new rural communities by expanding the rural frontier and filling in its interior



The formation of new rural communities by expanding the rural frontier and filling in its interior



(A) Protected areas by rural frontier (B) Small communities not allowed



Legend: values in the counterfactual scenario relative to those in the benchmark equilibrium

Back

Improving River Transport Infrastructure: Overview

- **High trade costs**:
  - Asymmetric transport costs due to river orientations
  - Seasonality of transport costs due to water level fluctuations
  - Slow speed of river boats
- **River transport infrastructure investments**:
  - Better quality boats
  - River dredging

    Amazon Waterway Project: Government scheme with Chinese investment to deepen and widen the central parts of rivers to allow larger ships to travel

Improving River Transport Infrastructure: Overview

- **High trade costs**:
  - Asymmetric transport costs due to river orientations
  - Seasonality of transport costs due to water level fluctuations
  - Slow speed of river boats
- **River transport infrastructure investments**:
  - Better quality boats
  - River dredging

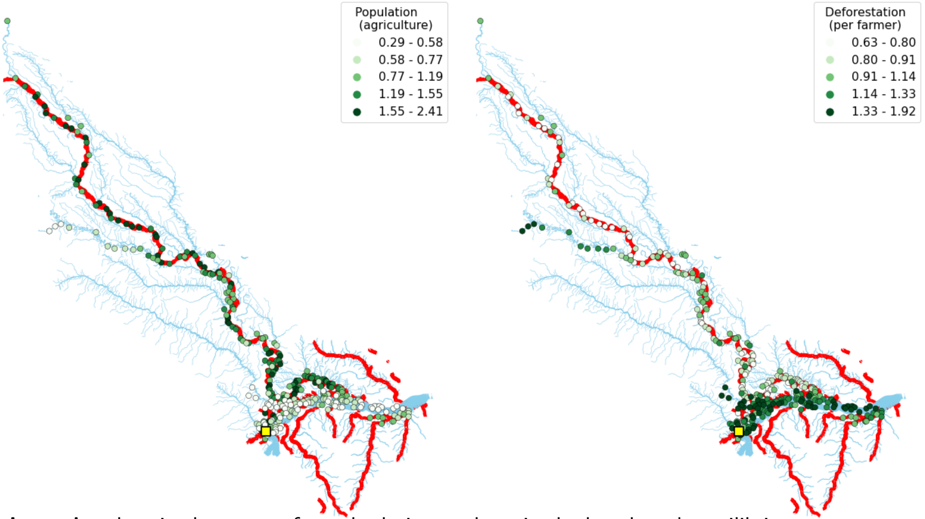
    Amazon Waterway Project: Government scheme with Chinese investment to deepen and widen the central parts of rivers to allow larger ships to travel
- Replace the **downstream-river-equivalent distance** with:

$$D_{od} = D_{od,initial,down} + \lambda_{up} D_{od,initial,up} + \lambda_{land} D_{od,land} + \lambda_{upgraded} D_{od,upgraded}$$

where $\lambda_{upgraded} = 0.8$ & symmetric transport cost in the "upgraded" part

Connecting Hinterlands to the Central Area of a Basin ⇒

- **Spatial reallocation of farmers toward remote areas**
- **Deforestation per farmer ↓ in remote areas**



Legend: values in the counterfactual relative to those in the benchmark equilibrium
Red lines: transport infrastructure improvement

Connecting Hinterlands to the Central Area of a Basin $\Rightarrow$

- **Welfare $\uparrow$**
- **Deforestation $\downarrow$**
- **Natural resource depletion $(\downarrow)$?**

Basin	Welfare	Deforestation	Natural resource depletion
Napo	+1.8%	-1.1%	-0.3%
UpperUcayali	+1.2%	-1.1%	+0.3%

- **The agglomeration benefits spread more evenly across the basin with more moderate-sized settlements**
- Total deforestation decreases by reducing the variance of the settlement size in the river basin, given the structure of congestion forces in access to land (negative and convex between ag population and per-farmer deforestation)

★ The overall impact on natural resource depletion is unclear through GE effects

Back