Are Workers Moving to High Wage Regions?

Ying Feng
National University of Singapore

Tatjana Kleineberg
The World Bank

Tommaso Porzio
Columbia University, NBER, and CEPR

Radine Rafols
The World Bank

STEG

Nairobi, January 2026

Is within-country labor mobility an engine of economic growth?

- Labor mobility can increase productivity through two channels
  1. Absolute advantage: locations differ in productivity
     - Gains from urban-rural migration (Gollin et al., 2014; Lagakos, 2020)

Is within-country labor mobility an engine of economic growth?

- Labor mobility can increase productivity through two channels
  1. Absolute advantage: locations differ in productivity
     - Gains from urban-rural migration (Gollin et al., 2014; Lagakos, 2020)
  2. Comparative advantage: some jobs exist only in certain locations
     - Migration can improve worker-job match (Bryan and Morten, 2019)

Is within-country labor mobility an engine of economic growth?

- Labor mobility can increase productivity through two channels
  1. Absolute advantage: locations differ in productivity
     - Gains from urban-rural migration (Gollin et al., 2014; Lagakos, 2020)
  2. Comparative advantage: some jobs exist only in certain locations
     - Migration can improve worker-job match (Bryan and Morten, 2019)
- Focusing on 1., we study the effects of worker reallocation from low- to high-income regions and we find that:
  - Migration flows are large and somewhat directed to high-income regions...
  - but provide only a modest contribution to aggregate productivity growth

Overview

- New dataset: merge internal migration flows with regional wage and GDP data

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%
  2. Universal internal mobility: 1 in 4 live outside their birth region by age 40

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%
  2. Universal internal mobility: 1 in 4 live outside their birth region by age 40
  3. Migration partly directed: 2/3 of movers go to higher-wage regions

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%
  2. Universal internal mobility: 1 in 4 live outside their birth region by age 40
  3. Migration partly directed: 2/3 of movers go to higher-wage regions
- Accounting framework to quantify gains from internal migration:

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%
  2. Universal internal mobility: 1 in 4 live outside their birth region by age 40
  3. Migration partly directed: 2/3 of movers go to higher-wage regions
- Accounting framework to quantify gains from internal migration:
  - Observed labor mobility makes modest contribution to countries' development
  - Results similar across development spectrum
  - Robust to other income measures, samples, and regional aggregations

Overview

- New dataset: merge internal migration flows with regional wage and GDP data
- Document three key facts:
  1. Ubiquitous regional wage gaps: 80th-20th percentile gap averages 30%
  2. Universal internal mobility: 1 in 4 live outside their birth region by age 40
  3. Migration partly directed: 2/3 of movers go to higher-wage regions
- Accounting framework to quantify gains from internal migration:
  - Observed labor mobility makes modest contribution to countries' development
  - Results similar across development spectrum
  - Robust to other income measures, samples, and regional aggregations
  - Channels: 1. distance, 2. downward moves, 3. more directed migration

Data Construction

Data Construction: Sources

1. Lifetime migration flows
    - IPUMS census data

Data Construction: Sources

1. Lifetime migration flows
    - IPUMS census data
2. Regional wage data
    - Representative household/labor force surveys from World Bank I2D2 and Global Labor Database (GLD), some censuses from IPUMS

Data Construction: Sources

1. Lifetime migration flows
   - IPUMS census data
2. Regional wage data
   - Representative household/labor force surveys from World Bank I2D2 and Global Labor Database (GLD), some censuses from IPUMS
3. Satellite-based local GDP measures
   - Estimates from Rossi-Hansberg and Zhang (2025)

Data Construction: Sources

1. Lifetime migration flows
   - IPUMS census data
2. Regional wage data
   - Representative household/labor force surveys from World Bank I2D2 and Global Labor Database (GLD), some censuses from IPUMS
3. Satellite-based local GDP measures
   - Estimates from Rossi-Hansberg and Zhang (2025)
4. Movers in Panel Data (in progress)
   - Administrative sources and panel labor force surveys

Data Construction

- Migration flows from IPUMS
  - Map non-harmonized birth-place codes to current harmonized location codes using location labels, organic text, and map searching
  - 1st sub-national unit (geolev1): 76 countries, 224 country-years, 1960-2016
  - 2nd sub-national unit (geolev2): 30 countries

Data Construction

- Migration flows from IPUMS
  - Map non-harmonized birth-place codes to current harmonized location codes using location labels, organic text, and map searching
  - 1st sub-national unit (geolev1): 76 countries, 224 country-years, 1960-2016
  - 2nd sub-national unit (geolev2): 30 countries
- Wage data from labor force surveys
  - Micro-data on wages, hours worked, demographics, schooling, and location
  - Allows estimating raw and residualized regional wages (net of worker charact.)
  - Sample: 39 countries, 156 country-year, 1960-2024

Data Construction

- Migration flows from IPUMS
  - Map non-harmonized birth-place codes to current harmonized location codes using location labels, organic text, and map searching
  - 1st sub-national unit (geolev1): 76 countries, 224 country-years, 1960-2016
  - 2nd sub-national unit (geolev2): 30 countries
- Wage data from labor force surveys
  - Micro-data on wages, hours worked, demographics, schooling, and location
  - Allows estimating raw and residualized regional wages (net of worker charact.)
  - Sample: 39 countries, 156 country-year, 1960-2024
- Satellite-based GDP measure
  - Global coverage, annual 2012-2021, fine geographic disaggregation
  - Map 0.25° resolution to first and second sub-national units (geolev1/2)

Data Construction: Merging Cross-sectional Datasets and Coverage

- Merge data sets at country-year-region level
  - Harmonize geographical units over time (LFS) and across datasets
  - Adjust spelling, levels of aggregation, border changes, etc.

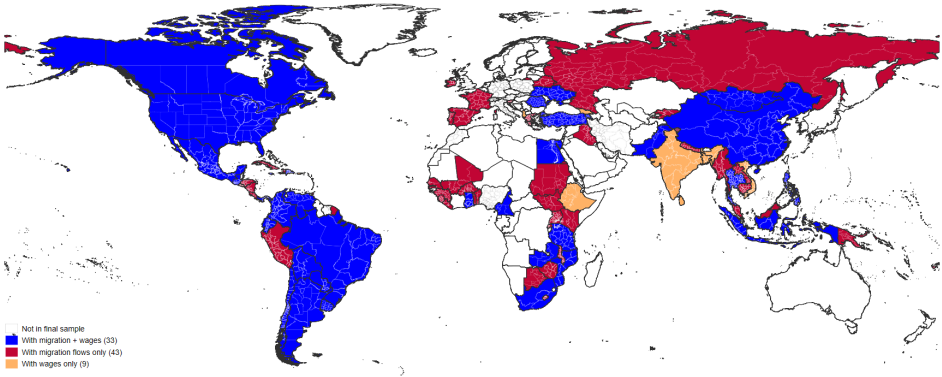
Data Construction: Merging Cross-sectional Datasets and Coverage

- Merge data sets at country-year-region level
  - Harmonize geographical units over time (LFS) and across datasets
  - Adjust spelling, levels of aggregation, border changes, etc.
- Data Cleaning and Validation
  - Inspect time-series of wage data at country-year-region level
  - Trim irregularities and spikes
  - Strong correlation between regional wages and GDP measure Figure

Data Construction: Merging Cross-sectional Datasets and Coverage

- Merge data sets at country-year-region level
  - Harmonize geographical units over time (LFS) and across datasets
  - Adjust spelling, levels of aggregation, border changes, etc.
- Data Cleaning and Validation
  - Inspect time-series of wage data at country-year-region level
  - Trim irregularities and spikes
  - Strong correlation between regional wages and GDP measure Figure
- Coverage of merged datasets
  - → Migration and wage data: 33 countries (low- and high-income) Details
  - → Migration and GDP data at 1st subnational level: 58 countries Details
  - → Migration and GDP data at 2nd subnational level: 30 countries Details

Sample Coverage



→ 33 countries with migration and wage data at geolev1 level (blue)

Data Construction: Panel data (in progress)

- Current sample:
  1. Brazil: RAIS for formal wage workers
  2. Indonesia: Indonesia Family Life Survey (IFLS)
  3. South Africa: National Panel Survey
  4. Tanzania: National Panel Survey
  5. United States: Panel Study of Income Dynamics (PSID-SHELF)

  $\rightarrow$ Any others?

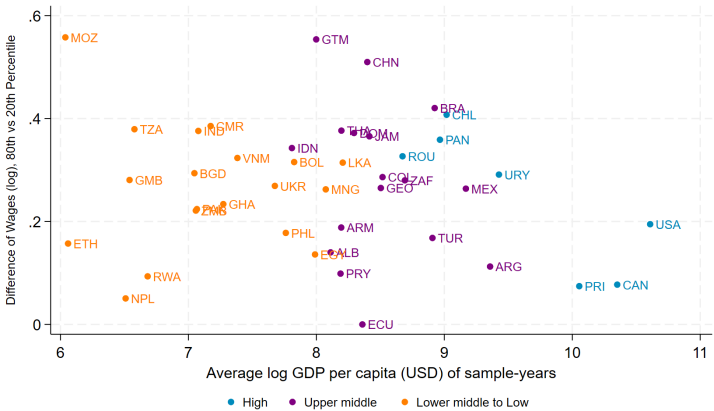
Data Construction: Panel data (in progress)

- Current sample:
  1. Brazil: RAIS for formal wage workers
  2. Indonesia: Indonesia Family Life Survey (IFLS)
  3. South Africa: National Panel Survey
  4. Tanzania: National Panel Survey
  5. United States: Panel Study of Income Dynamics (PSID-SHELF)

  $\rightarrow$ Any others?
- Data Cleaning and Merging
  - Identify cross-region movers btw rounds (not always possible, Mexico ENOE)
  - Validation: Moving flows correlate in panel data and IPUMS Flows
  - Merge with wage data from labor force surveys

Empirical Facts

Fact 1a: Large within-country Variation in Hourly Wages (30% for 80-20 pct.)



Variation in schooling | Imputed Wages

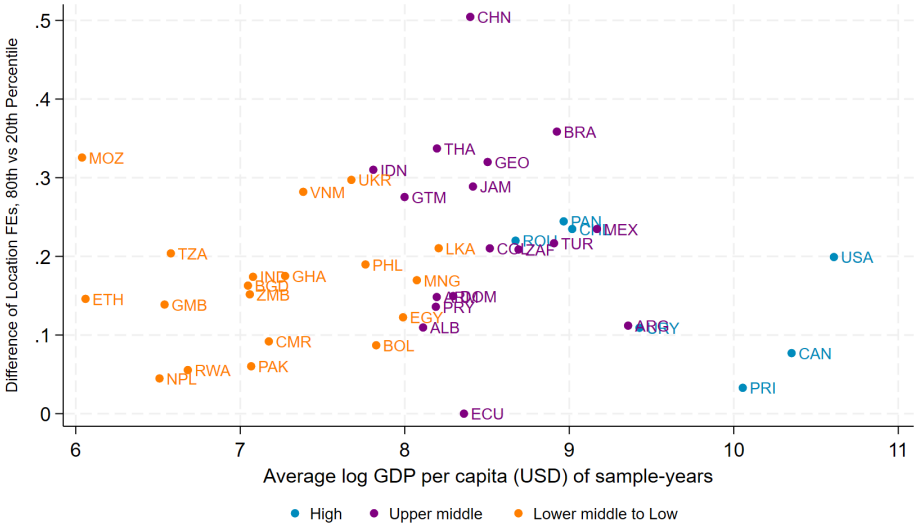
Accounting for Differences in Worker Characteristics

- Measure regions' residualized wage after accounting for worker characteristics
- Estimate Mincerian wage regression in micro-data:

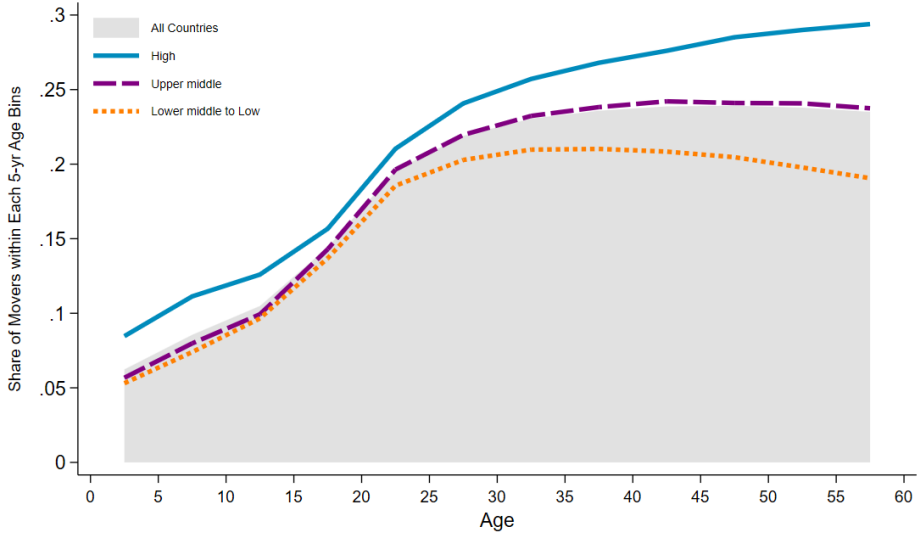
$$\log(\text{wage}^i) = \beta_0 + \beta_1 \text{Educ}^i + \beta_2 \text{Exp}^i + \beta_3 \text{Male}^i + \delta_j + u^i,$$

- Individual-level data on hourly wages ($wage^i$), years of schooling (Educ^i), and experience (Exp^i, constructed as age minus schooling-years)
- Region fixed effects (δ_j) capture residualized wages
- Sample: workers aged 25-55

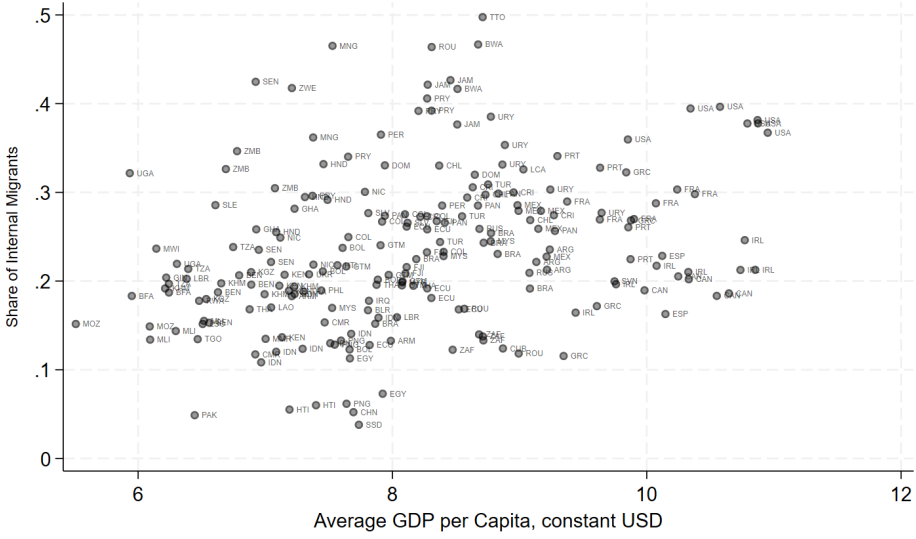
Fact 1b: Large within-country Variation Remains With Region Fixed Effects (20%)



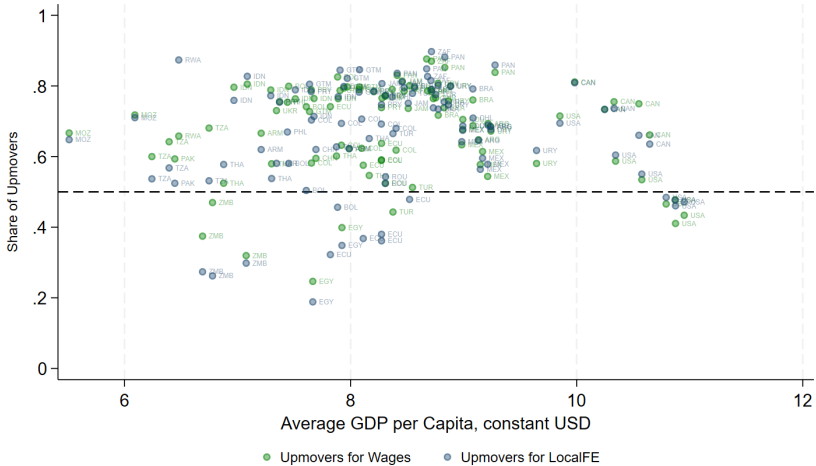
Fact 2a) High Labor Mobility: 1 in 4 Live Outside their Birthplace by Age 40



Fact 2b: Labor Mobility Slightly Positively Correlated with GDP pc (40-44 cohort)

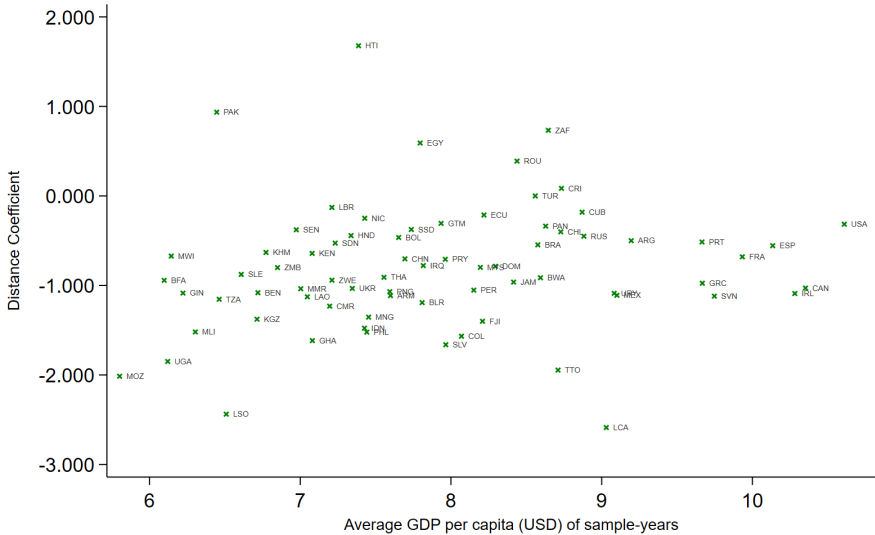


Fact 3a: Partly Directed Migration: 2/3 Move to Higher-wage Regions



This sample conditions on lifetime movers among the 40-44 cohorts. Coefficiens of In- and Outflows

Fact 3b: Gravity Matters: Doubling Distance Decreases Migration by 36%

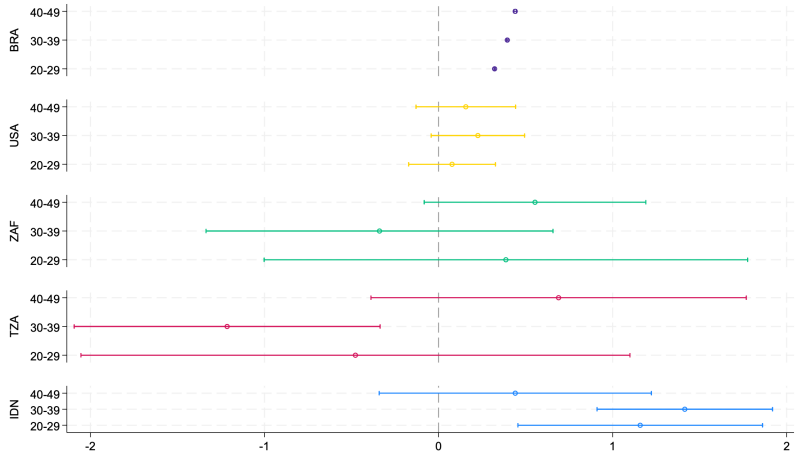


Fact 3b: Minimal Heterogeneity in Migration Frictions

		Age		HH composition		Education	
	All	**Young**	**Old**	**w/Child≤6yrs**	**no Child**	**non-College**	**College**
Distance	-0.362***	-0.328***	-0.309***	-0.348***	-0.313***	-0.336***	-0.350***
	(0.027)	(0.026)	(0.029)	(0.026)	(0.031)	(0.028)	(0.035)
Origin x Year FE	Y	Y	Y	Y	Y	Y	Y
Dest x Year FE	Y	Y	Y	Y	Y	Y	Y
N countries	74	74	74	74	74	74	66
N regions	1253	1253	1253	1253	1253	1253	1140

Fact 4: Actual Mover Gains Relate to Origin-Destination Wage Gaps (Panel Data)

$$\Delta(\log w_i) = \widehat{\beta}(\overline{w}_d^{\text{cross}} - \overline{w}_o^{\text{cross}}) + X_i'\gamma + \epsilon_i$$



Accounting Framework

Accounting Framework to Measure Gains from Internal Migration

- Assume wage differences across regions reflect marginal product of labor

Accounting Framework to Measure Gains from Internal Migration

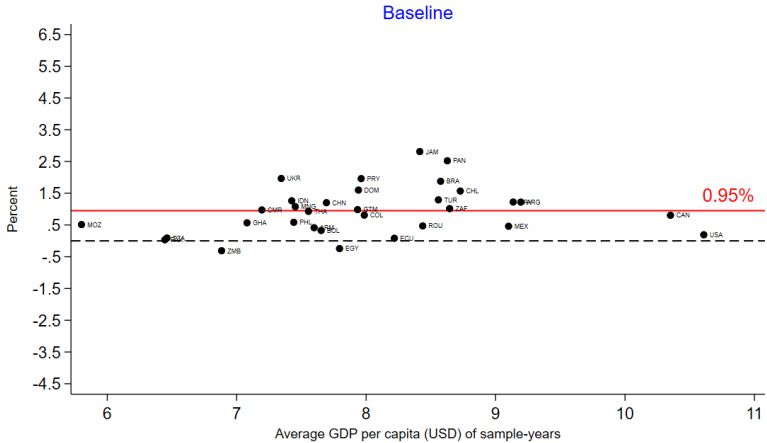
- Assume wage differences across regions reflect marginal product of labor
- For each country-year, calculate (percent) wage gains from observed labor mobility:

$$\Delta \overline{Y} = \underbrace{\log\left(\sum_j s_j^{\text{actual}} \overline{Y}_j\right)}_{\text{avg wage with migration}} - \underbrace{\log\left(\sum_i s_i^{\text{birth}} \overline{Y}_i\right)}_{\text{avg wage without migration}},$$

$\rightarrow$ Gains from migration compare pop-weighted average wage to a counterfactual average wage that weights regions by birthplace populations (no-migration case)

- $\overline{Y}$ are national ($\overline{Y}_i$ regional) average wages, s_i^{actual} are regions' actual, post-migration population shares; s_i^{birth} are birthplace, no-migration population shares

Wage Gains from Observed Internal Migration Flows (wage measure: region FE)



- Across countries: Aggregate gain 1%, avg gain per mover 6.3% or US$10,178

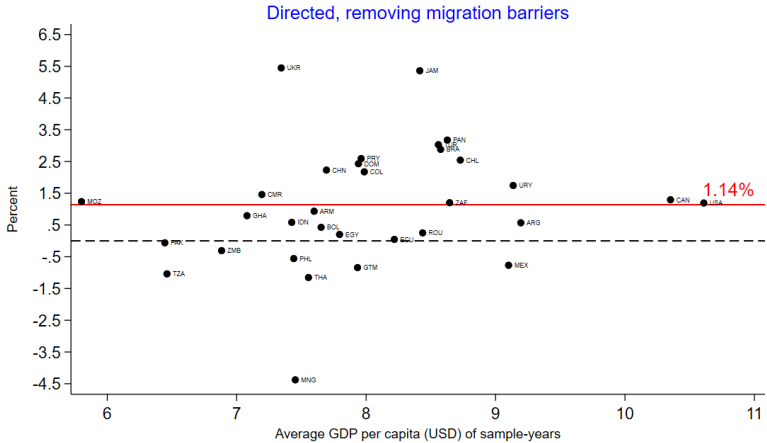
Examining Factors that Determine or Limit Gains from Internal Migration

- Compute potential wage gains from counterfactual migration scenarios:

$$\Delta \overline{Y} = \underbrace{\log\left(\sum_j s_j^{\text{actual}} \overline{Y}_j\right)}_{\text{actual avg wage}} - \underbrace{\log\left(\sum_i s_i^{\text{counterfactual}} \overline{Y}_i\right)}_{\text{avg wage with CF pop. allocation}},$$

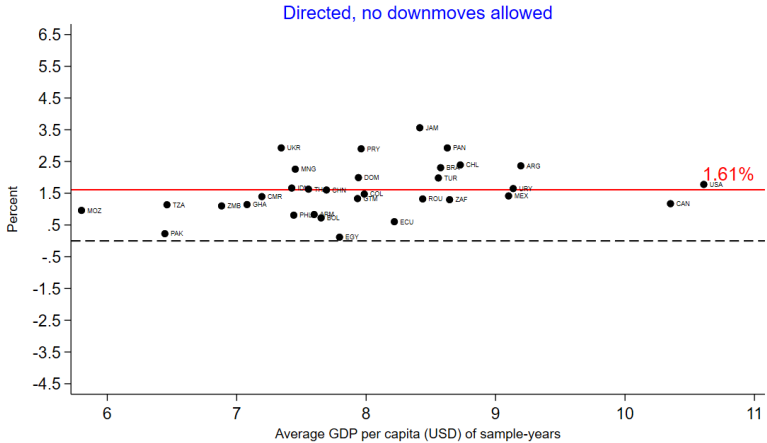
- Evaluate migration scenarios by changing regions' population weights s_i:
  1. Distance barrier: Remove distance effect from gravity relationship
  2. No down-movers: Allocate all down-movers back to birthplace
  3. Increase directed migration: Allocate movers to +80th percentile of regions

Wage Gains from Counterfactual Migration: Removing Distance Effect



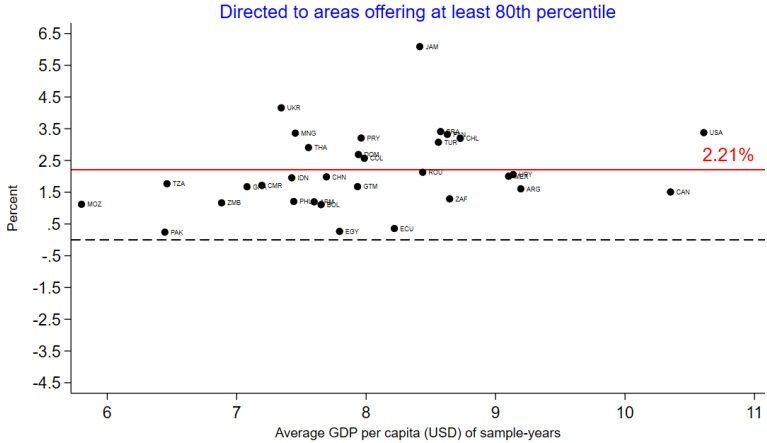
- Across countries: Aggregate gain 1.14%, avg gain per mover 7.6% or US$14,325.

Wage Gains from Counterfactual Migration: Removing Down-Movers



- Across countries: Aggregate gain 1.61%, avg gain per mover 10.6% or US$17,902.

Wage Gains from Counterfactual Migration: Assigning Movers to $\geq$ 80pct



- Across countries: Aggregate gain 2.2%, avg gain per mover 14% or US$24,550.

CohortGains | Conditional on Movers | Other Outcomes

Modest Effects Robust to Income Measures, Samples, and Regional Aggregation

	Wage Data (LFS)		Satellite-based GDP Data (RHZ)		
	Residual Wages	Raw Wages	Same Sample	Expanded Sample	GEOLEV2
	(1)	(2)	(3)	(4)	(5)
Baseline	0.95	2	2.95	3.17	4.63
No Gravity	1.15	2.51	4.97	5.87	6.02
No downmoves	1.61	3.01	4.52	4.63	6.58
>80pctile	2.21	3.97	5.8	5.99	10.26
N Countries	33	33	33	58	30

- Effects of net worker reallocation modest across samples and income measures
- Most income variation lies between 1st sub-national units Decomposition Raw Wages

Conclusion and Next Steps

- Net worker reallocation towards higher-wage regions is present but not a major driver of development, suggesting other channels
  - Removing barriers increase effects but may require overcoming preferences
  - Directing flows even with optimistic scenarios yields only modest effects (3%)

Conclusion and Next Steps

- Net worker reallocation towards higher-wage regions is present but not a major driver of development, suggesting other channels
  - Removing barriers increase effects but may require overcoming preferences
  - Directing flows even with optimistic scenarios yields only modest effects (3%)
- Policies that simply encourage moving might not have large growth effects unless policies can change regions' underlying productivity (wages)
  - E.g., agglomeration, education, worker-job matches, local search frictions, etc.

Conclusion and Next Steps

- Net worker reallocation towards higher-wage regions is present but not a major driver of development, suggesting other channels
  - Removing barriers increase effects but may require overcoming preferences
  - Directing flows even with optimistic scenarios yields only modest effects (3%)
- Policies that simply encourage moving might not have large growth effects unless policies can change regions' underlying productivity (wages)
  - E.g., agglomeration, education, worker-job matches, local search frictions, etc.
- Next steps: shed more light on why gains are muted
  - Panel data to see rationales for moves where private gains are limited
  - ... other suggestions?

References

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229–2268.

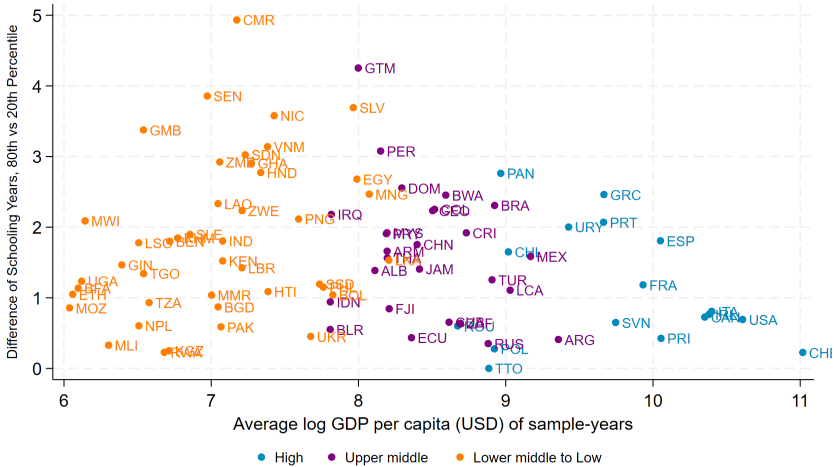
Gollin, D., D. Lagakos, and M. E. Waugh (2014). Agricultural productivity differences across countries. *American Economic Review 104*(5), 165–70.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic perspectives 34*(3), 174–192.

Rossi-Hansberg, E. and J. Zhang (2025). Local gdp estimates around the world. Technical report, National Bureau of Economic Research.

Thank you

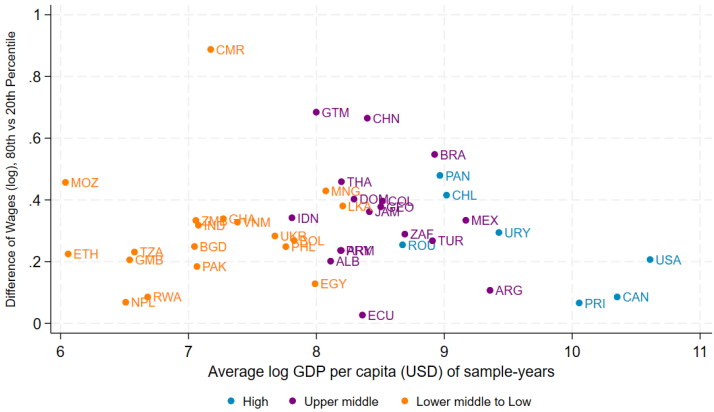
Robustness Fact 1: Within-country Variation in Schooling



Back

Robustness Fact 1: Within-country Variation in Wages (imputed)

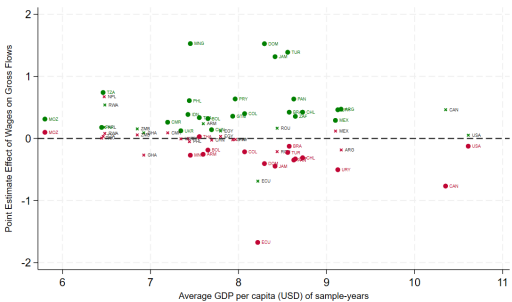
Takes the regional average of imputed wages for persons with missing hourly wages but are active in the labor force (employed, self-employed and unpaid).



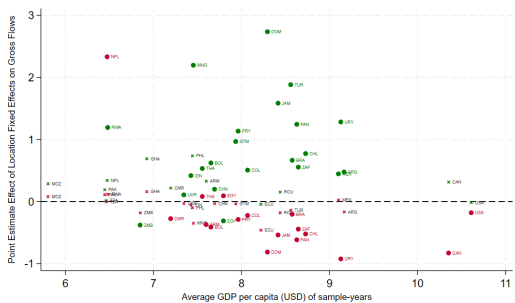
Back

Robustness Fact 3: Coefficients of Gross Inflows (green) and Outflows (red)

Wages



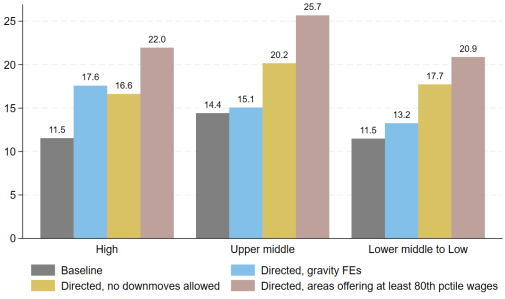
Regional FE



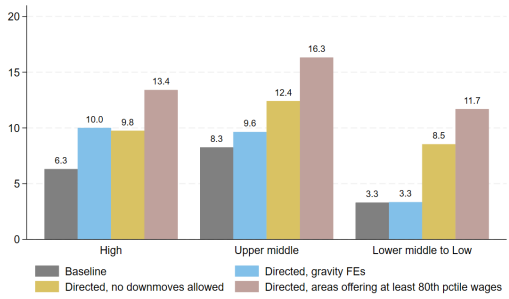
Back

Gains Conditional on Movers

Wages

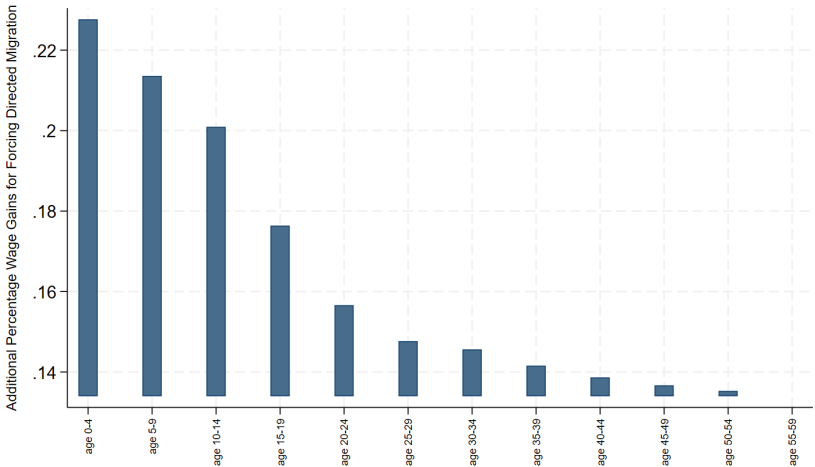


Regional FE



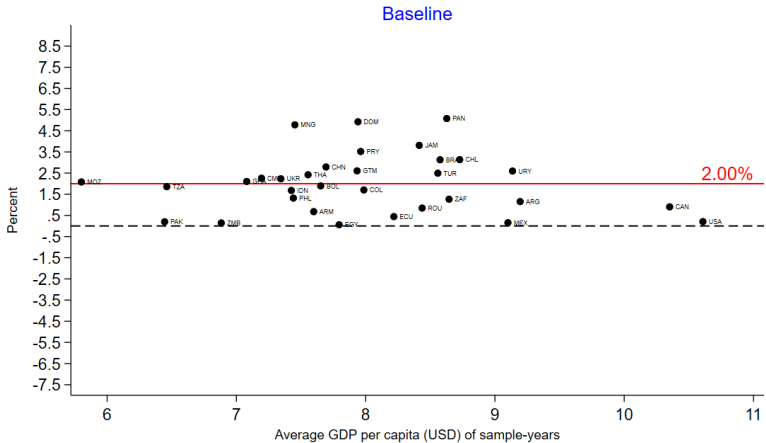
Back

Cohort Gains from Directed Moves



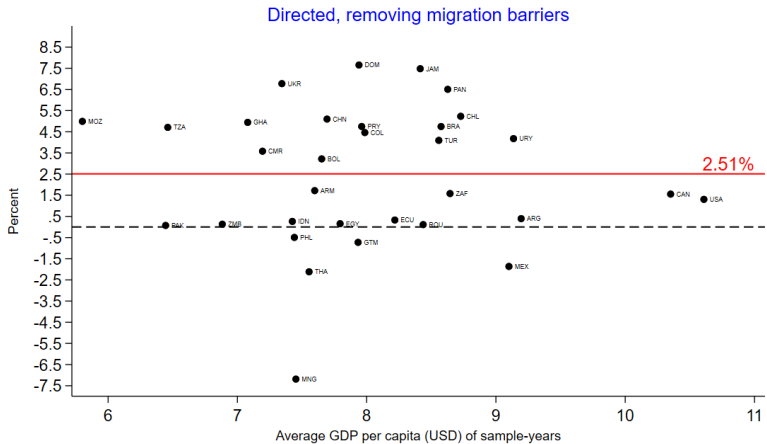
Back

% Wage Gains due to Migration - Actual Flows



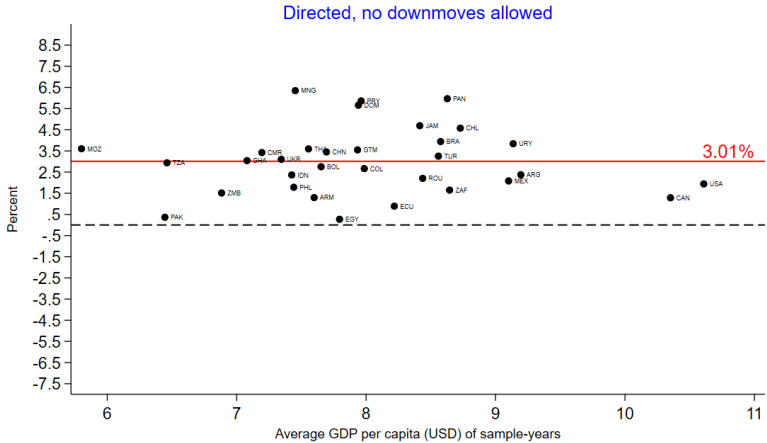
- The average gain of movers is 8.3%, or US$11,892, across countries.

% Wage Gains with Counterfactual Flows: Removing Geo-barriers (Distance)



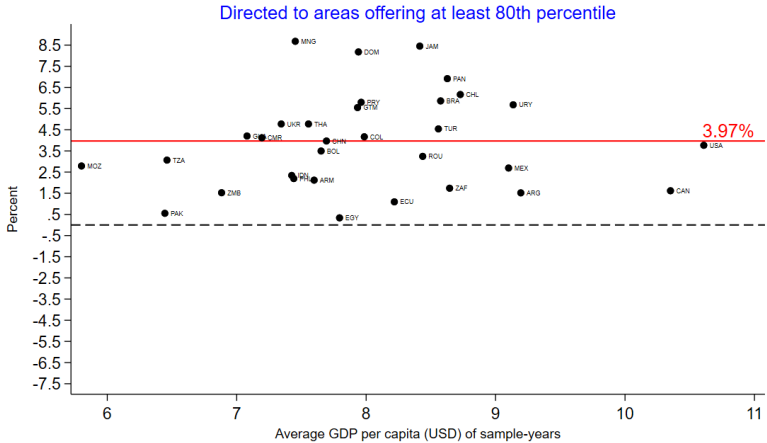
- The average gain of movers is 9.6%, or US$14,666, across countries.

% Wage Gains with Counterfactual Flows: Removing Down-movers



- The average gain of movers is 14.7%, or US$22,158, across countries.

% Wage Gains with Counterfactual Flows: Assigning all movers to $\geq$ 80p



- The average gain of movers is 18.8%, or US$29,531, across countries. Back

Impact on other socio-economic outcomes

Relative to their birthplace, people relocate to regions with:

- higher **paid employment** share by 0.91 percentage points
- lower **agricultural** employment share (1.5 p.p.)
- lower shares of **own-account/self-employed** workers (1.0 p.p.)

Back

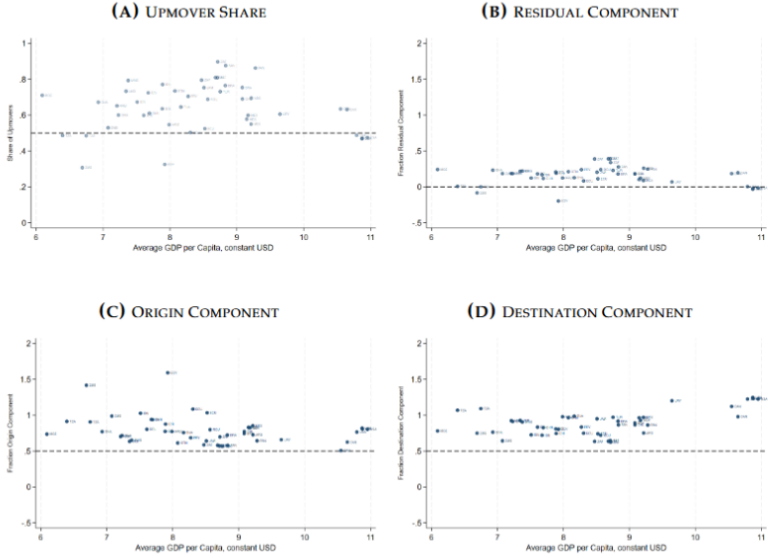
Decomposing Directed Migration

Each region's upmover share is defined as

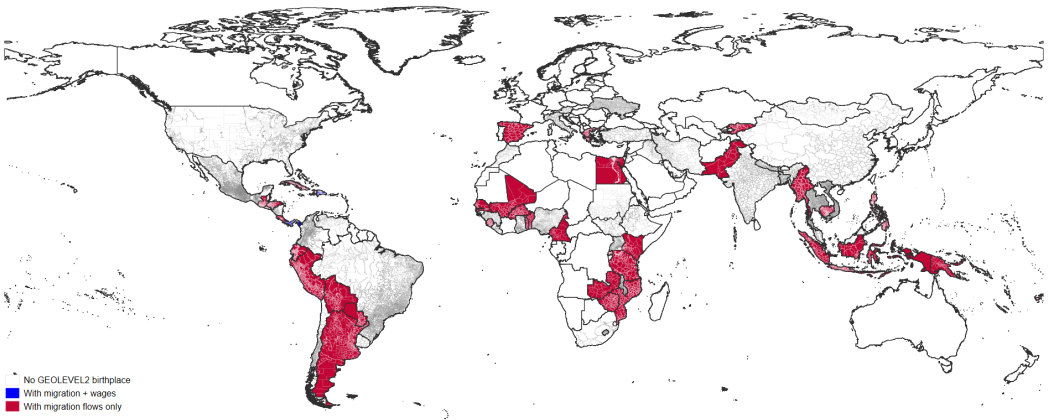
$$\sum_i \sum_j \lambda_{ij}\, \mathbb{1}\{w_j > w_i\} = \underbrace{\tfrac{1}{2}\sum_i \bar{\lambda}_i\, \theta_i}_{\text{Origin composition}} + \underbrace{\tfrac{1}{2}\sum_j \bar{\lambda}_j\, \theta_j}_{\text{Destination composition}} + \underbrace{\sum_{ij}\left(\lambda_{ij} - \tfrac{1}{2}\bar{\lambda}_i - \tfrac{1}{2}\bar{\lambda}_j\right)\mathbb{1}\{w_j > w_i\}}_{\text{Residual (sorting)}}$$

... where λ_i is origin i's share of all out-migration, θ_i are number of destinations paying more than origin i, λ_j is destination j's share of all in-migration, and θ_j are number of alternative locations paying less than destination j.

Decomposing Directed Migration

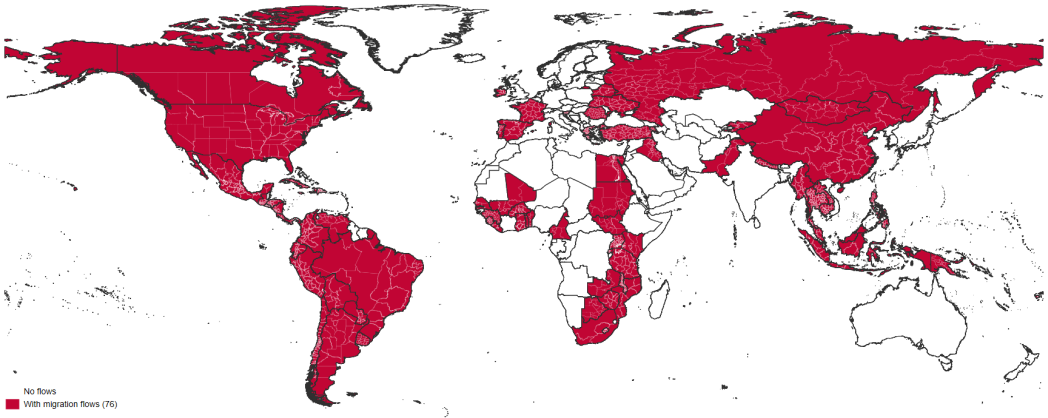


Countries with birthplace information at GEOLEV2



Back

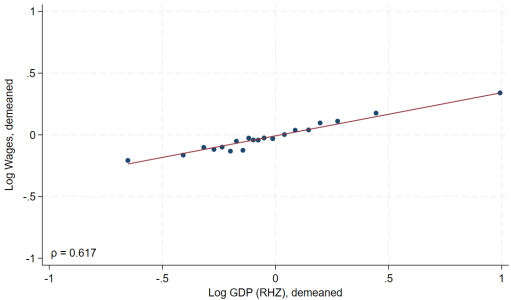
Countries with birthplace information (RHZ sample)

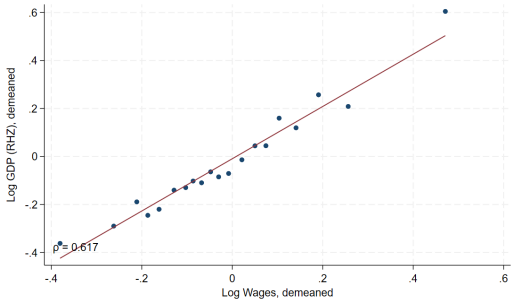


Back

Data validation w/ Rossi-Hansberg & Zhang (2025)

Binscatter plots:





Back

Decomposing Income Inequality across Different Levels of Geographic Aggregation

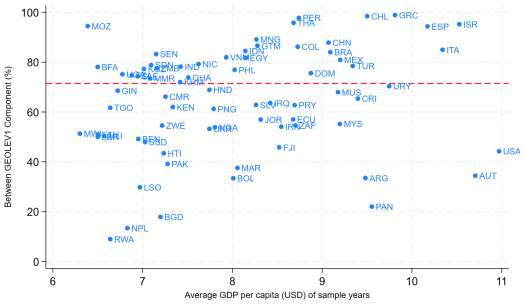
- High-resolution satellite data allows decomposing income variation across levels:
  1. Decompose income variation across geolev2 units into: Between-geolev1 and between-geolev2 within-geolev1 (IPUMS definitions)
  2. Decompose income variation across finest grids into: Between geolev1, between geolev2 within-geolev1, and between-grids within-geolev2

$$\sum_g s_g(\overline{Y}_g - \overline{Y})^2 = \underbrace{\sum_j s_j \left(\overline{Y}_j - \overline{Y}\right)^2}_{\text{between GEOLEV1}} + \underbrace{\sum_j \sum_{g \in j} s_g \left(\overline{Y}_g - \overline{Y}_j\right)^2}_{\text{between GEOLEV2 within GEOLEV1}}$$

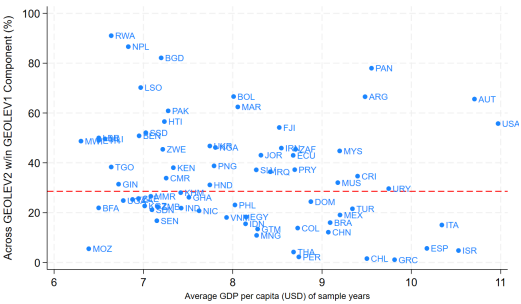
- j are states/provinces (geolev1) and g are smaller units (geolev2)
- s_g are regions' population share; $\overline{Y_g}$ regions' average income

2-level Decomposition: 70% of Variation Explained between States/Provinces (avg)

Between GEOLEV1

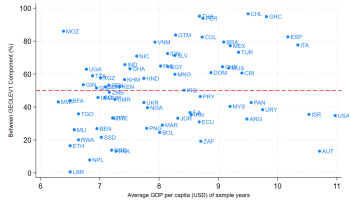


Across GEOLEV2 within GEOLEV1

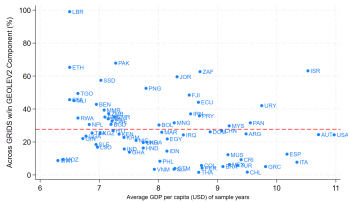


3-level Decomposition: 50% of Variation Explained between States/Provinces (avg)

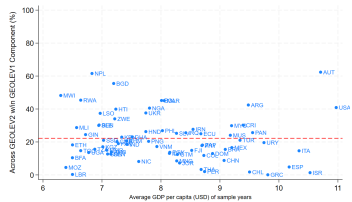
Between GEOLEV1



Across GEOLEV2 within GEOLEV1



Across GRIDS within GEOLEV2



Back

Country Coverage by Income Classification

Income Category	Pop'n (in billions)	Total countries	Countries w/ flows only	Countries w/ wage only	Countries w/ flows+wage
High	676	17	13	7	6
Low	406	15	13	5	1
Lower Middle	2668	28	24	15	10
Upper Middle	2733	28	26	18	16

Back

Region	Countries
East Asia and Pacific	Cambodia, China, Fiji, Indonesia, Laos, Malaysia, Mongolia, Myanmar, Papua New Guinea, Philippines, Thailand, Vietnam
Europe and Central Asia	Albania, Armenia, Belarus, France, Georgia, Greece, Ireland, Italy, Kyrgyzstan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Switzerland, Turkey, Ukraine
Latin America and Caribbean	Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Jamaica, México, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Saint Lucia, Trinidad and Tobago, Uruguay, Venezuela
Middle East and North Africa	Egypt, Iraq
North America	Canada, United States
South Asia	Bangladesh, India, Nepal, Pakistan, Sri Lanka
Sub-Saharan Africa	Benin, Botswana, Burkina Faso, Cameroon, Ethiopia, Gambia, Ghana, Guinea, Kenya, Lesotho, Liberia, Malawi, Mali, Mozambique, Rwanda, Senegal, Sierra Leone, South Africa, South Sudan, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe

Back

Country Coverage by Region

Region Category	Pop'n (in billions)	Total countries	Countries w/ flows only	Countries w/ wage only	Countries w/ flows+wage
East Asia and Pacific	2108	12	11	6	5
Europe and Central Asia	561	18	13	6	4
Latin America and Caribbean	654	24	23	15	14
Middle East and North Africa	160	2	2	1	1
North America	375	2	2	2	2
South Asia	1909	5	2	5	1
Sub-Saharan Africa	716	25	23	10	6

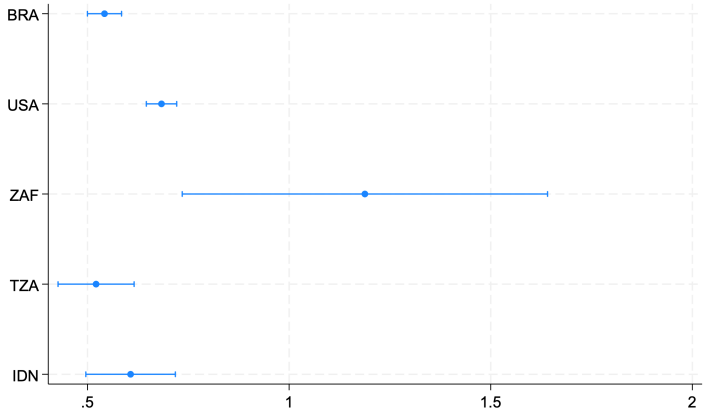
Back

Income Category	Countries
High	Canada, Chile, France, Greece, Ireland, Italy, Panama, Poland, Portugal, Puerto Rico, Romania, Slovenia, Spain, Switzerland, Trinidad and Tobago, United States, Uruguay
Upper Middle	Albania, Argentina, Armenia, Belarus, Botswana, Brazil, China, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Fiji, Georgia, Guatemala, Indonesia, Iraq, Jamaica, Malaysia, México, Paraguay, Peru, Russia, Saint Lucia, South Africa, Thailand, Turkey, Venezuela
Lower Middle	Bangladesh, Benin, Bolivia, Cambodia, Cameroon, Egypt, El Salvador, Ghana, Honduras, India, Kenya, Kyrgyzstan, Laos, Lesotho, Mongolia, Myanmar, Nepal, Nicaragua, Pakistan, Papua New Guinea, Philippines, Senegal, Sri Lanka, Tanzania, Ukraine, Vietnam, Zambia, Zimbabwe
Low	Burkina Faso, Ethiopia, Gambia, Guinea, Haiti, Liberia, Malawi, Mali, Mozambique, Rwanda, Sierra Leone, South Sudan, Sudan, Togo, Uganda

Back

Comparing Bilateral Moving Flows in Panel and Cross-sectional Data

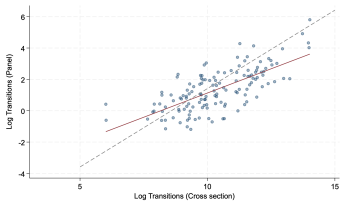
Flow elasticity: $\log \pi_{od}^{\text{panel}} = \alpha + \widehat{\beta} \log \pi_{od}^{\text{cross}} + \mathit{epsilon}_{od}, \qquad o \neq d$



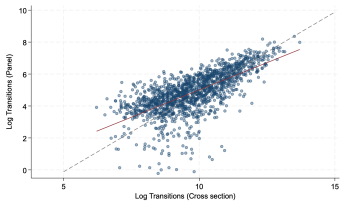
Back

Visualizing at o-d-year (data validation)

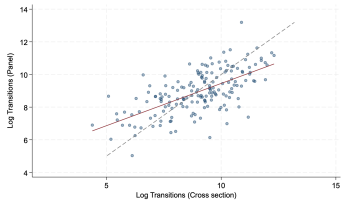
Indonesia



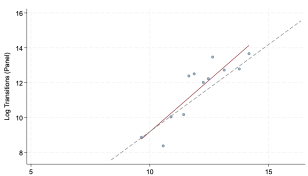
United States



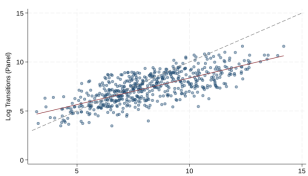
Tanzania



South Africa



Brazil



Back