The Dynamics of Internal Migration: A New Fact and its Implications

Greg Howard   Hansen Shao

University of Illinois

IEB Urban Conference
2024

Two facts about internal migration

1. Internal migration is rare
2. People move to nearby and populous places (gravity)

Two facts about internal migration

1. Internal migration is rare
2. People move to nearby and populous places (gravity)

How should we understand these two facts?

(a) Moving costs?

(b) Persistent preferences?

Two facts about internal migration

1. Internal migration is rare
2. People move to nearby and populous places (gravity)

How should we understand these two facts?

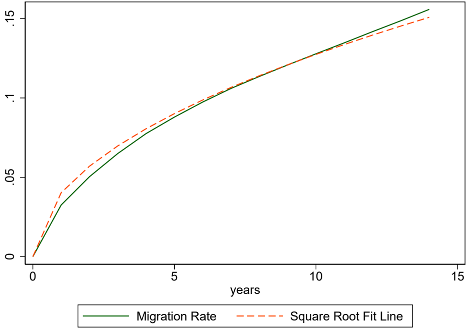
(a) Moving costs?

(b) Persistent preferences?

Does it matter how we understand these two facts?

- Literature has emphasized moving costs
  - Tractable
  - Easily matches both facts
  - Natural extension of the trade literature

- Persistent preferences is consistent with a new fact about the dynamics of migration:
  3. t-year migration rate is proportional to $\sqrt{t}$



Outline

Main Question

Does the way we model internal migration have consequences for how we answer important spatial economic questions?

Outline

Main Question

Does the way we model internal migration have consequences for how we answer important spatial economic questions?

1. New fact about internal migration
   - Hard to reconcile with existing models

Outline

Main Question

Does the way we model internal migration have consequences for how we answer important spatial economic questions?

1. New fact about internal migration
   - Hard to reconcile with existing models
2. New model
   - Idiosynchratic preferences correlated across space and time
   - Model can match new fact, while also generating rare migration and the gravity relationship

Outline

Main Question

Does the way we model internal migration have consequences for how we answer important spatial economic questions?

1. New fact about internal migration
   - Hard to reconcile with existing models
2. New model
   - Idiosynchratic preferences correlated across space and time
   - Model can match new fact, while also generating rare migration and the gravity relationship
3. New Implications of the model
   - Short-run elasticities are still similar to existing models
   - Long-run elasticities are very different
   - Speed of population adjustment is very different

Contributions to the Literature

- Spatial dynamics
  - Rise and decline of regional economies
    Blanchard and Katz (1992); Caliendo, Dvorkin and Parro (2019); Allen and Donaldson (2020); Morris-Levenson and Prato (2022); Glaeser and Gyourko (2005); Kleinman, Liu and Redding (2023); Amior and Manning (2018); Davis, Fisher and Veracierto (2021)
  - Macro adaptation to local shocks
    Tombe and Zhu (2019); Hao, Sun, Tombe and Zhu (2020); Eckert and Peters (2018); Giannone (2017); Heise and Porzio (2021); Bryan and Morten (2019); Cruz and Rossi-Hansberg (2021); Oliveira and Pereda (2020); Schubert (2021)
- How to model migration
  - Modifications of the dynamic logit
    Kennan and Walker (2011); Kaplan and Schulhofer-Wohl (2017); Giannone, Li, Paixao and Pang (2020); Porcher (2020); Mangum and Coate (2019); Monras (2018); Coen-Pirani (2010); Davis et al. (2021); Zerecero (2021)
  - Persistent preferences
    Bayer and Juessen (2012)
  - Empirical evidence
    Saks and Wozniak (2011); Farrokhi and Jinkins (2021); Koşar, Ransom and Van der Klaauw (2021); Fujiwara, Morales and Porcher (2022)
- Multinomial probits
  - Butler and Moffitt (1982); Keane (1992); Geweke, Keane and Runkle (1994)

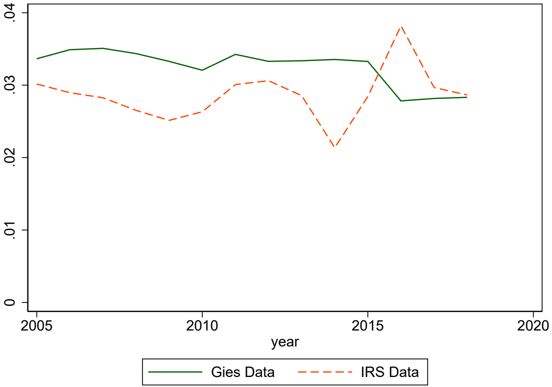
3 Facts about Internal Migration

Data

- Gies Consumer and Small Business Credit Panel (GCCP)
  - Credit data from one of the leading providers of credit reports
  - 1 percent of Americans with credit reports
  - Includes state of residence
  - Panel data, 2004-2018
- IRS Migration Data
  - Based on tax filings
  - Aggregated flows of state-to-state migration

Fact #1

Migration is rare



Comparison of interstate migration rates in IRS and GCCP

Fact #2

Migration follows a gravity pattern

Poisson regression:

$$\log m_{i\to j} = \beta \log \text{distance}_{ij} + \alpha \log p_i + \gamma \log p_j + \epsilon_{ij}$$

	(1) Migration (IRS)	(2) Migration (Credit)
Log Distance	-0.736***	-0.744***
	(0.0572)	(0.0515)
Log Origin Population	0.900***	0.923***
	(0.0832)	(0.0797)
Log Destination Population	0.822***	0.893***
	(0.0976)	(0.0799)
Observations	2550	2550

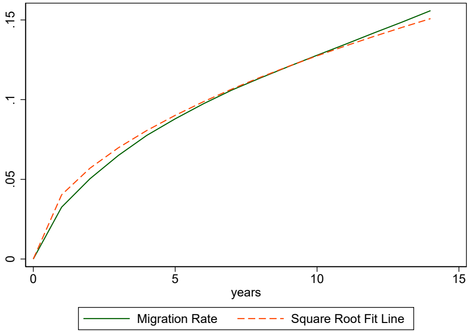
Standard Errors are two-way clustered by origin and destination states

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Define t-year interstate migration rate as the share of people who live in a different state than they did t years ago

Fact #3 (New)

t-year interstate migration rate is proportional to $\sqrt{t}$



- Implies a high rate of return or onward migration
- Suggestive of persistent preferences PSID

Standard Dynamic Logit

I locations indexed by i, N individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(i) = \max_j v_{jt} - \delta_{ij} + \epsilon_{jnt} + \mathbb{E}[V_{nt+1}(j)]$$

- ϵ_{jnt} is i.i.d. and has an extreme value distribution

Comparison to Standard Model

- In standard model, migration is Markov
  - State variable is current location
  - When migration is rare, t-year migration proportional to t
- Can be reconciled...
  - ...with flexible tenure-dependent moving costs
  - ...or with location attachment
  - ...but requires many fine-tuned parameters

The SPACE Model

Model

I locations indexed by i, continuum of individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(\vec{\epsilon}_{nt}) = \max_i \{v_{it} + \epsilon_{int}\} + \beta \mathbb{E}[V_{nt+1}(\vec{\epsilon}_{nt+1}) | \vec{\epsilon}_{nt}]$$

  - No moving costs
  - State variable is match-specific idiosyncratic preference for every location

Model

I locations indexed by i, continuum of individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(\vec{\epsilon}_{nt}) = \max_i \{v_{it} + \epsilon_{int}\} + \beta \mathbb{E}[V_{nt+1}(\vec{\epsilon}_{nt+1}) | \vec{\epsilon}_{nt}]$$

  - No moving costs
  - State variable is match-specific idiosyncratic preference for every location
- Personal utility is persistent

$$\epsilon_{int} = \rho \epsilon_{in,t-1} + \left(\sqrt{1-\rho^2}\right) \eta_{int}$$

Model

I locations indexed by i, continuum of individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(\vec{\epsilon}_{nt}) = \max_i \{v_{it} + \epsilon_{int}\} + \beta \mathbb{E}[V_{nt+1}(\vec{\epsilon}_{nt+1}) | \vec{\epsilon}_{nt}]$$

  - No moving costs
  - State variable is match-specific idiosyncratic preference for every location
- Personal utility is persistent

$$\epsilon_{int} = \rho \epsilon_{in,t-1} + \left(\sqrt{1-\rho^2}\right) \eta_{int}$$

- Personal utility is spatially-correlated

$$\vec{\eta}_{nt} \sim N(0, \Sigma), \qquad \Sigma_{ij} = \exp(-A \text{ distance}_{ij})$$

Model

I locations indexed by i, continuum of individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(\vec{\epsilon}_{nt}) = \max_i \{v_{it} + \epsilon_{int}\} + \beta \mathbb{E}[V_{nt+1}(\vec{\epsilon}_{nt+1}) | \vec{\epsilon}_{nt}]$$

  - No moving costs
  - State variable is match-specific idiosyncratic preference for every location
- Personal utility is persistent

$$\epsilon_{int} = \rho \epsilon_{in,t-1} + \left(\sqrt{1-\rho^2}\right) \eta_{int}$$

- Personal utility is spatially-correlated

$$\vec{\eta}_{nt} \sim N(0, \Sigma), \qquad \Sigma_{ij} = \exp(-A \text{ distance}_{ij})$$

- Spatially and Persistently Auto-Correlated Epsilons (SPACE)

Model

I locations indexed by i, continuum of individuals indexed by n, and discrete time indexed by t:

- Agents choose location that maximizes utility

$$V_{nt}(\vec{\epsilon}_{nt}) = \max_i \{v_{it} + \epsilon_{int}\} + \beta \mathbb{E}[V_{nt+1}(\vec{\epsilon}_{nt+1}) | \vec{\epsilon}_{nt}]$$

  - No moving costs
  - State variable is match-specific idiosyncratic preference for every location

- Personal utility is persistent

$$\epsilon_{int} = \rho \epsilon_{in,t-1} + \left(\sqrt{1-\rho^2}\right) \eta_{int}$$

- Personal utility is spatially-correlated

$$\vec{\eta}_{nt} \sim N(0, \Sigma), \qquad \Sigma_{ij} = \exp(-A \text{ distance}_{ij})$$

- Spatially and Persistently Auto-Correlated Epsilons (SPACE)

Model can match all three facts

- Migration is rare
  - When ρ is high, the migration rate is low

Model can match all three facts

- Migration is rare
  - When ρ is high, the migration rate is low
- Gravity
  - When the correlation between an individual's i and j shocks are high, then people who live in i are more likely to be close to indifferent about living in j

Model can match all three facts

- Migration is rare
  - When ρ is high, the migration rate is low
- Gravity
  - When the correlation between an individual's i and j shocks are high, then people who live in i are more likely to be close to indifferent about living in j
- Square root fact

Proposition 2

As $\rho \to 1$, the t-year migration rate is proportional to $\sqrt{t}$

Parameterization

- Two parameters: persistence ρ, and spatial correlation A
- Target: 1-year migration rate, gravity equation
- Simulate 10 million people for two periods, fifty U.S. states and D.C.
    - v_i matches population

Matching the Facts Quantitatively

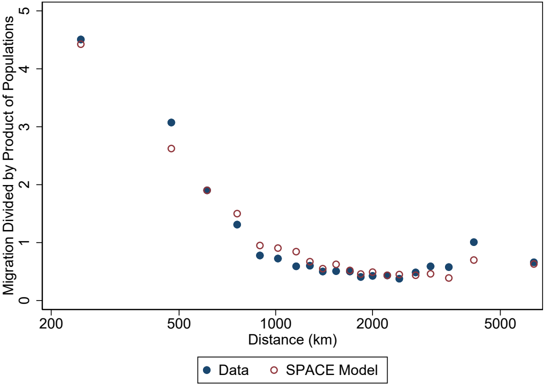
- Persistence: $\rho = .9996$,
- Spatial correlation: $A = .000299$ km^{-1}
- Targets (and hits) 3.34 percent migration rate

	(1) Migration (IRS)	(2) Migration (Credit)	(3) Simulated Migration
Log Distance	-0.736***	-0.744***	-0.744***
	(0.0572)	(0.0515)	(0.0396)
Log Origin Population	0.900***	0.923***	0.892***
	(0.0832)	(0.0797)	(0.0486)
Log Destination Population	0.822***	0.893***	0.889***
	(0.0976)	(0.0799)	(0.0501)
Observations	2550	2550	2550

Standard Errors are two-way clustered by origin and destination states

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Gravity

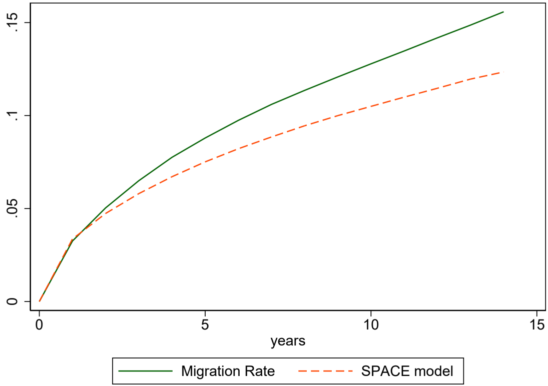


- Curvature is untargeted

Square Root Fact

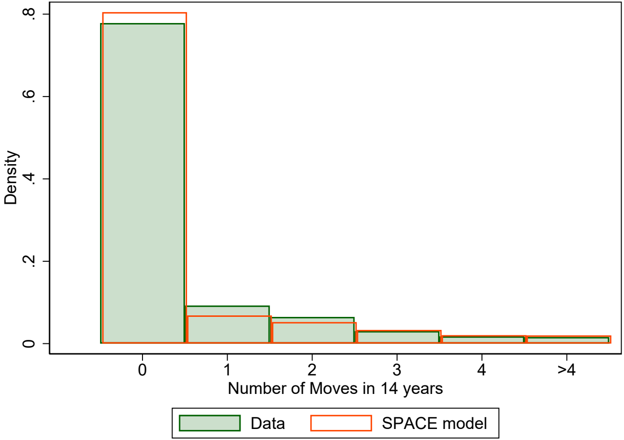
Proposition 2

As $\rho \to 1$, the t-year migration rate is proportional to $\sqrt{t}$.



5-year Calibration

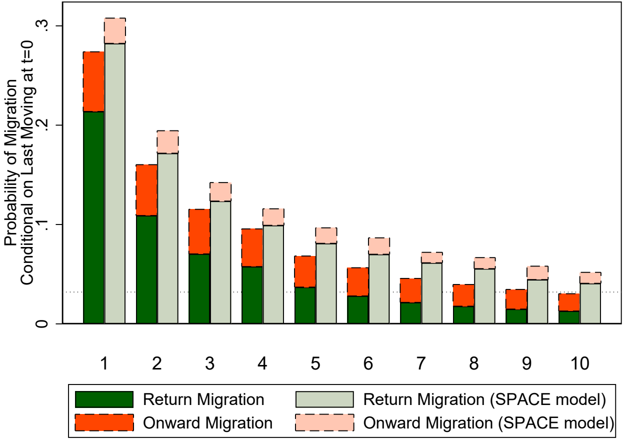
Frequency of Migration



5-year Calibration

Return Migration

- Conditional probability of moving after previous move



5-year Calibration

Implications of the Model

1. Population elasticities
2. Population dynamics
3. Moving Costs

Is the model useful for counterfactuals and welfare?

- Need to be able to calculate population elasticities
- One reason for skepticism: multinomial probits do not have a closed-form solution for these elasticities as a function of parameters

Is the model useful for counterfactuals and welfare?

- Need to be able to calculate population elasticities
- One reason for skepticism: multinomial probits do not have a closed-form solution for these elasticities as a function of parameters

Proposition 3

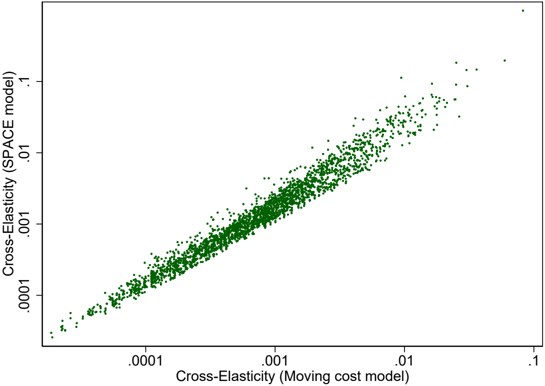
As $\rho \to 1$, the semi-elasticity of the population in i with respect to u_j is

$$\frac{\partial \log p_i}{\partial v_j} = -\lim_{\rho \to 1} \frac{m_{i \to j}}{p_i} \frac{1}{\sqrt{1 - \Sigma_{ij}}} \sqrt{\frac{\pi}{1 - \rho^2}}$$

- If you know migration, distance, and the parameters, sufficient to calculate these elasticities

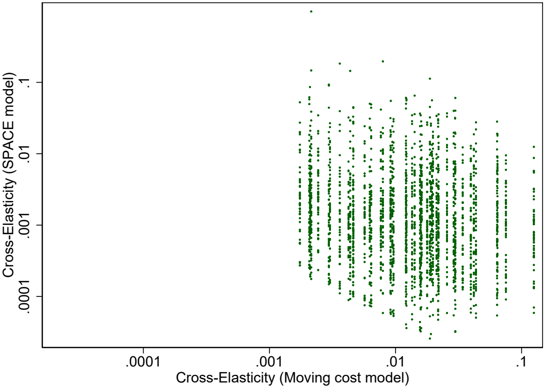
Short-run Population Cross-elasticity

- In the short-run, elasticities from moving cost model and SPACE model are quite similar
- Both primarily depend on gross migration

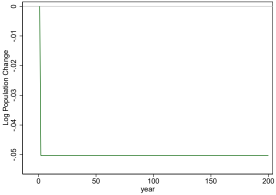


Long-run Population Cross-elasticity

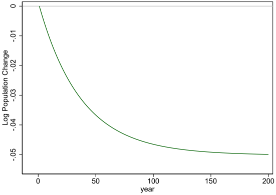
- In the long-run, very different
- Dynamic logit converges to static logit
  - Cross-elasticities proportional to population of shocked state



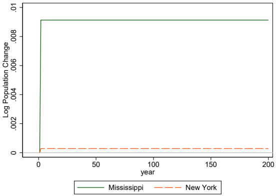
Population dynamics after a one-time permanent change in $v_{\text{Louisiana}}$



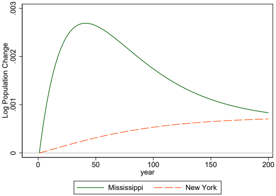
(a) Louisiana, SPACE model



(b) Louisiana, Dynamic Logit model



(c) Mississippi and New York, SPACE model



(d) Mississippi and New York, Dynamic Logit model

Aggregate Dynamics

- In dynamic logit, permanent utility increase in j increases the migration rate into j, population adjusts very slowly
- In SPACE, permanent utility increase in j raises population immediately
- How does SPACE interpret persistent migration flows?
  - Persistent utility changes from feedback from the housing market, labor market adjustments, search, etc.

Different Implications of SPACE and Dynamic Logit

1. Moving costs need not be large
2. Implied utility changes
3. Aggregate dynamics
4. Long-run Population Elasticities

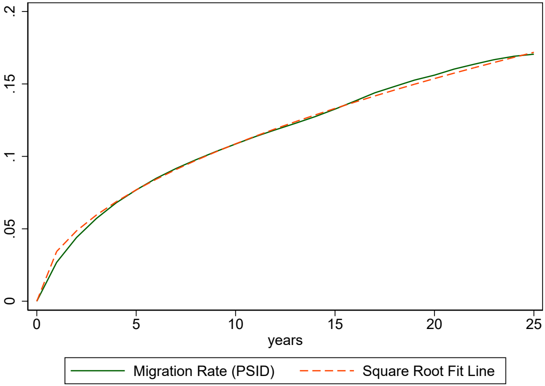
Moving Costs Need Not Be Large

- Kennan and Walker (2011): average moving cost is $312,146 (in 2010 USD)
  - $0 in SPACE model
- Common counterfactual is to change moving costs in some way

Takeaways

1. New fact: t-year internal migration rates are proportional to $\sqrt{t}$
2. Persistent preferences match dynamic moments of migration *and* gravity
3. Persistent preferences has different implications for long-run adjustments, population dynamics, and estimates of moving costs

Square Root Rule, PSID

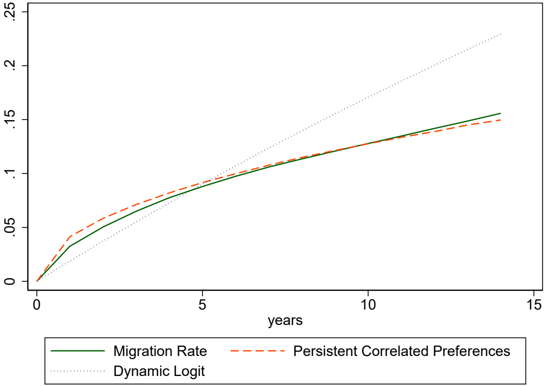


Return

Square Root Rule, 5-year calibration

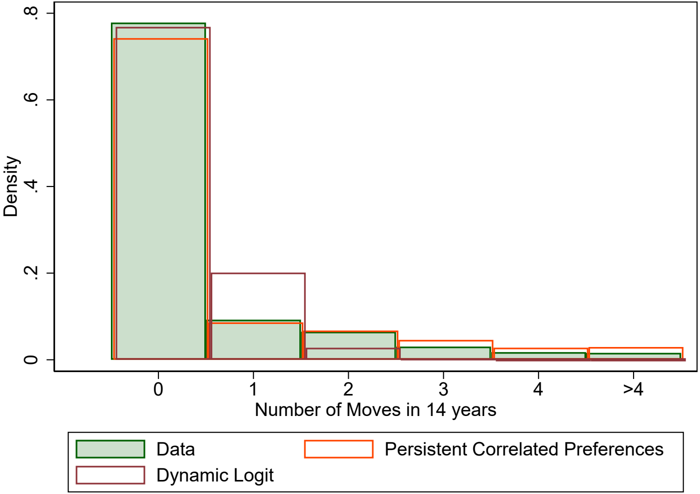
Proposition 2

As $\rho \to 1$, the t-year migration rate is proportional to $\sqrt{t}$.



Return

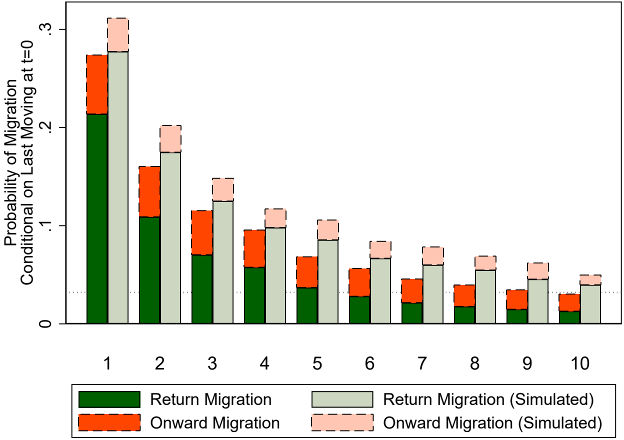
Frequency of Migration, 5-year calibration



Return

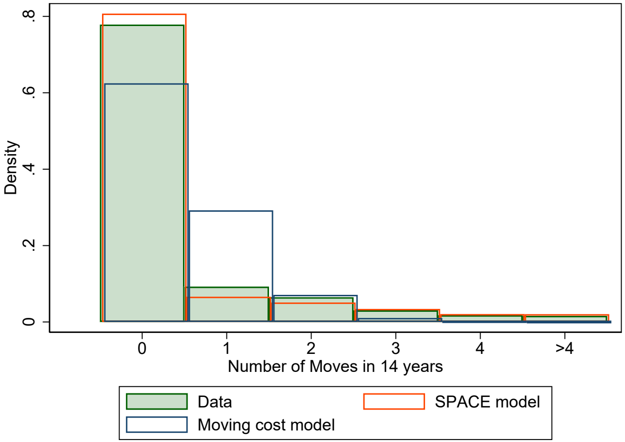
Return Migration, 5-year calibration

- Conditional probability of moving after previous move



Return

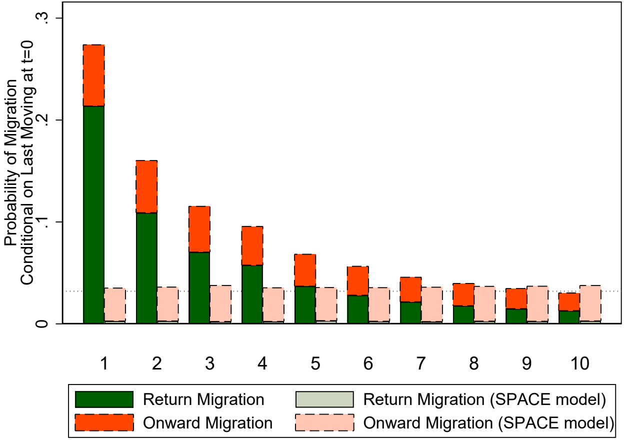
Frequency of Moves, Dynamic Logit Model



Return

Return Migration, Dynamic Logit Model

- Conditional probability of moving after previous move



Return

Bibliography

Allen, Treb and Dave Donaldson, "Persistence and path dependence in the spatial economy," 2020. National Bureau of Economic Research Working Paper.

Amior, Michael and Alan Manning, "The persistence of local joblessness," *American Economic Review*, 2018, *108* (7), 1942–70.

Bayer, Christian and Falko Juessen, "On the dynamics of interstate migration: Migration costs and self-selection," *Review of Economic Dynamics*, 2012, *15* (3), 377–401.

Blanchard, Olivier Jean and Lawrence F Katz, "Regional evolutions," *Brookings Papers on Economic Activity*, 1992, *1992* (1), 1–75.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Butler, John S and Robert Moffitt, "A computationally efficient quadrature procedure for the one-factor multinomial probit model," *Econometrica*, 1982, pp. 761–764.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 2019, *87* (3), 741–835.

Coen-Pirani, Daniele, "Understanding gross worker flows across US states," *Journal of Monetary Economics*, 2010, *57* (7), 769–784.