Behavior Modeling Summer School 2026

Where Will You Move? or Stay?

A Residential Choice Analysis considering Societal Attachment

Tokyo, September 7-9, 2026

Hiroshima University – Team B

- Damar Pratita Dewayanto
- Kevianwillie Handoyo
- Melanton Hendra Siregar
- Rizqi Ramadhan

Table of Contents

1

Background

Background

126,515

Net in-migrants into Tokyo region in 2023

-1.31%

Tohoku population decline, fastest since 2001.
Akita, Aomori, Iwate lead nationwide

70,000

2011 earthquake. People from Iwate, Miyagi & Fukushima evacuated outside their home prefectures

Tohoku is shrinking, but why some people leave while others stay is not obvious.

Source: 2026 Summer School distributed dataset - Tohoku panel data, n=3,037 individuals, 2005-2024; Report on Internal Migration Based on the Basic Resident Registration, 2023 Results (published Jan 2024), p.11, Section IV-1. Statistics Bureau of Japan, Ministry of Internal Affairs and Communications. https://www.stat.go.jp/data/idou/2023np/jissu/pdf/gaiyou.pdf; https://www.bousai.go.jp/kaigirep/hakusho/h25/honbun/1b_1s_02_01.htm

Research Question

What factors determine an individual's decision to relocate or stay in a place?

Objective

- Measuring the pull of each location to know how attractive a zone is
- Measuring the resistance in moving to the other region
- Measuring whether peers affected decisions to choose location of residence

2

Data

Data Characteristic

Brief description: A **panel data** consisting **residential history** of survey respondents.

Observation Period: **2005 – 2024**

Number of Individuals: **3,037 individuals** (410 evacuee + 2,627 non-evacuee)

Alternatives: **10 cities** in Tohoku and Kanto region

Data used:

initial.csv

Initial state of individual at the start of observation

df.history.csv

Individual states and changes over 20 years

stay_var.csv

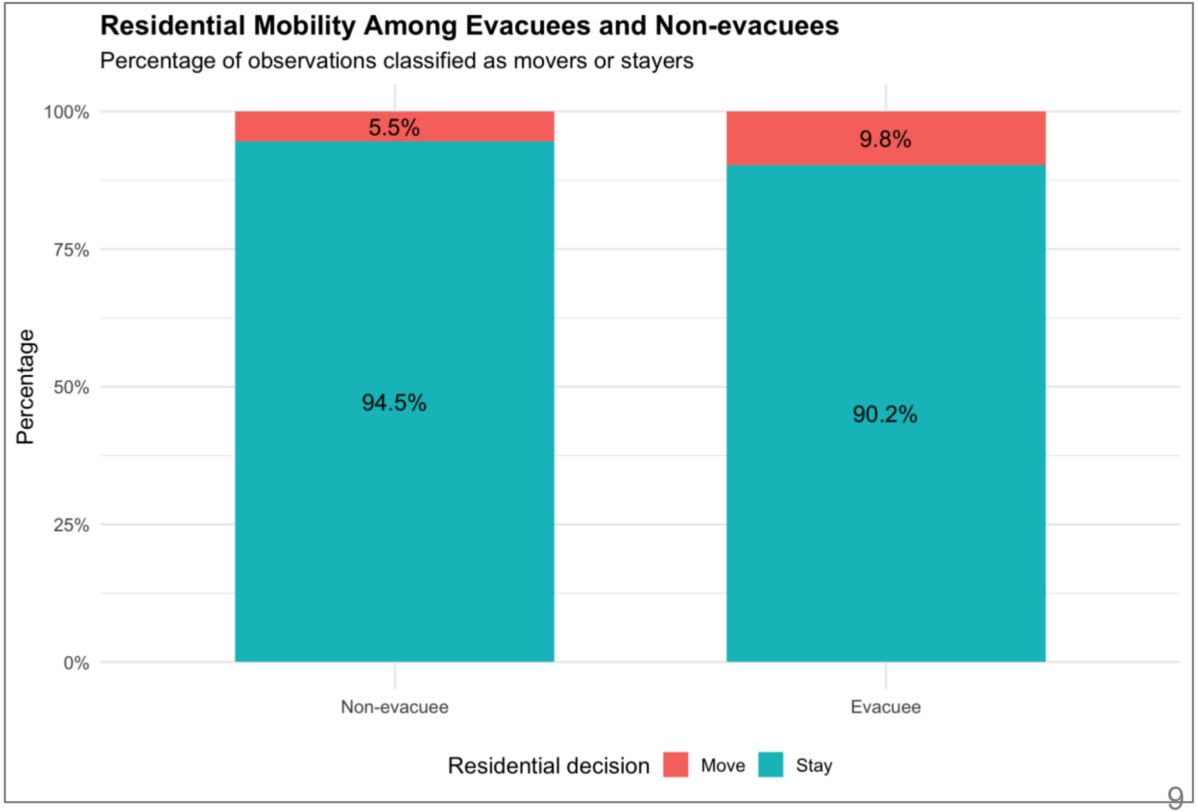
Variables that describes individual’s attachment to current location (activity, family proximity, duration of staying in current residence.)

ex_var.csv

Variables that describes all other destination’s attractiveness (distance from current residence location to each alternative, density of each alternative, commute distance).

A little context before we proceed…

Naturally, most people would choose to stay in their current residence.



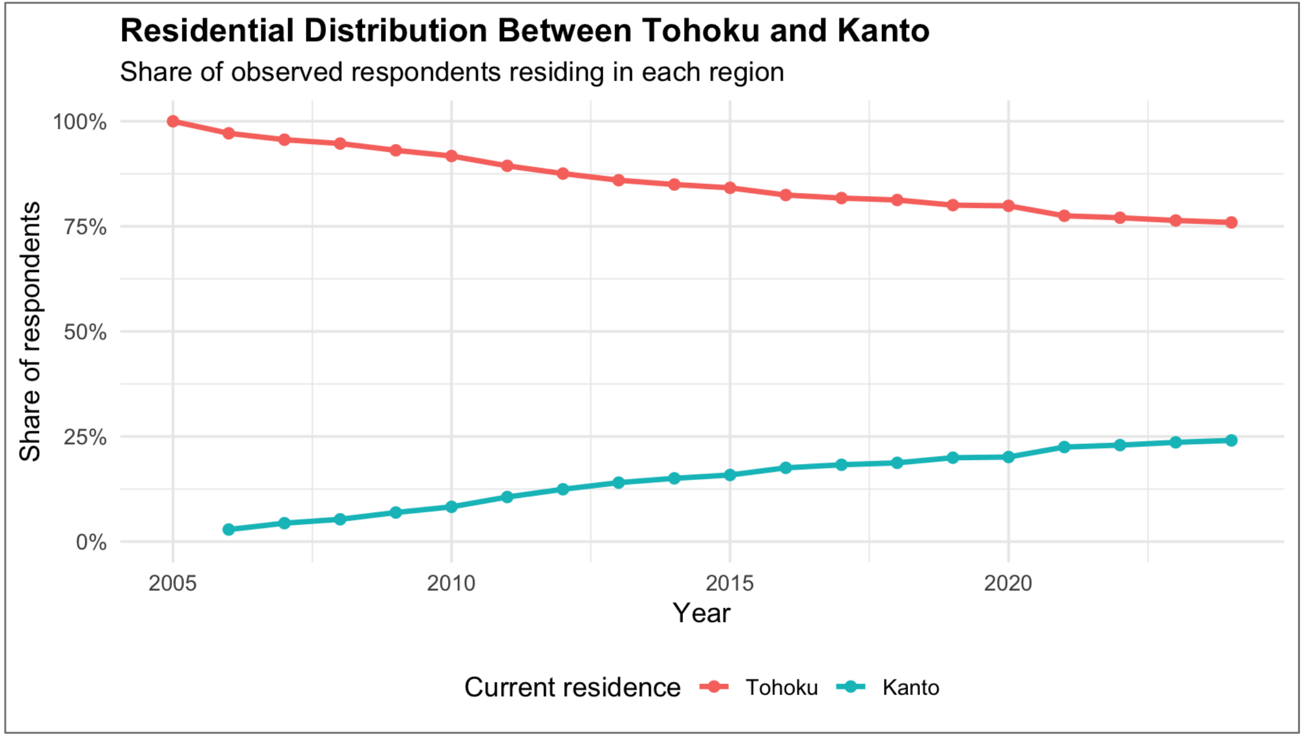
Only 38.7% of moves are inter-municipality

Share of type of moves

Types of Move	Share
Inter-Municipality	38.7%
Other	61.3%

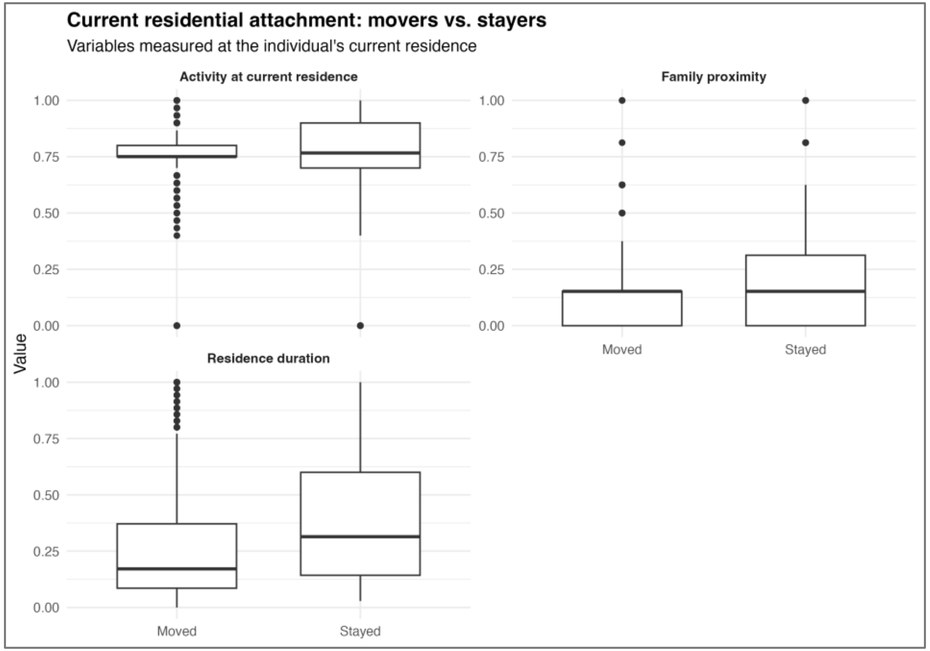
What does the data show?

Respondents residing in Tohoku keeps declining in each year, and Tohoku out-migration rate is consistently above in-migration rate from Kanto.



Why do people move or stay?

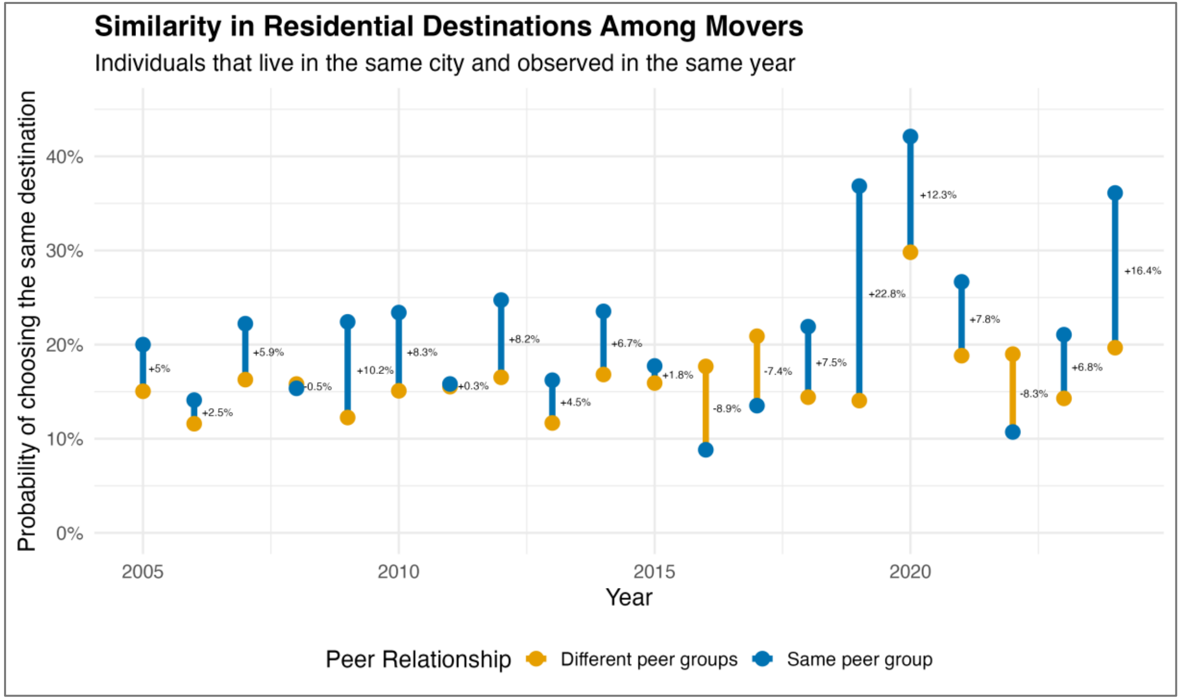
Residential attachment's characteristic between movers and stayers is somewhat different, especially for residence duration.



Individuals that **chooses to stay** tend to have **longer time staying** in their current residence than those who moves, which may suggest association between residential attachment and residential location choice.

Why do people move or stay?

Movers tend to choose similar residential destination with individuals in the same peer group rather than different peer group.



In 16 of 20 years of observation, movers in the **same peer group has higher probability to choose the same residential destination** than different peer groups. This finding motivates testing whether peer behavior influences residential location choice.

How did we calculate similarity?
Pairwise comparisons were conducted among movers sharing the same city and year. Pairs were classified as same- or different-peer-group and same- or different-destination, and the share of same-destination pairs was compared across peer-group types.

3

Model

Nested Logit – Mathematical Setup

$$J = \bigcup_B J_B\,, \qquad J_B \cap J_{B'} = \emptyset \text{ for } B \neq B'$$

$U_{ijB} = V_{ijB} + \varepsilon_{ijB}, \qquad j \in B$ # utility function

$\varepsilon_{ijB} \sim$ GEV, allowing correlation within nest B, independence across nests

$P_{ijB} = P_{iB} \cdot P_{ij|B}$ # choice probability

$$P_{ij|B} = \frac{\exp\left(V_{ijB}/\lambda_B\right)}{\sum_{k\in B} \exp(V_{ikB}/\lambda_B)}$$ # conditional (within-nest) probability

$$P_{iB} = \frac{\exp(W_{iB} + \lambda_B\, IV_{iB})}{\sum_{B'} \exp(W_{iB'} + \lambda_{B'}\, IV_{iB'})}$$ # upper-level (nest) probability

$$IV_{iB} = \ln \sum_{k\in B} \exp(V_{ikB}/\lambda_B)$$ # inclusive value (Logsum) of nest B

$0 < \lambda_B \leq 1$ # nesting parameter constraint

Nested Logit Specifications

We will use **nested-logit model**: **upper nest** is stay vs. move; **lower nest** is which of 9 zones when moving; connected by a dissimilarity parameter λ. ***Move*** here is defined as the zone-changing displacement (`target != residence`).

- `λ = 1` collapses this to the old two-part model, verified to 1.9e-9 log-likelihood agreement against `src.models.mnl.fit_mnl()` on the identical feature set, which means load-bearing correctness check for the whole estimator.
- **Row grain**: person-PERIOD-at-risk (8,113 occasions, 3,037 people); the old design used person-year (60,740 rows). A person contributes one occasion per period (`pre:` 2005–2010; `shock:` 2011–2012; `post:` 2013–2024) until their first zone-change, then exits the risk set.
- Choice set: the person's own zone (`` `stay` ``) plus the other 9 (`` `move` ``). `` `s` `` = own zone, `M` = the 9 move zones.
- 17 parameters (5 generic + 11 move-nest + λ) on 8,113 occasions: 477 observations per parameter. `` `peer_share` `` enters all 10 alternatives, including `` `stay` ``. Every other move-nest term is gated by $move\_cost = \mathbb{1}[j \neq s]$, so it shifts the stay/move margin alone.

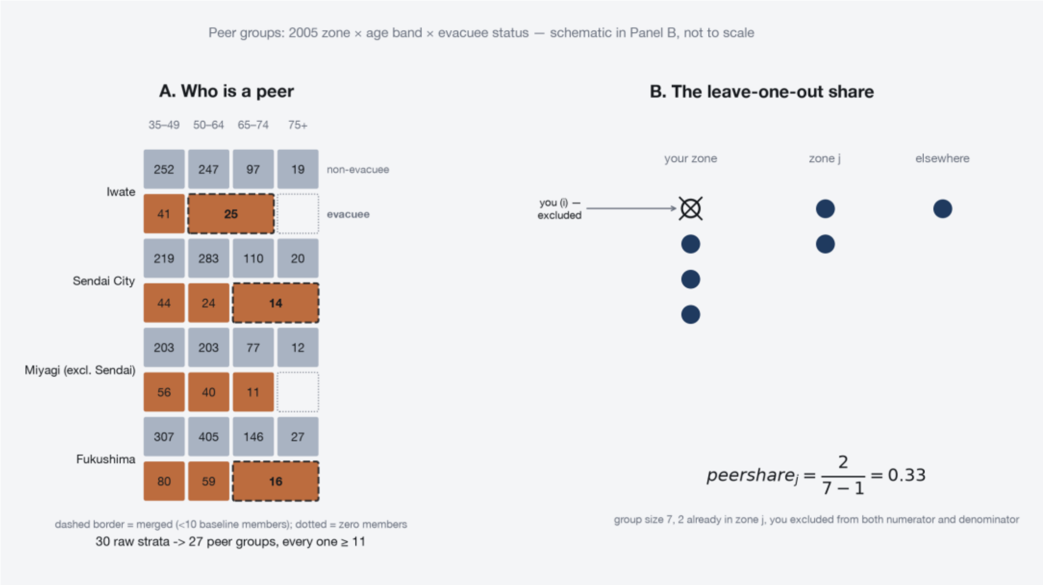
$$V_j = \beta_{dist}\, distance_j + \beta_{dens}\, density_j + \beta_{mcom}\, manCommute_j + \beta_{wcom}\, womanCommute_j + \beta_{peer}\, peerShare_j$$
$$+\mathbb{1}[j \neq s]\big(\alpha_0 + \alpha_{dis} disaster + \alpha_{sh} shock + \alpha_{po} post + \alpha_{ds}(disaster \times shock) + \alpha_{dp}(disaster \times post) + \alpha_{own} own + \alpha_{dur} dur + \alpha_{act} act + \alpha_{fam} fam + \alpha_{age} age\big)$$

$$(\text{move}) = \frac{e^{\lambda \cdot IV}}{e^{V_s} + e^{\lambda \cdot IV}}, \quad IV = \ln \sum_{j \in M} e^{V_j/\lambda}, \quad P(j \mid \text{move}) = \frac{e^{V_j/\lambda}}{\sum_{k \in M} e^{V_k/\lambda}}$$

What counts as a "peer," and is the effect exogenous

We build peer groups from 2005 **zone, age band,** and **evacuee status**. The grouping is finer than prefecture and coarser than city. We treat peer share as predetermined and contextual, exogenous by construction. We read it from the data and never solve for it as a self-consistent belief.

This peer construction follows Dugundji and Gulyas (2013) who build groups the same way under the same assumptions: *people are influenced by others of similar background regardless of exact location.* However, we differ by using *leave-one-out* which excludes the respondent’s own choice



Panel A: the real 2005 grid (30 strata, 3 merges, 2 empty cells), live from `build_persons()`.
Panel B: a schematic of the leave-one-out mechanic.

$$peer\_share_{i,j,t} = \frac{n_{j,t}^{(g)} - \mathbb{1}[z_{it} = j]}{|g| - 1},$$
$$g = zone_i^{2005} \times age\_band_i \times evacuee_i$$

- Group = 2005 `zone` × `age band` × `evacuee status` → 27 groups (30 raw strata; we merge any stratum under 10 members into a neighboring age band, same zone and status)
- Share = leave-one-out share of your group already in zone j, excluding you
- Zone, not prefecture, not city: Sendai City differs enough from the rest of Miyagi to split, and 3,037 people can't fill city-level groups
- Predetermined, not solved: fixed at 2005, read from the data. This is the contextual, exogenous kind in Manski (1993)
- What that costs: we can show association with a predetermined share, not real-time contagion

4

Result

Headline result: λ is real, and the honest ρ^2

$\hat{\lambda}$ = 0.502, well below 1 and inside (0, 1], the destination choice and the stay/move decision are linked. The 0.690 adjusted ρ^2 **overstates** this: `move_asc` alone scores 0.62 against the honest null. The other 15 parameters add about 0.07

Table 1. Full results, headline spec

	Parameter	t-value
Constant (move vs. stay)	1.77	2.97
Distance to zone [normalized 0–1]	-0.03	-0.18
Population density at zone [normalized 0–1]	0.88	9.36
Male commute distance to zone [normalized 0–1]	-1.13	-14.30
Female commute distance to zone [normalized 0–1]	-0.93	-10.74
Peer group's leave-one-out share already in zone [0–1]	3.00	5.05
Evacuee status, on choosing to move	-1.29	-6.54
Shock period (2011–2012), on choosing to move	-2.10	-11.13
Post period (2013–2024), on choosing to move	-0.17	-0.99
Evacuee × shock period, on choosing to move (DiD)	2.94	10.83
Evacuee × post period, on choosing to move (DiD)	1.59	6.25
Home ownership [0–1], on choosing to move	-1.52	-15.96
Duration of residence [normalized 0–1], on choosing to move	-0.74	-4.81
Community activity level [normalized 0–1], on choosing to move	-0.17	-0.54
Proximity to family [normalized 0–1], on choosing to move	-1.07	-2.83
Age [decades], on choosing to move	-0.11	-2.85
Nesting parameter λ (move nest, t vs. 1)	0.50	-12.13
Number of observations	8,113	
Initial log-likelihood (uniform null)	-18,680.9	
Final log-likelihood	-4,244.7	
Rho-squared	0.692	
Adjusted rho-squared	0.690	

Table 2. Fallback vs. headline (lambda)

	λ = 1 (fallback)	Headline: λ free
Log-likelihood	-4275.2	-4244.7
λ	1 (fixed)	0.502 (se 0.041)
t(λ vs. 1)	n/a	-12.13
N	8,113	8,113

$\hat{\lambda}$ = 0.502 confirmed stable from 7 different optimizer starting points (λ_0 = 0.1 to 3.0, all converge to 0.5022).

λ's t-test runs against 1: λ = 1 means no nesting structure at all, the null worth testing, following Dugundji & Gulyás (2013, Table 3)'s reporting convention.

`move_asc` is overstated because it is the constant on "move vs. stay", the person-specific status-quo dummy. Fit alone, it scores adjusted ρ^2 = 0.62 against the same market-share null.

Rho-squared/Adjusted rho-squared are measured against the market-share (constants-only) null. Initial log-likelihood is the uniform null (all coefficients = 0), the textbook Biogeme convention: the two rows use different baselines on purpose.

Societal attachment: what anchors someone to a zone

Home ownership, years of residence, family proximity, and age **each pull against moving on their own**, all fitted inside the same nested-logit move-nest branch as the destination attributes. Ownership is **the strongest anchor**; community activity level is the one that **doesn't reach significance**.

Model applied: the same nested logit, reading the move-nest block of the utility equation. Each of these four terms is gated by `move_cost`, so it shifts a single margin: whether this person moves at all. Which *zone* they'd choose is governed by the *generic terms* alone (`distance`, `density`, `commute`, `peer share`).

TABLE 3. Attachment factors

Attachment factor	Coefficient	t-value	Read
Home ownership [0–1]	-1.52	-15.96	strongest anchor
Duration of residence [normalized]	-0.74	-4.81	anchor, moderate
Proximity to family [normalized]	-1.07	-2.83	anchor, moderate
Age [decades]	-0.11	-2.85	anchor, small but real
Community activity level [normalized]	-0.17	-0.54	not significant

All four significant terms carry the expected sign: **more attachment, lower probability of moving**. Community activity level is the outlier, both smaller in magnitude and indistinguishable from zero in this sample.

Equilibrium check, the survival-bias caveat, and equivalence unit

Our peer share is predetermined, so the estimation never faces the equilibrium-multiplicity problem self-consistent models do; a stress test still lands well inside the safe region. What this design cannot rule out is people already sharing a 2005 zone may share an unobserved reason for having settled there together, a correlated-effects concern separate from the exogenous/endogenous question.

Equilibrium check

- `peer_share` is a predetermined 2005-baseline share, not solved as a self-consistent fixed point → outside Brock & Durlauf's (2003) multiple-equilibria problem
- Stress test against their threshold: multiple equilibria risk when β·J/L > 1
- $\frac{\beta_{peer}}{L} = \frac{3.00}{10} = 0.30 \quad \text{(under the 1.0 threshold)}$
- Compare Dugundji & Gulyás (2013): β≈ 2.79, L=3 → ratio ≈0.93, near 1, why they had to check for multiplicity
- Our 10-zone choice set sits inside the well-behaved region → Slide 6's policy sim is well-posed

Survival-bias / correlated-effects caveat (Manski's third category)

- Identification requires errors independent of individual and group characteristics (Brock & Durlauf condition vii)
- 2005 baseline rules out reverse causation: a later move can't change 2005 group membership
- Doesn't rule out a shared unobserved reason for the original 2005 location: family business, cultural tie, shared risk tolerance. Dugundji & Gulyás (2013, p.193) flag residential choice as the hard case
- An unsolved limitation for the model

Equivalence unit

- IQR increase in `peer_share` among post-period movers (≈0.059) equals about 0.16 of a full 0→1 swing in commute distance, same normalized scale: a modest social-pull effect
- `distance_array` (the planned km-equivalence ruler) is unusable: indistinguishable from zero (Slide 3). A first attempt priced `peer_share` at −62,358 "km"; we discarded it before it reached a slide
- No real-world (minutes) scale exists for `man_commute`, so this stays normalized-scale only

5

Policy Recommendation

Policy simulation - Promoting home ownership

Renters converted	Ownership rate	P(stay)	P(move)
0 % (baseline)	50.0%	50.8%	49.2%
10%	55%	52.3%	47.7%
25%	62.5%	54.5%	45.5%
50%	75.0%	58.3%	41.%7
100%	100%	65.8%	34.2%

Highlight result

+15 points

aggregate move probability, baseline vs. full homeownership (49.2% → 34.2%). A moderate 25% renter-conversion target cuts mobility to 3.7 points.

Mechanism: coefficients fixed at the estimated model (γ_own = −1.52, λ = 0.50); simulation converts a share of renting households to owners (0%, 10%, 25%, 50%, and 100% and recomputes move probability

Creating A Better Condition

1 **Strengthen Housing Ownership**
Enhance the possibility to own houses in the region.

2 **Strengthen Community Activities**
Encourage community activities to strengthen residents' connection to their local area

Supporting home ownership and encourage community activities to help residents stay connected to their local area.

6

Conclusion

Conclusion

1 Home ownership is the strongest factor encouraging residents to stay.

2 Longer residence and closer family ties also support staying decisions

3 Promoting housing ownership by converting renters to owners cuts move probability by up to 15 percentage points.

References

Brock, W. A., & Durlauf, S. N. (2001). Discrete Choice with Social Interactions. *The Review of Economic Studies*, *68*(2), 235–260. https://doi.org/10.1111/1467-937X.00168

Brock, W. A. and Durlauf, S. N. (2003). Multinomial Choice with Social Interactions. *NBER Working Paper, t0288.* https://ssrn.com/abstract=385026

Dugundji, E. R., & Gulyás, L. (2013). Structure and emergence in a nested logit model with social and spatial interactions. *Computational and Mathematical Organization Theory*, *19*(2), 151–203. https://doi.org/10.1007/s10588-013-9157-y

2026 *Summer School distributed dataset - Tohoku* panel data, n=3,037 individuals.

Statistics Bureau of Japan, Ministry of Internal Affairs and Communications. (2013). *Disaster Prevention White Paper 2013 Edition*. https://www.bousai.go.jp/kaigirep/hakusho/h25/honbun/1b_1s_02_01.htm.

Statistics Bureau of Japan, Ministry of Internal Affairs and Communication. (2025). *Report on Internal Migration Based on the Basic Resident Registration 2005–2024*, 2023 Results (published Jan 2024), p.11, Section IV-1. https://www.stat.go.jp/data/idou/2023np/jissu/pdf/gaiyou.pdf.

Thank you!