The Effect of Expected Income on Individual Migration Decisions

John Kennan and James R. Walker

Econometrica, 2011

Jed Armstrong, Spatial Economics Reading Group, Apr. 08 2020

Previous models of migration choice

- Blanchard & Katz (1992) Income differentials play a large role in inter-state migration decisions
- Gallin (2004) Migration as a function of expected wages, but individual decision problem un-modeled
- Holt (2009) Dynamic discrete choice of migration, but move / stay not location-specific
- Dahl (2002) Many alternatives but only a single life-time migration choice

Motivation

Interstate Migration, NLSY 1979-1994

Movers (%)	32.3
Moves per mover	1.98
Repeat moves (% of moves)	49.5
Movers returning home (%)	50.2

- Migration decisions aren't one-off or isolated
- Need a dynamic model to study migration decisions

Overview of the paper

- Tractable dynamic search model of migration decisions based on wage differentials
- Estimated using NLSY79

- Expected income change, distance, home / previous location, population size, age, climate
- Empirical estimate of moving costs (\$312k+) and home premium (\$23k)

Model – environment

- Finitely-lived individuals i
- Get linear utility from wages and amenities in one of J locations
- Pay a moving cost to try a new location
- Make migration decisions to maximize lifetime utility
- Wage earned in l is best offer available – only chance of permanent wage change is l'

Model – recursive structure

$$V(x,\zeta) = \max_j \left(u(x,j) + \zeta_j + \beta \sum_{x'} p(x'|x,j)\, \mathbb{E}_{\zeta'} V(x',\zeta') \right)$$

- Finite horizon Bellman equation
- State vector is x – current and previous locations $(l^0, l^1, \ldots)$, age, home location
- j is the location choice
- ζ_j is preference / moving-cost shock $\sim$ type I extreme value

Model – specification

- Individual wakes up aged a in location l^0, gets paid and enjoys amenities, decides where to move j in the evening

$$u(x,j) = \alpha_0 w(x) + \sum_k \alpha_k A_k(x) + \alpha^H \mathbb{1}_{\{l^0=h\}} - \Delta_\tau(x,j) + \xi_{l^0}$$

- ξ_{l^0} is an individual-location utility fixed effect
- Moving costs $\Delta_\tau(x,j)$ depend on distance, adjacency, recent occupation, etc.

$$w_i(x) = \mu_{l^0} + \eta_i + \nu_{il^0} + G(X_i, a, t) + \varepsilon_i(x)$$

- $G(X_i, a, t)$ is time-trend + effects of observables

Model – identification of income effect

$$w_i(x) = \mu_{l^0} + \eta_i + \nu_{il^0} + G(X_i, a, t) + \varepsilon_i(x)$$

- Wage effect on location decisions $\mu_{l^0} + \nu_{il^0}$

- Probability of moving depends only on ν: $\rho(\text{move}|\nu)$
- Shape of $\rho(\text{move}|\nu)$ determines income effect on migration
- Try to identify the choice probability by Bayes theorem
$$f_\nu(\nu|\text{move}) = \frac{\rho(\text{move}|\nu)f_\nu(\nu)}{\Pr(\text{move})}$$

Model – identification of income effect

- Just need to identify distributions of ν for movers and stayers
- Use Kotlarski's lemma for contaminated observations

Let X_1, X_2, and θ be three independent real-valued random variables and define $Y_1 = X_1 + \theta$ and $Y_2 = X_2 + \theta$. Then the joint distribution of (Y_1, Y_2) determines the distributions of X_1, X_2, and θ.

- Define the wage residual for i at t in location j as

$$y(i,j,t) = w(i,j,t) - \mu(j) - G(X_i, a, t) = \eta(i) + \nu(i,j) + \varepsilon(i,j,t)$$

Model – identification of income effect

$$y(i,j,t) = \eta(i) + \nu(i,j) + \varepsilon(i,j,t)$$

- For stayer i in location j

$$y(i,j,t) = \underbrace{\eta(i) + \nu(i,j)} + \varepsilon(i,j,t)$$
$$y(i,j,t') = \underbrace{\eta(i) + \nu(i,j)} + \varepsilon(i,j,t')$$

Kotlarski means that distributions of $\eta + \nu$ and ε are identified!

Model – identification of income effect

$$y(i,j,t) = \eta(i) + \nu(i,j) + \varepsilon(i,j,t)$$

- For mover i from j to j'

$$y(i,j,t) = \underbrace{\eta(i)} + \nu(i,j) + \varepsilon(i,j,t)$$
$$y(i,j',t') = \underbrace{\eta(i)} + \nu(i,j') + \varepsilon(i,j',t')$$

Kotlarski means that distributions of η and $\nu + \varepsilon$ are identified!

Model – identification of income effect

$$f_\nu(\nu|\text{move}) = \frac{\rho(\text{move}|\nu) f_\nu(\nu)}{\Pr(\text{move})}$$

- By de-convolution, can identify $f_\nu(\nu|\text{move})$ and $f_\nu(\nu) \to \rho(\text{move}|\nu)$

Empirical implementation

- Fact that wage information in all past locations is known makes state space grow
- With J locations and n points of support for each wage distribution $J(n+1)^J$ states for each person

- Assume only recently-observed (last $M < J$) locations are known
- Individuals with identical recent histories have same state, irrespective of older histories
- M-vector l of locations in state

Data

- NLSY79 white non-Hispanic, HS grads, no post-secondary education
- $n = 432$ people, 4,274 person-years, 123 interstate moves
- μ_j estimated with PUMS data from 1990 Census

Estimation - distribution assumptions

- Wage residual $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$
- Use constant-probability discrete distributions of ν, η, ξ, σ_ε
- e.g. for ν_{ij} the distribution with $n_\nu = 3$ is

$$\Pr(\nu_{ij}) = \begin{cases} 0 & \text{w.p. } 1/3 \\ \tau_\nu & \text{w.p. } 1/3 \\ -\tau_\nu & \text{w.p. } 1/3 \end{cases}$$

where τ_ν is a parameter to be estimated

Estimation – likelihood function

- If individual i visits N_i locations in the data there are $n_\eta n_\sigma (n_\xi n_\nu)^{N_i} = 7{\cdot}4{\cdot}(3{\cdot}3)^{N_i}$ values ω_i, all equally likely
- A draw is a wage-location pair – get information from both

$$\varepsilon_{it}(\omega^i|\theta_\tau) = w_{it} - \mu_{l^0(it)} - G(X_i, a_{it}|\theta_\tau) - \nu(\omega^i) - \eta(\omega^i)$$

$\rho(l(i,t), \omega^i|\theta_\tau)$ is CCP of $l(i,t)$

- MLE by integrating over all ω_i

Estimation – likelihood function

- Given parameters θ, likelihood of an individual history for a person of type τ is

$$L_i(\theta_\tau) = \frac{1}{7 \cdot 4 \cdot (3 \cdot 3)^{N_i}} \times \sum_{\omega_i} \left(\prod_{t=1}^{T} \Pr(\varepsilon_{it}(\omega^i|\theta_\tau))\rho(I(i,t),\omega^i|\theta_\tau) \right)$$

- Types τ for 'movers / stayers' (Johnston, 1971)
- If fraction π_τ is of type τ then the sample log likelihood is

$$\Lambda(\theta) = \sum_{i=1}^{N} \log \left(\sum_{\tau=1}^{K} \pi_\tau L_i(\theta_\tau) \right)$$

Results – goodness of fit

GOODNESS OF FIT

Moves	Binomial		NLSY		Model	
None	325.1	75.3%	361	83.6%	36,257	83.9%
One	91.5	21.2%	31	7.2%	2,466	5.7%
More	15.4	3.6%	40	9.3%	4,478	10.4%
Movers with more than one move	14.4%		56.3%		64.5%	
Total observations	432		432		43,201	

Results – goodness of fit

RETURN MIGRATION STATISTICS

Movers	NLSY	Model
Proportion who		
Return home	34.7%	37.5%
Return elsewhere	4.8%	6.2%
Move on	60.5%	61.9%
Proportion who ever		
Leave home	14.4%	13.7%
Move from not-home	40.0%	42.5%
Return from not-home	25.7%	32.3%

Results

- Home premium is \$23,000 per year
- Average move away from a bad match increases income \$8,400. Move to better state increases by \$9,500
- Need high moving costs to fit lack of movement (NPV of gain is \$312,000)

Young (20)	\$384,743
'Old' (24.4)	\$312,146

Contribution

- First tractable model of lifetime migration decisions
    - Improves on previous literature by explicitly modeling choice procedure and increasing time- and location-dimensions
- First structural estimation of moving costs, but implausibly(?) high (\$312K+)
- First structural estimation of home premium