Reassessing Migration and Health: Analysis Using U.S. County Level Data

Animals weigh the costs and benefits associated with available habitats when making migration decisions (Hutto et al.; Gilliam et al.). Habitat selection models are a staple in biology, but do they apply to human migration behavior. The central idea in rational theories of human migration is that individuals and families do weigh the costs and benefits of their location options and migrate when the benefits outweigh the costs (Molloy et al. 2011). We argue that average life expectancy encapsulates survival benefits of a certain location but would be correlated with other contributions to survival like labor opportunities (unemployment rates) as well as other costs such as tax rates. All things being equal, we would expect to see out-migration from locations with low average life expectancies and in-migration into locations with high average life expectancies.

We use empirical data to test this simple model of migration in the context of internal migration (across counties) in the United States. Counties in the United States show a wide variation in terms of in and out migration, average life expectancies, tax rates, unemployment as well as the availability of amenities (Wang et al. 2013, Desmet and Fafchamps 2005). The use of life expectancy as a measure of benefits at the county level is supported by its universal acceptance as an important component of human development and its close association with measures of well-being (Arora et al. 2016). While migration models differ in the treatment of factors that determine costs they usually incorporate costs of living, probability of employment, local amenities and tax rates (Molloy et al. 2011). We will be controlling for factors such as unemployment, poverty, owner-occupied housing, urbanization, race distributions and education levels in our empirical model.

Our primary sources of data include data on returns filled from the Internal Revenue Services (IRS 2018) to compute the number of in-migrants and out-migrants and life expectancy measures at the county level maintained by the Institute of Health Metrics and Evaluation (IHME 2015) for the years 1994 to 2010. We estimate regression models both with and without lagged predictors (1-year, 2-year and 3-year lags) to ensure that non-immediate effects are captured. Results from preliminary analyses have been presented in Tables 1-3 below. Table 1 describes our key dependent and independent variables. Table 2 estimates an example of a fixed-effects regression specification at lag 1. This includes models estimated for all counties as well as sub-group analysis carried out for (i) Regions (North-East - NE, Mid-West - MW, South - S and West - W), (ii) Quartiles of Urban Population Proportions, (iii) Quartiles of Hispanic Population Proportions, (iv) Quartiles of Median Family Incomes, and (v) Quartiles of College Education Population Proportions. For the sake of brevity, we have shown specifications only for all counties. We carry out similar analyses for the other sub-groups described. The results across all specifications run are summarized in Tables 3.

We find that our data does not support the predictions of the model of migration or the Rational Habitat Selection Model. We find that the association between out-migration and life expectancy as well as that between in-migration and life expectancy differs by quartiles of counties stratified

according to urbanization, education, Hispanic populations and incomes as well as across regions of the United States. While the association between out-migration and life expectancy supports the theory for lower quartiles of incomes and urban and college educated population proportions and higher quartiles of Hispanic population proportions for the model without lags, this association vanishes for lagged models. Similarly, the association between in-migration and life expectancy follows the theory for higher quartiles of incomes, urban and college educated population proportions and lower quartiles of Hispanic populations. We find that some of these associations sustain over specifications with different lags of the independent variables.

How can we explain our findings? One important consideration is omitted variables that could be correlated with both life expectancy and migration that are not captured in our models. Another possibility could be that migration decisions are made in time periods beyond 0 to 3 years. We will explore and discuss these further in later analyses.

Table 1: Summary Statistics

Variable	Mean	SD	N
Out-Migration (%)	0.03	0.01	63861
In-Migration (%)	0.02	0.01	63861
Life Expectancy (Years)	76.18	2.09	54420
Asian (%)	0.01	0.03	63772
Black (%)	0.09	0.14	63772
American-Indian/Eskimo/Aleut (%)	0.02	0.08	63772
White (%)	0.84	0.17	63772
Hispanic (%)	0.07	0.12	63723
Owner-occupied Housing (%)	0.73	0.08	63772
Poverty Rate (%)	0.14	0.06	63772
Inflation-adjusted Median Family Income	62046.94	15188.38	63772
Per-capita Health Expenditure	99.16	156.21	60577
Per-capita Hospital Expenditure	340.28	705.24	60577
Per-capita Revenue Generated (Own Sources)	2521.87	2001.6	60577
Unemployment Rate (%)	0.07	0.03	63795
Urban (%)	0.41	0.32	63765
Completed College (%)	0.17	0.08	63772

Table 2: Predictors of In-migration and Out-migration

	(1) **Dependent Variable: In-migration**	(2) **Dependent Variable: Out-migration**
Life Expectancy (Years)	0.00000549	0.0000624*
Asian (%)	-0.00259	-0.0126***
Black (%)	-0.0103***	-0.00409**
American-Indian/Eskimo/Aleut (%)	-0.00373	-0.00285
Hispanic (%)	-0.0140***	-0.0180***
Owner-occupied Housing (%)	-0.0122***	-0.0194***
Poverty Rate (%)	-0.00146	0.00375***
Inflation-adjusted Median Family Income	6.79e-08***	9.62e-08***
Per-capita Health Expenditure	0.000000367*	0.000000188
Per-capita Hospital Expenditure	-0.000000143**	-4.87e-08
Per-capita Revenue Generated (Own Sources)	0.000000211***	6.82e-08***
Unemployment Rate (%)	-0.0338***	0.00683***
Urban (%)	-0.00130***	-0.00177***
Completed College (%)	0.00424**	0.00736***
Constant	0.0315***	0.0262***
Observations	51489	51638

p-values in parentheses

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

Note: We don't show coefficients for year dummies.

Table 3: Summary of Results (All models)

3a. Migration regressed on Life Expectancy at Lag 0

	IN		**OUT**	
	Violates Theory	Obeys Theory (Sign Positive)	Violates Theory	Obeys Theory (Sign Negative)
All Region*	NE,MW		MW	W
Urban	Q1	Q4	Q4	Q1
Hispanic	Q1,Q3		Q2	Q3,Q4
Income	Q1	Q3	Q3	Q2
Education	Q1	Q4	Q2,Q4	Q1

* East (E), Northeast (NE), Midwest (MW), West (W), South (S)

3b. Migration regressed on Life Expectancy at Lag 1

	IN		OUT	
	Violates Theory	Obeys Theory (Sign Positive)	Violates Theory	Obeys Theory (Sign Negative)
All			YES	
Region	NE,MW		MW,S	W
Urban	Q1	Q4	Q2,Q4	Q1
Hispanic			Q1,Q2	
Income			Q3	
Education	Q1	Q4	Q2,Q4	Q1

3c. Migration regressed on Life Expectancy at Lag 2

	IN		OUT	
	Violates Theory	Obeys Theory (Sign Positive)	Violates Theory	Obeys Theory (Sign Negative)
All			YES	
Region	NE,MW	S	MW,S	
Urban	Q1	Q4	Q2,Q3,Q4	
Hispanic			Q1,Q2,Q3	
Income		Q1	Q1,Q3,Q4	
Education		Q2,Q4	Q2,Q3,Q4	

3d. Migration regressed on Life Expectancy at Lag 3

	IN		OUT	
	Violates Theory	Obeys Theory (Sign Positive)	Violates Theory	Obeys Theory (Sign Negative)
All		YES	YES	
Region	NE	S	MW,S	
Urban	Q1	Q2,Q3,Q4	Q2,Q3,Q4	
Hispanic		Q1	Q1,Q2,Q3,Q4	
Income		Q1,Q3	Q1,Q3,Q4	
Education		Q2,Q4	Q2,Q3,Q4	

Bibliography

Arora, A., Spatz, E., Herrin, J., Riley, C., Roy, B., Kell, K., & Krumholz, H. M. (2016). Population well-being measures help explain geographic disparities in life expectancy at the county level. *Health Affairs*, *35*(11), 2075-2082.

Desmet, K., & Fafchamps, M. (2005). Changes in the spatial concentration of employment across US counties: a sectoral analysis 1972–2000. *Journal of economic geography*, *5*(3), 261-284.

Gilliam, J. F., & Fraser, D. F. (1987). Habitat selection under predation hazard: test of a model with foraging minnows. *Ecology*, *68*(6), 1856-1862.

Hutto, R. L. (1985). Habitat selection by nonbreeding, migratory land. *Habitat selection in birds*, 455.

Institute for Health Metrics and Evaluation, 2015. US Data for Download Life expectancy Estimates. Retrieved from: http://www.healthdata.org/us-health/data-download (Accessed March 2016).

Internal Revenue Services. (2018). *SOI Tax Stats: U.S. Population County-to-County Migration Data 1990-2016*. Retrieved from:https://www.irs.gov/statistics/soi-tax-stats-migration-data(Accessed May 2018).

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic perspectives*, *25*(3), 173-96.

Wang, H., Schumacher, A. E., Levitz, C. E., Mokdad, A. H., & Murray, C. J. (2013). Left behind: widening disparities for males and females in US county life expectancy, 1985–2010. *Population health metrics*, *11*(1), 8.