Migratory Responses to Environmental Variability in the United States

A MULTILEVEL APPROACH

Shuai Zhou and Guangqing Chi, Pennsylvania State University
2023 Annual Meeting of the Population Association of America
New Orleans, LA, April 12 – 15, 2023

Background and knowledge gap

- Globally, previous studies primarily focused on environmental changes in the developing world

- In the US, studies on slow-onset environmental variabilities used aggregated data at the regional level or crude level

  Gutmann et al. (2005): Great Plains region, 1930-1990
  Poston et al. (2009): The entire US at the state level, 1995-2000
  Feng et al. (2012): Corn belt region, 1970-2009

- There is a knowledge gap regarding the impact of slow-onset environmental variabilities on migration at the individual level in developed setting

Research objectives

- Explore individuals' migratory responses to slow-onset environmental variabilities (precipitation, temperature, air quality, and environmental amenity)

- Examine the heterogeneous environmental impacts on migration across two demographic groups (age group 15-64 and age group 65+)

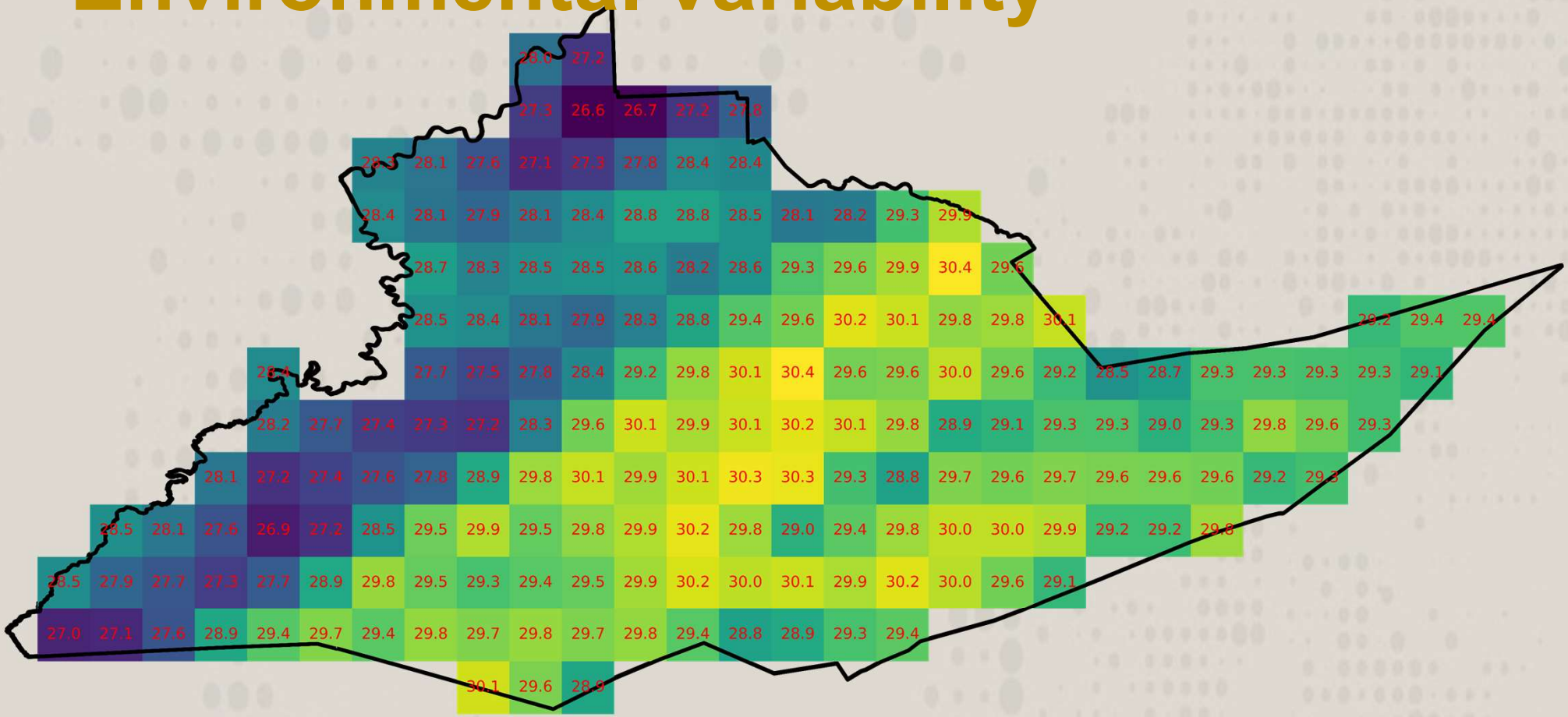
Data

- ✓ The American Community Survey (ACS) Microdata
- ✓ The Parameter-elevation Regressions on Independent Slopes Model (PRISM)
- ✓ The Atmospheric Composition Analysis Group (ACAG)
- ✓ The National Oceanic and Atmospheric Administration (NOAA)

Migration definition

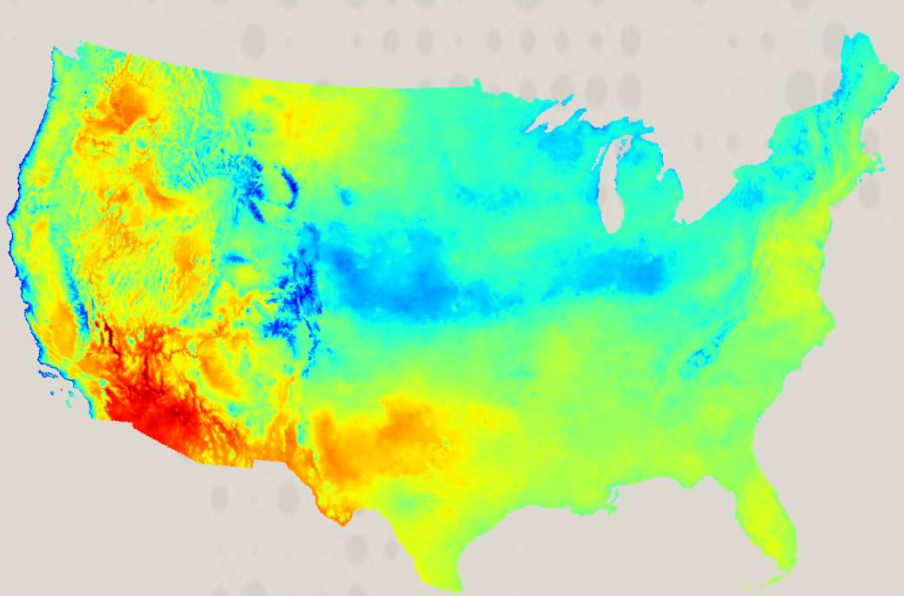
Migrations are moves cross counties/Public Use Microdata Areas (PUMAs) between the ACS years

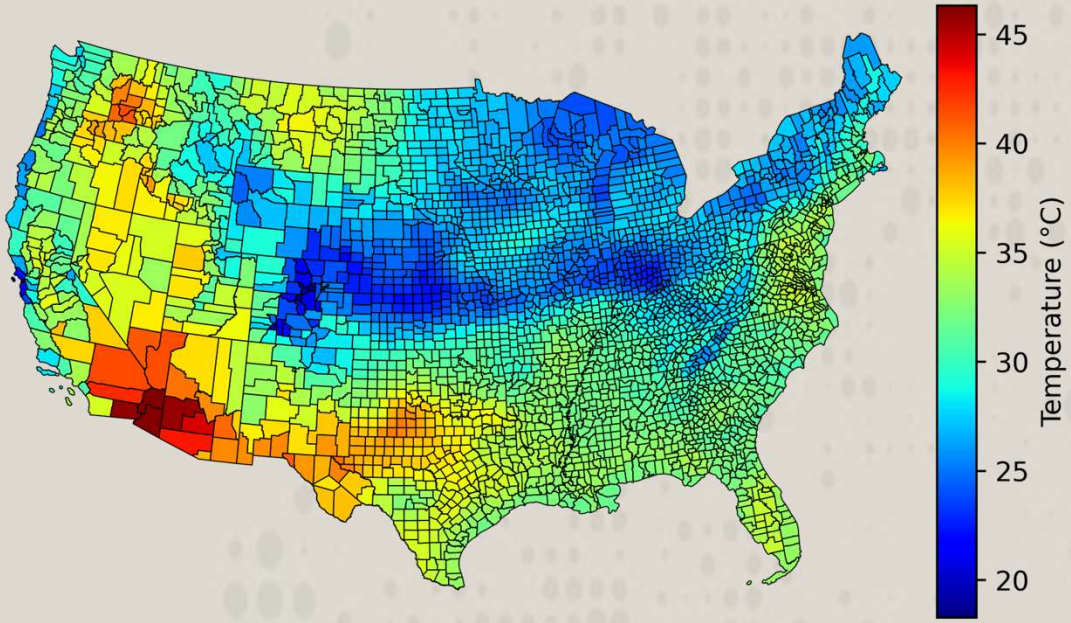
Environmental variability



County-level temperature = 28.97°C

Environmental variability (cont.)





Environmental variability (cont.)

$$\boldsymbol{Climate\ anomaly_{i,t}} = \frac{\boldsymbol{Level_{i,t} - \mu_i^{LR}}}{\boldsymbol{\sigma_i^{LR}}}$$

$Level_{i,t}$ = Annual average in county *i* at time *t* (2010-2020)

μ_i^{LR} = Long-run (30-year, 1980-2009) average in county *i*

σ_i^{LR} = Long-run (30-year, 1980-2009) standard deviation in county *i*

Two-level logistic regression

$$Logit(\Pr(Y_{ij} = 1)) = \alpha_0 + \alpha_{0j} + \alpha_1 X_{1ij} + \cdots + \alpha_k X_{kij} + \beta_1 Z_{1j} + \cdots + \beta_m Z_{mj}$$

Level-1 (individual) variables:
Age
Personal income
Gender
Marital status
Race
Education

Level-2 (county) variables:
Climate anomalies
Household income
Housing price
Employment rate
Homeownership
Metro status

Note: Climate anomalies include anomalies in precipitation, temperature, PM2.5, and Normalized Difference Vegetation Index (NDVI).

Descriptive statistics

	Mean	SD	Min	Max
Dependent variable				
Migration status	**0.42**	0.49	0	1
Level-1 variables				
Age	37.52	17.78	**15**	96
Personal income ($1,000)	32.45	52.97	-14.10	1,378.00
Gender	0.50	0.50	0	1
Marital status	0.33	0.47	0	1
Race	0.59	0.49	0	1
Education	0.53	0.50	0	1

Note: N = 2,243,336.

Race distribution: Non-Hispanic White (59%), Non-Hispanic Black (14%), Hispanics (17%), Others (10%).

Descriptive statistics (cont.)

	Mean	SD	Min	Max
Level-2 variables				
Precipitation anomaly	**0.09**	0.34	-0.84	1.55
Temperature anomaly	**0.06**	0.11	-0.40	0.48
NDVI anomaly	**-0.01**	0.17	-1.12	0.53
PM2.5 anomaly	**-1.07**	0.46	-2.12	1.13
Household income ($1,000)	89.47	20.94	48.97	178.22
Housing price ($1,000)	296.00	181.35	81.88	1,111.50
Employment rate	91.95	2.48	81.49	97.58
Homeownership	61.25	10.88	18.97	87.44
Metropolitan status	**0.99**	0.09	0	1

Note: N = 2,243,336.

General models

Level-1 variables	
Age	-0.010***
Personal income	-0.001***
Gender, Male (Ref. = Female)	0.097***
Marital status, Married (Ref. = Unmarried)	-0.071***
Race, NHB (Ref. = NHW)	-0.209***
Race, Hispanics (Ref. = NHW)	-0.335***
Race, Others (Ref. = NHW)	0.022***
Education, College and above (Ref. = Below college)	0.170***
Level-2 variables	
Precipitation anomaly	0.017**
Temperature anomaly	0.075***
NDVI anomaly	-0.249***
PM2.5 anomaly	-0.006***

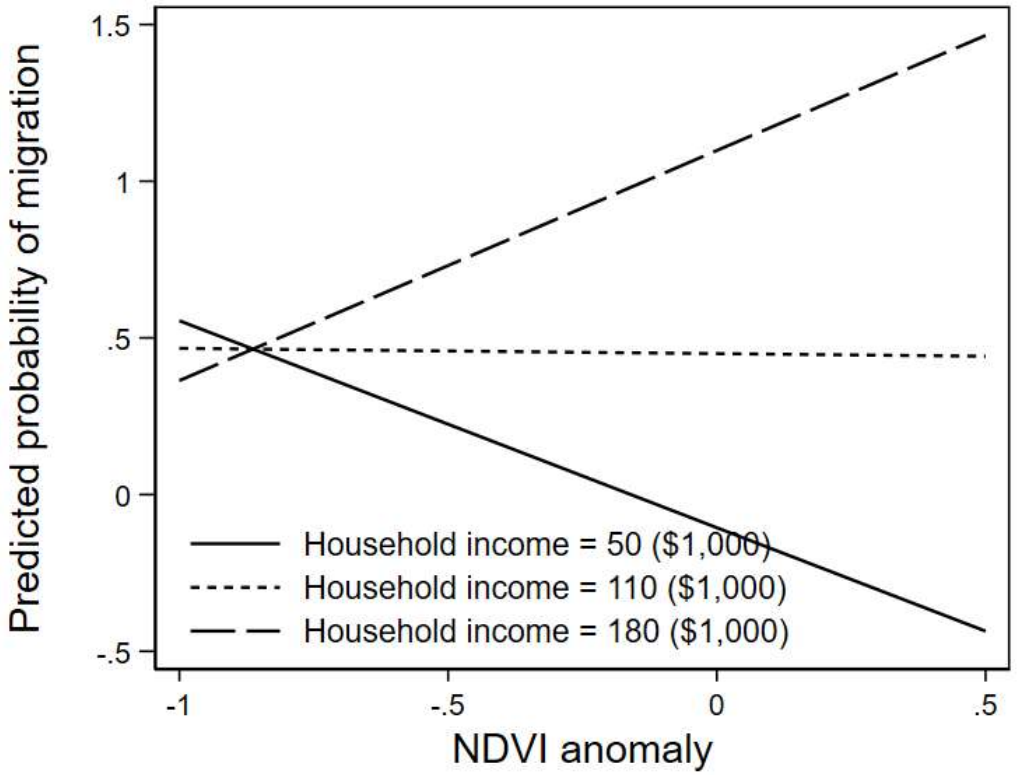
Note: *** p<0.001, ** p<0.01, * p<0.05. NHB=Non-Hispanic Black, NHW=Non-Hispanic White. Level-2 sociodemographic factors and model diagnostics are not show.

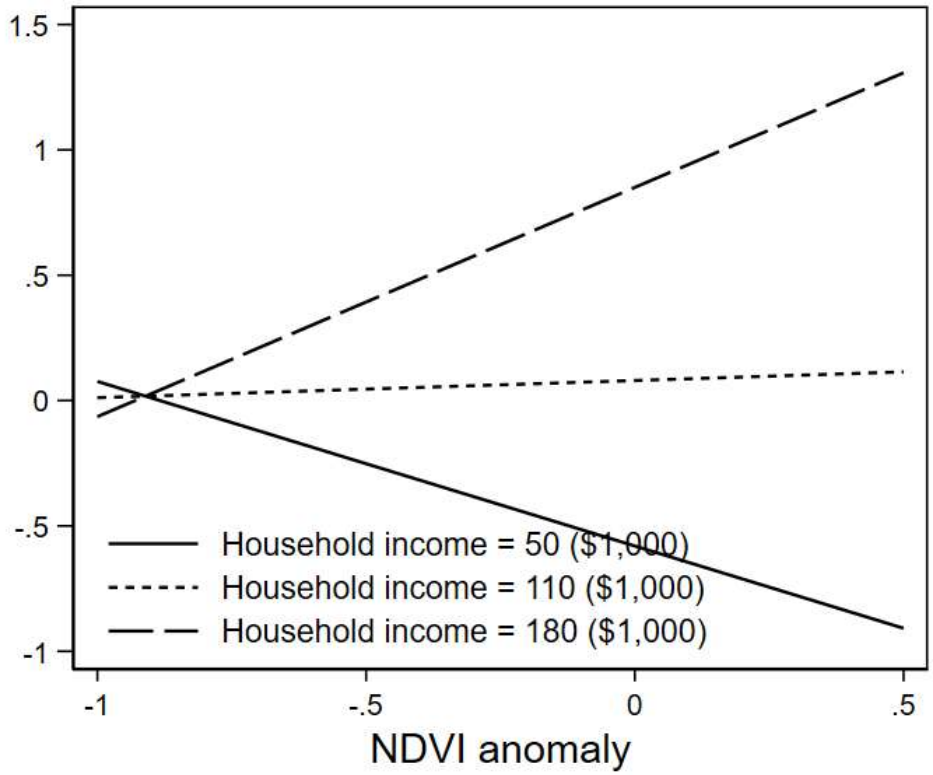
Age-specific models

	Mig_{15-64}	Mig_{65+}
Level-2 variables		
Precipitation anomaly	0.016**	0.033
Temperature anomaly	0.093***	-0.127*
NDVI anomaly	-0.247***	-0.193***
PM2.5 anomaly	-0.092***	-0.001

Note: *** p<0.001, ** p<0.01, * p<0.05. Level-1 variables, Level-2 sociodemographic factors and model diagnostics are not show.

Climate-Income interaction

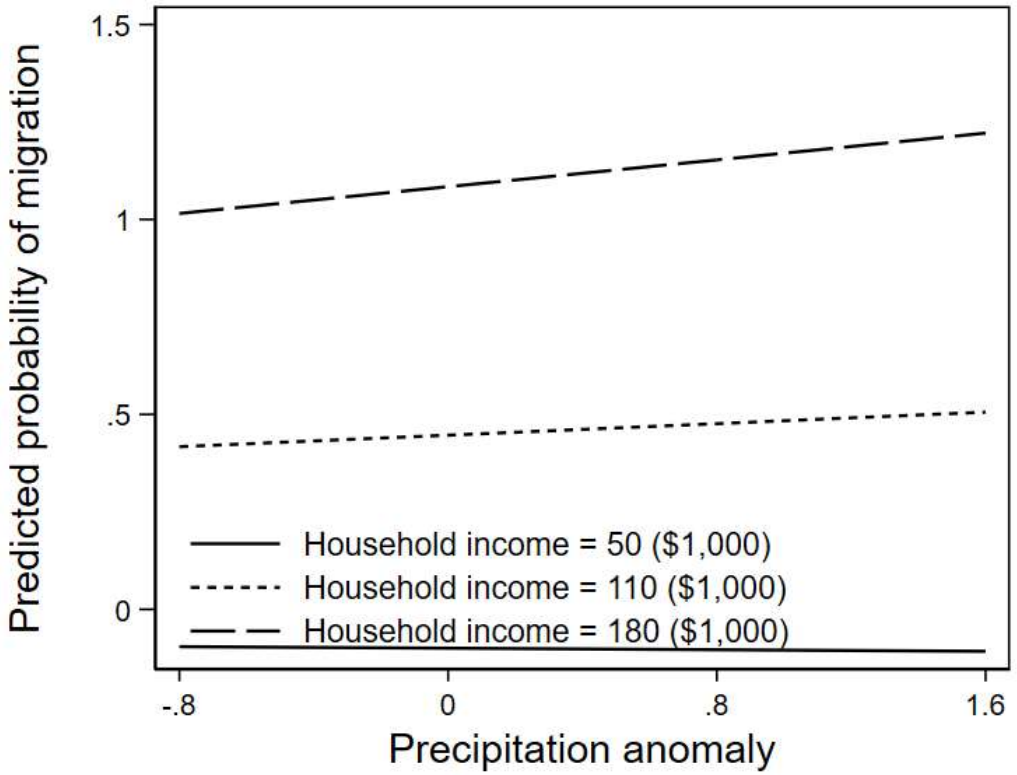




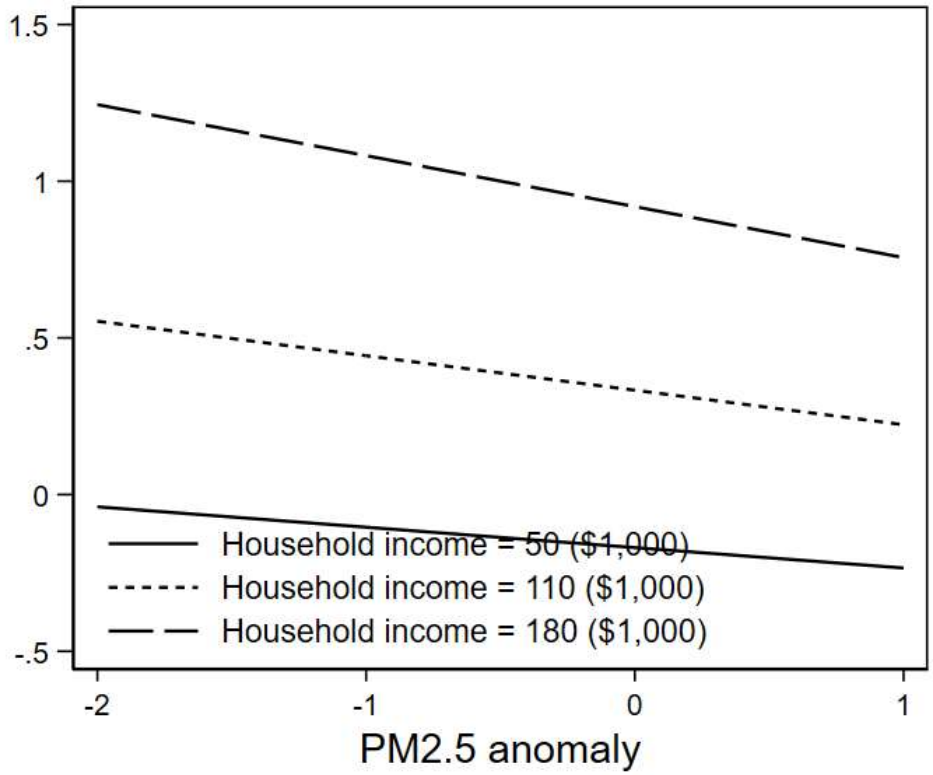
Migration$_{15\text{-}64}$

Migration$_{65+}$

Climate-Income interaction (cont.)



Migration$_{15\text{-}64}$



Migration$_{15\text{-}64}$

Findings

- Being male, non-Hispanic white, and highly educated increased migration probability

- Precipitation and temperature anomalies generally increased migration probability, while PM2.5 and NDVI anomalies decreased migration probability

- The elder generation was responsive to temperature and environmental amenity, while the younger generation preferred places with environmental amenity, economic well-being, and affordable living costs

Limitations

1. The ACS microdata from 2010 to 2020 were treated as cross-sectional rather than longitudinal data

2. Only 523 counties (~ 1/6 US counties) were identified through matching county and PUMA, among which 99% are metro counties

EDeN Environmental Demography Network

Thank you

Shuai Zhou, Ph.D. candidate
Rural Sociology and Demography (Dual Title)
The Pennsylvania State University
801 Oswald Tower, University Park, PA 16802
sxz217@psu.edu

Appendixes

Table 1: Data sources, variables, and coding scheme

Variable	Coding scheme	Source
Dependent variable		
Migration status	Migrant = 1; Stayer = 0	ACS
Level-1 variables		
Age	Continuous variable	ACS
Personal income	Continuous variable	ACS
Gender	Female = 1; Male = 0	ACS
Marital status	Married = 1; Unmarried = 0	ACS
Race	Non-Hispanic White = 1; Non-Hispanic Black = 2; Hispanics = 3; Others = 4	ACS
Education	College and above = 1; Below college = 0	ACS
Level-2 variables		
Precipitation anomaly	Continuous variable	PRISM
Temperature anomaly	Continuous variable	PRISM
NDVI anomaly	Continuous variable	NCAR
PM2.5 anomaly	Continuous variable	ACAG
Household income	Continuous variable	ACS
Housing price	Continuous variable	ACS
Employment rate	Continuous variable	ACS
Homeownership	Continuous variable	ACS
Metropolitan status	Metro = 1; Nonmetro = 0	OMB

Table 3: Two-level logistic regression predicting migration status in the U.S., 2010-2020

	Model 1	Model 2	Model 3
Level-1 variables			
Age		-0.010***	-0.010***
Personal income		-0.000***	-0.001***
Gender, Male (Ref. = Female)		0.093***	0.097***
Marital status, Married (Ref. = Unmarried)		-0.078***	-0.071***
Race, NHB (Ref. = NHW)		-0.222***	-0.209***
Race, Hispanics (Ref. = NHW)		-0.326***	-0.335***
Race, Others (Ref. = NHW)		0.050***	0.022***
Education, College and above (Ref. = Below college)		0.172***	0.170***
Level-2 variables			
Precipitation anomaly			0.017**
Temperature anomaly			0.075***
NDVI anomaly			-0.249***
PM2.5 anomaly			-0.006***
Household income			0.010***
Housing price			0.001***
Employment rate			-0.047***
Homeownership			0.004***
Metro county (Ref. = Nonmetro county)			0.132***
Constant	0.170	0.513***	3.115***
Year effect	—	—	Controlled
County effect	—	—	Controlled
Observations	2,243,336	2,243,336	2,243,336
ICC	0.675	0.676	0.673
LR test	—	23,823***	20,932***

Table 4: Two-level logistic regression predicting age-specific migration status in the U.S., 2010-2020

	Mig_{15-64}	Mig_{65+}
Level-1 variables		
Age	-0.014***	-0.015***
Personal income	-0.001***	0.001***
Gender, Male (Ref. = Female)	0.104***	-0.007
Marital status, Married (Ref. = Unmarried)	-0.068***	0.083***
Race, NHB (Ref. = NHW)	-0.192***	-0.237***
Race, Hispanics (Ref. = NHW)	-0.331***	-0.227***
Race, Others (Ref. = NHW)	0.029***	-0.027
Education, College and above (Ref. = Below college)	0.170***	0.231***
Level-2 variables		
Precipitation anomaly	0.016**	0.033
Temperature anomaly	0.093***	-0.127*
NDVI anomaly	-0.247***	-0.193***
PM2.5 anomaly	-0.092***	-0.001
Household income	0.009***	0.011***
Housing price	0.001***	0.001***
Employment rate	-0.048***	-0.042***
Homeownership	0.006***	-0.020***
Metro county (Ref. = Nonmetro county)	0.160***	-0.128
Constant	3.123***	4.823***
Year effect	Controlled	Controlled
County effect	Controlled	Controlled
Observations	2,029,092	214,244

Table 5: Two-level logistic regression predicting migration status in the U.S. with interactions between environmental factors and household income, 2010-2020

	Mig_{All}	$Mig_{15\text{-}64}$	Mig_{65+}
Level-1 variables			
Age	-0.010***	-0.014***	-0.015***
Personal income	-0.001***	-0.001***	0.001***
Gender, Male (Ref. = Female)	0.097***	0.104***	-0.007
Marital status, Married (Ref. = Unmarried)	-0.071***	-0.068***	0.083***
Race, NHB (Ref. = NHW)	-0.209***	-0.192***	-0.237***
Race, Hispanics (Ref. = NHW)	-0.337***	-0.332***	-0.228***
Race, Others (Ref. = NHW)	0.021***	0.028***	-0.028
Education, College and above (Ref. = Below college)	0.170***	0.170***	0.230***
Level-2 variables			
Precipitation anomaly	-0.025	-0.04	0.152†
Temperature anomaly	0.125†	0.166*	-0.348
NDVI anomaly	-1.238***	-1.197***	-1.261***
PM2.5 anomaly	-0.015	-0.028	0.092
Household income	0.009***	0.008***	0.010***
Housing price	0.001***	0.001***	0.001***
Employment rate	-0.048***	-0.048***	-0.045***
Homeownership	0.004***	0.006***	-0.020***
Metro county (Ref. = Nonmetro county)	0.118***	0.146***	-0.137
Interaction terms			
Precipitation anomaly * Household income	0.001*	0.001**	-0.001
Temperature anomaly * Household income	-0.000	-0.001	0.002
NDVI anomaly * Household income	0.011***	0.011***	0.012***
PM2.5 anomaly * Household income	-0.001***	-0.001***	-0.001
Constant	3.288	3.275***	5.161***
Year effect	Controlled	Controlled	Controlled
County effect	Controlled	Controlled	Controlled
Observations	2,243,336	2,029,092	214,244
Numbers of county	523	523	523
ICC	0.670	0.668	0.675