

New Estimates of Geographic Mobility and the Returns to Migration from Individual Level Tax Records, Linked Administrative and Census Data (1969-1989)

John Sullivan

Geographic mobility in the United States has been declining since the late 1960s. In this paper, we focus on the years around the decline and create a longitudinal panel built on the universe of IRS Form 1040 tax returns from 1969, 1974, 1979, 1984 and 1989 and supplemented with information from additional administrative records and the 2000 decennial census. We present estimates of 5-year mobility rates for both short and long-distance moves, compare our estimates to those from other sources and disaggregate mobility rates by age, race, and income. We also estimate short and long-term economic returns to domestic migration using earnings from the tax data and home values from the 2000 census. We compare these outcomes for inter-county and inter-state moves, as well as moves from county and state of birth. We consider selection issues in our data sources and in our analysis of the effects of migration on economic outcomes. Specifically, we recognize the omission of the non-tax filing population and residents without a social security number from our sample and consider the conditions that result in some people choosing to migrate and others choosing to remain.

Introduction

Internal migration in the United States can be an important avenue for social and economic opportunity. The ability to move from one place to another allows individuals to seek social and economic opportunities and can be an important driver of racial integration, for example, African Americans moving to the North during the Great Migration (Alexander et al. 2017). Further, a workforce that is mobile can move to high productivity areas, taking advantage of better job opportunities and wages and increasing regional and even national productivity (Hsieh and Moretti 2019).

Evidence from surveys like the Current Population Survey (CPS) show a slow but steady decline in internal migration in the United States since the late 1960s. The decline can be attributed in part to compositional changes in the US population (Foster 2017, Boznick 2021), increases in housing cost (Jia et al. 2023), and, potentially, changes in the returns to migration (Purcell 2020). Yet, questions remain about which groups experienced the decline first and most severely and for which groups the returns to migration changed.

Measuring migration at a population level is challenging because migration data is often available only in surveys with relatively small samples or in aggregate estimates with few contextual measures (like public county to county migration flows produced from tax returns). Further, measurement of later life economic outcomes (accuracy, life-cycle, and period related variability in income (Torche 2013)) and the endogeneity of migration decisions present major challenges to the measurement of the returns to migration.

In the 1970s, the Internal Revenue Service (IRS) began providing individual level income tax returns to the Census Bureau to support income measurement and the production of population projections. We utilize this information on the universe tax fillers from the years 1969, 1974, 1979, 1984 and 1989 to present novel estimates of migration rates at multiple geographic levels. We explore declining migration by tracing changing migration rates for age, race and income groups and consider changing returns to migration in terms of two outcomes, income and housing value. In doing so, we can better understand why migration has declined in the US, for whom it declined first and at what geographic level the decline has been most severe.

Data and Methods

Our estimates primarily rely on location information from individual 1040 tax returns for the universe of tax fillers from the years 1969, 1974, 1979, 1984 and 1989. These data have been processed by the Census Bureau's Person Validation System (PVS) which assigns unique identifiers, called Protected Identification Keys (PIKs), to individual level records. Using PIKs, we link tax fillers not only to their filing locations in subsequent years, but also to a rich body of measures from the 2000 Census and Social Security Administration (SSA) administrative records. This person level linkage allows us to longitudinally measure geographic and income mobility and to contextualize those measurements with characteristics otherwise unavailable from tax records, notably, race/ethnicity, household structure, mortality, place of birth and home value.

From the tax data, we use address information to estimate 5-year migration rates at multiple geographic levels (state, county and tract), allowing us to see at which geographic levels mobility decline occurred. For example, did interstate mobility decline before or to a greater extent than movement between

census tracts? We make these estimates for the full universe of tax filers, but also for age, race and income groups, observing for which groups migration rates fell first.

Additionally, we will present estimates of the economic returns to various types of domestic migration. To address the selectivity of migrants, we use gradient boosted models to estimate propensity scores. These scores allow us to pair the population of migrants with a comparable group of non-migrants. We then compare those that moved to those that did not on two measures, income and home value (from the 2000 Census long form sample). We estimate the short-term effect of migration on income for each 5-year period, but also estimate the effect on longer-term earnings.

Conclusion

In this paper, we will present new estimates of migration during a period of declining internal migration in the United States. Using restricted-use IRS data, combined with additional administrative and survey records, we examine migration during this period and estimate returns to migration. This work will provide insights into current debates on the importance of variation in the returns to migration in migration decisions and the selectivity of samples built upon administrative records like tax and social security data.

Citations

Alexander, J. T., Leibbrand, C., Massey, C., & Tolnay, S. (2017). Second-generation outcomes of the great migration. *Demography*, 54(6), 2249-2271.

Bleckley, D. A., Genadek, K. R., Alexander, J. T. (2023). Assigning Contemporaneous Census Tracts to Historical Income Tax Data. Census Bureau ADEP Working Paper 2023-03.
https://www.census.gov/content/dam/Census/library/working-papers/2023/econ/censustract_taxdata.pdf

Bozick, R. (2021). Age, period, and cohort effects contributing to the Great American Migration Slowdown. *Demographic Research*, 45, 1269-1296.

Foster, T. B. (2017). Decomposing American immobility: Compositional and rate components of interstate, intrastate, and intracounty migration and mobility decline. *Demographic Research*, 37, 1515-1548.

Hsieh, C. T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1-39.

Jia, N., Molloy, R., Smith, C., & Wozniak, A. (2023). The economics of internal migration: Advances and policy questions. *Journal of economic literature*, 61(1), 144-180.

Purcell, P. J. (2020). Geographic mobility and annual earnings in the United States. *Soc. Sec. Bull.*, 80, 1.

Torche, Florencia. (2013). How do we characteristically measure and analyze intergenerational income? Stanford Center on Poverty and Inequality Working Paper Series.