Master’s Capstone Topic Proposal

Tyler Frankenberg

DATA 698

Dr. Arthur O’Connor

February 26, 2023

City University of New York

This paper will explore the influence of the economic factors that drove population growth or decline through domestic in/out migration at the United States county level over the period 2011 – 2019. Using simple and multiple linear regression models, the analysis will considers including but not necessarily limited to personal tax burden, labor market health, and housing and rental market affordability and their effect on county percent net population growth from domestic migration.

Domestic migration within the United States is of heightened public interest in the present political and economic climate. The 2020 Census revealed a substantial acceleration of migration from established population centers in the Northeast and on California’s West Coast to the so-called Sun Belt states in the South and Southwest. This migration was politically significant due to the shift in congressional representation it heralded, from Democratic Party strongholds in California and New York to more traditionally Republican controlled Sun Belt states, and has been hailed by conservative political commentators as vindication of these Red states’ light tax burden on citizens when compared with the progressive tax regimes of population losers. On the other hand, the same demographic shifts may portend a leftward shift in the politics of some of these states and localities.

In tandem with this trend, the 2020 COVID-19 pandemic’s lockdowns and transitions to remote work spurred a disruption in patterns of residential settlement away from urban centers, with vast implications for local economies, for the health and affordability of local residential and commercial real estate markets, as well as for the size and composure of state and municipal tax bases, and the public services they support. With such significant economic forces at work, and with so much political and economic power at stake, States, Counties, and Municipalities want to know how they can attract and retain population. Yet, it can be difficult to discern actual causes and effects of population change out of the highly politicized discourse. In order to accurately describe recent trends in domestic migration, determine their causes, and evaluate their effects, we must take a holistic view using multiple data sources, each placed in the appropriate context and considered for their relationship to the whole.

This paper will frame such a holistic ‘baseline’ view of the economic facts by conducting multiple linear regression analysis testing the influential variables derived from IRS, Census, and other federal agency datasets to explain causes of population change from domestic migration at the county level. Such analyses are commonly conducted across academic disciplines from Economics and American Politics, to Urban Planning, Sociology, and Public Policy.

This paper explores the influence of the economic factors that drove population growth or decline through domestic in/out migration at the United States County level over the years 2011-2019. Using simple and multiple linear regression models, the analysis considers factors including personal tax burden, labor market health, and housing and rental market affordability and their effect on county percent net population growth from domestic migration. I begin with the null hypotheses that all three variables are statistically insignificant as predictors of percent net population change from domestic migration, and test their respective individual influence through simple linear regression. I then explore the influence of the variables together using multiple linear regression, giving attention to potential interdependence and collinearity between them and attempting transformations as necessary to address each.

This paper will frame such a holistic ‘baseline’ view of the economic facts by conducting multiple linear regression analysis testing the influential variables derived from IRS, Census, and other federal agency datasets to explain causes of population change from domestic migration at the county level. Such analyses are commonly conducted across academic disciplines from Economics and American Politics, to Urban Planning, Sociology, and Public Policy. In dialogue with the work of urbanist Richard Florida in *The Rise of the Creative Class* (2002), McGranahan and Wojan (2007) considered different definitions of this sought-after titular class of educated economic migrants and their implications for the study of migration dynamics across Urban and Rural counties.

Their work moreover, incorporates McGranahan’s Natural Amenity Scale (1999) to characterize the geographic appeal of counties in the lower 48 United States, which this project expects likewise to examine as a control variable. Partridge et al (2012) analyze an apparent diminishing influence of labor market changes on county-to-county migration in the early 2000s when compared with the historic trend, while Jeanty, Partridge, and Iriwin (2010) develop a spatial simultaneous equation for the effect of housing market factors. Sociologists have examined the so-called ‘Tax Flight’ phenomenon in which high-income individuals migrate in search of a lower tax liability (Young, C., Varner, C., Lurie, I. Z., & Prisinzano, R, 2016).

The analysis will use as its dependent variable the county-by-county percent net population change from domestic migration, calculated from the Census Bureau’s annual population estimates (United States Census Bureau 2009, 2019, 2022). The analysis will test the following independent variables for a statistically significant relationship:

* State and Local Personal Tax Burden calculated as the proportion of the county’s per capita Adjusted Gross Income collected in the form of State or Local income taxes, personal property taxes, real estate taxes, or general sales taxes, minus refunds. This will be calculated from the Internal Revenue Service’s annual SOI Tax Stats County datasets (<https://www.irs.gov/statistics/soi-tax-stats-county-data>).
* Labor Market Health measured by the county-by-county unemployment rate, with a low unemployment rate signaling a healthy labor market and a high unemployment rate signaling a distressed labor market. The data source will be the US Department of Agriculture Economic Research Service’s summary of Department of Labor unemployment data ([https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/](https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/).)).
* Housing and rental market affordability, as measured by the Federal Housing Finance Administration annual residential real estate price index (<https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx#qat>)   
  and the Fair Market Rent price as established by the Department of Housing and Urban Development (<https://www.huduser.gov/portal/datasets/fmr.html#null>).

This paper will frame such a holistic ‘baseline’ view of the economic facts by conducting multiple linear regression analysis testing the influential variables derived from IRS, Census, and other federal agency datasets to explain causes of population change from domestic migration at the county level. Such analyses are commonly conducted across academic disciplines from Economics and American Politics, to Urban Planning, Sociology, and Public Policy. We will begin with the null hypotheses that all three variables are statistically insignificant as predictors of percent net population change from domestic migration, and test their respective individual influence through simple linear regression. The paper will then explore the influence of the variables together using multiple linear regression, giving attention to potential interdependence and collinearity between them and attempting transformations as necessary to address each. The paper will also attempt as is feasible to address the influence of confounding variables by controlling for, at minimum:

* The relative climate, weather, and geographic appeal of counties as measured by their Natural Amenity Score (McGranahan, 1999).
* The effect of progressive tax regimes, by considering the differential influence of tax burden at different levels of Adjusted Gross Income (Internal Revenue Service)

# References

Federal Housing Finance Administration (2022). Annual House Price Indexes. Accessed February 26, 2023 from <https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx#qat>.

Internal Revenue Service. SOI Tax Stats – County Data. Updated December 19, 2022. Accessed February 26, 2023 from <https://www.irs.gov/statistics/soi-tax-stats-county-data>.

United States Census Bureau, (2022). United States Population Estimates, Vintage 2022. Accessed from <https://www2.census.gov/programs-surveys/popest/datasets/>.

United States Census Bureau, (2019). United States Population Estimates, Vintage 2019. Accessed from <https://www2.census.gov/programs-surveys/popest/datasets/>.

United States Census Bureau, (2009). United States Population Estimates, Vintage 2009. Accessed from <https://www2.census.gov/programs-surveys/popest/datasets/>.

United States Department of Agriculture (2022). Unemployment and median household income for the U.S., States, and counties, 2000-2021. Economic Research Service. Last Updated June 3, 2022. Accessed February 26, 2023 from <https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/>

United States Department of Housing and Urban Development. Fair Market Rents (40th Percentile Rents). Office of Policy Development and Research. Accessed February 26, 2023 from <https://www.huduser.gov/portal/datasets/fmr.html>.