

# BUILDING A USE CASE ON DOMESTIC MIGRATION AND THE CURATED DATA ENTERPRISE

*Essential Elements*

---

*April 21, 2022*

## **Joseph Salvo**

Institute Fellow,  
Social and Decision Analytics Division  
<http://orcid.org/0000-0002-1552-3387>  
[ckp7gd@virginia.edu](mailto:ckp7gd@virginia.edu)

## **Stephanie Shipp**

Deputy Director and Professor,  
Social and Decision Analytics Division  
<https://orcid.org/0000-0002-2142-2136>  
[sss5sc@virginia.edu](mailto:sss5sc@virginia.edu)

## **Stephanie Zhang**

Intern,  
Social and Decision Analytics Division  
<https://orcid.org/0000-0001-9804-6522>  
[sz3wr@virginia.edu](mailto:sz3wr@virginia.edu)

**Acknowledgments:** This report was funded by U.S. Census Bureau Contract  
Census Agreement No. 01-21-MOU-06.

**Citation:** Salvo, J., Shipp, S., Zhang, S. (2022). Building a Use Case on  
Domestic Migration and the Curated Data Enterprise - Essential Elements.  
Proceedings of the Biocomplexity Institute. Technical Report 2022-027  
<https://doi.org/10.18130/bcwa-gt69>



**BIOCOMPLEXITY  
INSTITUTE**

## **Abstract**

The Curated Data Enterprise (CDE) is both an infrastructure and a continuously evolving ambition to empower and enable Census Bureau scientists and their data users to develop together new measures of the nation's people, places, and economy. The CDE could be an engine for better migration and mobility data, leading to improved planning for labor market changes, emergency response, and state revenue projections. State and local economic planners and demographers need information about the interplay between migration and local labor markets, especially given recent dramatic changes in the nature of work during the pandemic. Yet current measures do not keep pace with these changes.

Any area of the country affected by wildfires, extreme weather, flooding, and droughts associated with longer-term watershed issues, experience changes in their patterns of migration. These events precipitate movements that need to be monitored close to real time, so local and state planners can react quickly. When disaster strikes, emergency preparedness plans are needed, informed by reliable data on the size of affected populations, the composition of households, their access to travel or transit, to assess the potential for population displacement. State and local governments require population data for revenue projections, the distribution of resources, the identification of needs, and the establishment of priorities and strategies to address those needs. To be most useful, planners need access to methods to measure migration close to real time to discern the difference between temporary versus permanent shifts in population.

The CDE could support the integration of relevant data sources, including federal surveys such as the Current Population Survey (CPS) and the American Community Survey (ACS), combined with administrative data such as tax and address data, as well as third-party data, e.g., new home mortgage applications, real estate inquiries/transactions, private company mover records, cellphone detailed records, and utility company accounts. Technical challenges to overcome in using these data involve creating linkages between survey data, which have a substantial lag but are still valuable to gauge trends to shorter-term indicators, all for small geographic areas. Locally, migration patterns need to be studied as a way of addressing equity issues, for example involving access to housing, health care, and broadband, and displacement of populations based on race, ethnicity, and income. These observations not only rely on estimates of the total population, but the integral role played by population estimates and projections by age and sex, which undergirds measures of well-being, especially rates related to public health issues and poverty measures.

## Building a Use Case of Domestic Migration and the Curated Data Enterprise - Essential Elements

### I. Overview and Scope: Why a Use Case?

In the most general sense, demographers define *migrants* as those who move between “communities of residence,” as a subset of all *movers*, which include those who move between and within those communities. In practice, given the available data, most demographers in the U.S. define domestic migration as movement across a county boundary, which is just one part of the much larger picture of population mobility or all moves that occur from year to year.<sup>1</sup> Migration is the most volatile component of population change and the most difficult to estimate, especially for small areas and groups. It is tied to age, sex, and race, can vary significantly with local social, economic, and housing conditions and is impacted by local and national policies. Even more challenging, there is no single, comprehensive source of migration data in the U.S.

Still, identifying migration patterns affecting localities is among the most critical aspects of local planning. While seasonal migration patterns are fairly predictable, the increasingly unpredictable nature of factors affecting the movement of people – weather events, wildfires, and pandemics – have made migration an important issue for disaster mitigation and preparation strategies. Moreover, with more people working remotely and becoming untethered to a physical office location, local officials need to monitor their populations’ movements. Measures of migration are needed to evaluate the impact on current population estimates and population projections looking forward, all of which creates a context for critical decisions affecting strategic plans for business development, housing, disaster preparedness, and a whole range of policies affecting the evaluation of equity in the distribution of resources.

### II. What We Heard at the Listening Sessions

Following some of the issues above, our Listening Sessions consistently revealed a need for data on domestic migration. These can be divided into three main areas of concern on the part of the data user community:

- 1. The interplay between migration and local labor markets, given changes in the nature of work, especially during the pandemic.** Local impacts of changes in the larger economy and the world of work are reflected in migration patterns to and from different parts of the nation. Yet current measures are incapable of keeping pace with these changes. Even monthly data related to home sales are largely “after the fact” and

---

<sup>1</sup> For an overview of definitions, see Edmonston (2012).

not actionable in real time. People's actions, characteristics, and preferences are important for local strategies to deal with the impact of larger economic changes.

The pandemic has made real-time migration data more important to monitor trends. For example, the pandemic has accelerated the migration of persons from California to other states – measuring the impact of this movement on local labor and housing markets is challenging without the aid of real-time information. There are the local effects brought on by the increasing propensity of people to work remotely – the physical separation from an office. Teleworking affects migration decisions, which affects demand for office space, local municipal budgets, and services. Also, the occupations and industries found in sending and receiving areas need to be explored to determine whether moves represent larger overall shifts in the nature of the economy. Regardless, migration needs to be measured in something close to real time to be useful. This is especially the case when local planners are trying to discern the difference between temporary vs. permanent moves.

**2. The impact of natural disasters, mainly due to extreme events related to climate change.** Areas of the country that have been affected by wildfires, extreme weather, flooding, and droughts associated with longer-term watershed issues, have all experienced changes in their migration patterns. These events precipitate movements that need to be monitored in something close to real-time so that local planners can react quickly. The fires in the Pacific Northwest, flooding in the south, and the longer-term watershed issues in the West are just cases in point as to why there need to be measures of migration that work in real-time. Perhaps some of this can be achieved by linking survey data with other records to provide a longer-term and shorter-term view of impacts. Reacting to the needs of a population when disaster strikes requires an emergency preparedness plan, which almost always is based on some estimate of the size of affected populations and the potential for population displacement. All of this is used as a frame for planned resource allocation in times of disaster.

**3. The effect of migration on population change – now and in the future.** State and local governments require population data for revenue projections, the distribution of resources, the identification of needs, and the establishment of priorities and strategies to address those needs. Given the volatile nature of migration, it is the most difficult component of population change to measure, especially in real-time. The flow of people to and from areas of the nation needs to be understood on a longer- and shorter-term basis, from moves related to job changes or the quest for more space, to shorter seasonal movement. This begs for linkages between survey data, which have a substantial lag but are still valuable to gauge trends to shorter-term indicators, all for small geographic areas. Locally, migration patterns need to be studied to address equity issues, for example, involving access to housing and displacement of populations based on race, ethnicity, and income. All these observations are based not only on estimates of the total population but also on the integral role of population estimates and



projections by age and sex, which undergirds measures of well-being, such as rates related to public health issues and poverty measures.

Common themes heard across these three areas of concern are the need for small area data and longitudinal data. For example, with the rise in remote work and the gig economy, location choices related to work likely have become more varied, increasing disparities between those who are tied to specific locations and those who are not. Longitudinal data can help to illustrate these patterns more clearly. Small area data on local migration patterns is in high demand, particularly for local decision-making. Oversampling small areas in current surveys is a partial solution. Using administrative records from other sources could also fill gaps in information. These include home sales data, digital platform data on homebuying queries, mortgage loan applications, information from moving companies, postal records, cell phone data, and utility information related to water, sewer, and electricity use. (Many of these data sources and others are discussed below.) Data could also be aggregated by geographic region, such as watershed areas, to provide more insight for climate and resource-related purposes.

### **III. The Measurement of Migration: Current measures**

#### **1. Surveys**

##### ***American Community Survey (ACS)***

The U.S. Census Bureau conducts the American Community Survey (ACS) as part of the decennial census program. The ACS officially began in 2005, following more than ten years of testing and development. It replaced the long form, which was last used in the 2000 Census.<sup>2</sup> The ACS uses a very large sample of the U.S. population consisting of 3.54 million addresses annually and provides representative estimates for geographic areas down to the block group level. Although sampling variability is high for very small estimates, the ACS sample's size and extensive content make it an important source of information on the population's demographic, social, and economic characteristics. This includes data on migration flows.

Based on the question shown below: "Where did this person live one year ago," the most recent ACS data show that about 6 percent of the U.S. population (1+ years of age) moved between counties each year over the 2015-2019 period. When the question is applied to a specific place, this method can be used to examine *gross flows*, or the stream of people moving in each direction relative to a specific geographic location. For example, for 2015-2019, the ACS showed that about 4,500 persons migrated from New York County (Manhattan) to Los Angeles County, California, each year -- on average -- for this period.<sup>3</sup> About 3,100 persons migrated

---

<sup>2</sup> For an overview and history of ACS development, see Anderson et. al., 2012:1-22.

<sup>3</sup> The data for the five years of sample are pooled and then controlled to an average of the five years of population estimates for 2015-2019. This is essentially the mid-year of that series unless there are large fluctuations. This "period estimate" is technically not an average; however, "average" is used here to provide a more intuitive picture of migration between these counties. For more information, see: [https://www.census.gov/newsroom/blogs/random-samplings/2022/03/period-estimates-american-community-survey.html?utm\\_medium=email&utm\\_source=govdelivery](https://www.census.gov/newsroom/blogs/random-samplings/2022/03/period-estimates-american-community-survey.html?utm_medium=email&utm_source=govdelivery)

from Los Angeles to New York County for each year in this period, yielding a net migration loss to Manhattan of 1,400 persons annually (which is statistically significant).

**15 a. Did this person live in this house or apartment 1 year ago?**

☐ Person is under 1 year old → *SKIP to question 16*

☐ Yes, this house → *SKIP to question 16*

☐ No, outside the United States and Puerto Rico – *Print name of foreign country, or U.S. Virgin Islands, Guam, etc., below; then SKIP to question 16*

☐ No, different house in the United States or Puerto Rico

**b. Where did this person live 1 year ago?**

**Address (Number and street name)**

**Name of city, town, or post office**

**Name of U.S. county or municipio in Puerto Rico**

**Name of U.S. state or Puerto Rico** **ZIP Code**

The ACS also allows for the analysis of migrants by a whole range of characteristics. However, given the modest numbers for most flows and resultant small samples, along with rules to maintain confidentiality, these data are not available for origin-destination pairs. In general, data from the ACS are limited in three ways:

- a. Despite its huge sample, the sample is not large enough to code very small areas (census tracts, small towns, etc.) even with five years of data aggregated. The example above for Los Angeles County and Manhattan represents a relatively large exchange of population for moves between counties in different states, with most exchanges actually much smaller and difficult to measure reliably as a result. Moreover, despite the detail requested in the migration question (shown earlier), complete data on origins and destinations are missing for a sizable subset of responses and, since migration only relates to a modest subset of the universe to begin with, estimates for small areas would have large margins of error. Thus, the Census Bureau does not code cases at these very low geographic levels (e.g., census tracts) for gross flows, although it is possible to look at the migration status of persons on the receiving end for broad categories of moves (e.g., lived in another state one year ago).

- b. The release of ACS data is not timely enough to affect local decision-making, given the “one year ago” question and the actual one-year lag in the release of the data. And, for smaller areas, the utility of a five-year period estimate of domestic migration is very limited – think about the usefulness of a 2015-2019 period estimate of migration, given the pandemic.
- c. Finally, this measure provides estimates on an annual basis, when data on quarterly, monthly, or even weekly frequencies are required for timely decision making.<sup>4</sup>

### ***Current Population Survey (CPS)***

The Current Population Survey (CPS) is conducted by the U.S. Census Bureau and is co-sponsored by the U.S. Bureau of Labor Statistics. Similar to the ACS, the Current Population Survey (CPS) collects information on migration status using a question about residence one year before the survey date and the geographic location of the respondent’s current residence.<sup>5</sup> One-year migration data are collected annually, as part of the Annual Social and Economic Supplement (ASEC). (Similarly, five-year migration status is based on residence five years ago compared to current residence. Five-year migration data are only collected in years ending in 0 or 5.) Movers are defined as having a different previous residence than current residence and are classified as to whether they were living in the same or different county, state, region, or moved from abroad.

The CPS ASEC data go back many decades and have been used to track the geographic mobility of the U.S. population. The richness of the CPS lies in its extensive information on the characteristics of movers, including the type of move (e.g., family, employment, housing), and the reasons for moving (e.g., new job, retirement, change in marital status, health, college, natural disaster), for general groups of movers (i.e. intra-county, inter-county, abroad one year ago).<sup>6</sup> The most recent CPS data show that geographic mobility was at an all-time low for the 2020-2021 period, with just 3.3 percent of the population migrating across counties.<sup>7</sup> The CPS data show migration rates falling substantially over the past few decades.<sup>8</sup> A major impediment is the lack of small area data, given the much smaller sample size than in the ACS – 74,000 households, which largely confines this kind of analysis to the national level.<sup>9</sup>

---

<sup>4</sup> <https://www.census.gov/acs/www/about/why-we-ask-each-question/migration/>

<sup>5</sup> <https://www.census.gov/programs-surveys/cps/technical-documentation/subject-definitions.html#migrationuniverse>

<sup>6</sup> <https://www.census.gov/data/tables/2021/demo/geographic-mobility/cps-2021.html>

<sup>7</sup> <https://www.brookings.edu/research/despite-the-pandemic-narrative-americans-are-moving-at-historically-low-rates/>

<sup>8</sup> <https://www.census.gov/data/tables/time-series/demo/geographic-mobility/historic.html>

<sup>9</sup> <https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.html>

### ***American Housing Survey (AHS)***

The American Housing Survey (AHS) is sponsored by the Department of Housing and Urban Development (HUD) and conducted by the U.S. Census Bureau. The survey has been the most comprehensive national housing survey in the United States since its inception in 1973, providing current information on the size, composition, and quality of the nation's housing and measuring changes in our housing stock as it ages. The AHS is a longitudinal housing unit survey conducted biennially in odd-numbered years, with samples redrawn in 1985 and 2015 <sup>10</sup> The 2021 AHS survey ended in October 2021, and will be released in Quarter 3 of 2022. This release will also include questions about housing search, intent to move and wildfire risk (a recent addition to the survey).<sup>11</sup>

The AHS provides information about housing and measuring changes over time, including questions about the condition of homes and neighborhoods, costs of maintenance and characteristics of people in households. For the purposes of studying migration, the AHS contains information on the characteristics of household members who have moved during the last two years. This includes whether the resident had moved more than 50 miles, the status of their previous residence, whether their housing costs increased/decreased and most importantly, reasons for moving. The insight provided on why people move, especially during difficult times, allows for a better understanding of the dynamics of migration. For example, these data have been used to examine shifts in population from cities to suburbs (Sanchez and Dawkins, 2001) and questions of displacement and gentrification in low-income neighborhoods (Ellen and O'Regan, 2010).

The AHS offers the benefit of providing information about why Americans chose to move, as well as where they moved from and the costs of living. A sample of units is visited every 2 years, offering a glimpse into who is living at a given unit over time as well as why families and individuals move. The AHS is a survey of housing units, with a total of 117,000 units included in the latest survey and oversampling is deployed for selected universes, such as renters receiving HUD assistance.<sup>12</sup> Selected oversampling is used to create estimates for selected large metropolitan areas, the smallest geographic units for which estimates are available. The challenges in using the AHS data are the absence of actual flow information and the very selective geographic coverage.

---

<sup>10</sup> <https://www.census.gov/programs-surveys/ahs/about.html>

<sup>11</sup> <https://www.census.gov/programs-surveys/ahs/about/upcoming-releases/2021-release.html>

<sup>12</sup> In 2015, a new sample was drawn, rendering a certain amount of incompatibility between pre-2015 and post-2015 samples.

## ***2. Migration as a Component of Population Change Using Administrative Records***

### ***The Census Bureau Population Estimates Program***

The Census Bureau produces a series of official estimates annually, in the years between censuses, for all governmental units. The Bureau builds the estimates starting with a base from the most recent census. A combination of administrative records and survey data are used to create estimates of the components of change for all counties in the U.S. These components are used to move the population forward and create a county estimate for an intercensal year (see Figure 1).

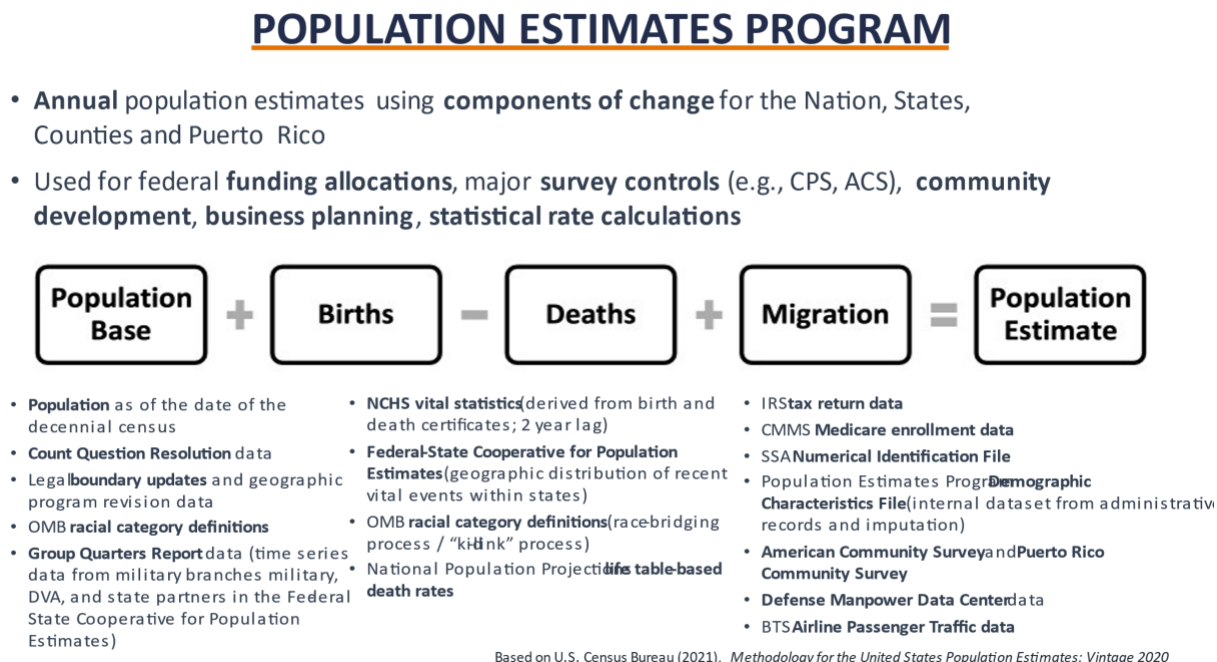
- Birth and death records from vital statistics are used to estimate natural increase.
- Data from the American Community Survey are used to estimate international migration.
- IRS records are used to estimate rates of domestic migration, based on changes of addresses on Income Tax Forms.
- Medicare enrollment data are used to estimate the population 65 years and over.

Estimates are for July 1 of each year and are created by age/sex and race/Hispanic origin at the county level. These serve as controls (i.e., weighting and post-stratification adjustments) for the ACS and a variety of other surveys internally, and for surveys by outside organizations.

Estimates of total population also are created for places and other governmental units at the subcounty level using what is called the “distributive housing unit method.” Data on housing units, occupancy rates, persons per household, and group quarters are all used to produce census tract estimates of total population, which are then aggregated to create estimates for legal, political, and administrative units (e.g., cities, towns). It is important to note that these subcounty estimates are controlled to county totals.



**Figure 1. A Summary Description of the US Census Bureau's Population Estimates Program**



### ***Issues and Limitations of Data from the IRS***

The data available from the IRS provides a window into migration between counties in the U.S. These data are used by the Census Bureau and by external researchers to gauge migration between counties, using data that are publicly available from the IRS. After geocoding the addresses on tax returns to the county level, comparing these geocodes from one year to the next provides an indication of movement across a county boundary, from the perspective of the receiving county (in-migrants) and from that of the origin county (out-migrants).<sup>13</sup>

The most recent data from the IRS are for 2018-2019, and includes the number of returns, personal exemptions, gross income, and migration flows. The number of migrating individuals is approximated from the number of returns and personal exemptions, and the size of migrating households.<sup>14</sup> In addition, county to county flows that have too few entries are categorized into seven other categories, including regional migration (migrating to Northeastern, Midwestern, Southern and Western states) and leaving the US. Corresponding data are available for migration into a given state or county.<sup>15</sup>

<sup>13</sup> [https://www.irs.gov/pub/irs-soi/99gross\\_update.doc](https://www.irs.gov/pub/irs-soi/99gross_update.doc)

<sup>14</sup> <https://www.irs.gov/statistics/soi-tax-stats-migration-data>

<sup>15</sup> It is important to note that the Census Bureau uses the absolute number of returns and persons to create *rates* of migration.

Data for 2019-2020 are expected to be released in May 2022 while data for 2020-2021 are expected to be released in November 2022. The data do not capture individuals who do not file income tax returns and their dependents. The delay in data release make it difficult to use for addressing real time or rapidly evolving issues such as COVID-19 migration patterns. In addition, due to confidentiality concerns, county-to-county flows with less than 20 returns were excluded. Finally, the location from which an individual is filing from may not actually reflect the county or state in which they currently reside.

While IRS migration data are used to create intercensal population estimates, these data are also used for survey design, business planning, community development and federal funding allocation at state and local levels. Additionally, when joined with aggregate data from the American Community Survey, county-level data allow for examinations of migration by age, between rural, suburban, and urban areas, and race and income. Migration data pieced together from income tax forms represents upwards of 90 percent of the American households and offer opportunities to measure annual trends (Foster, et. al., 2018).

Despite the near-complete enumeration, the use of income tax returns leaves several major gaps: namely economically marginalized groups, children (especially children born that year) and the elderly. Income tax returns do not capture informal economies and undocumented workers particularly well, both of which are prevalent in rural areas and/or areas with economies based around agriculture, leading to the undercount of migration in rural economies (Golding and Winkler, 2020). IRS data are also not good for capturing migration in counties with large university student populations. Finally, there was a notable decline in migration in 2015 in the IRS data, well below usual deviations in migration patterns. As a result, there are several sudden changes in migration numbers across the board between 2014-2016, a pattern that has been documented, but has not been well explained by the IRS (DeWaard, 2021).

## **2. Migration as a Residual**

This is more of a generic calculation based on the population balancing equation:

$$(\text{Pop } T_2 - \text{Pop } P_1) = (\text{Births} - \text{Deaths}) + \text{Net Migration}$$

Pop  $T_2$  = Population at time 2 (census)

Pop  $T_1$  = Population at time 1 (census)

Births and deaths from Vital Statistics

Therefore:

$$\text{Net Migration} = (\text{Pop } T_2 - \text{Pop } T_1) - \text{Natural Increase}$$

Estimates by age also can be derived using mortality, as reflected in estimates of survival derived from Life Tables (aka “survival rates”) and the change in population between two points in time. Thus:

$$\text{Net Migration} = (\text{Pop 40-44 } T_2) - (\text{Pop 30-34 } T_1 * \text{Survival Rate})$$

In other cases – for totals and age-specific estimates, net migration is assumed to be the residual population change that cannot be accounted for by the natural increase/mortality.

This “residuals approach” has two limitations. For this approach to work:

1. Accurate records from a vital statistics registry are required – data on births and deaths have been the beneficiary of substantial investments to standardize reporting and are generally reliable. However, even when overall counts are considered to be of high quality, data on the characteristics of mothers and the deceased can vary, based on the quality of reporting.
2. Accurate population counts for each decennial census – changes in population coverage (aka “undercounts” and “overcounts”) can be substantial and affect the calculation of net migration. Such was the case in New York City between 1990 and 2000, when the city’s population increased by 686,000 or 9.4 percent. Improvements in coverage associated with the first Local Update of Census Addresses (LUCA) program in 2000 was responsible for substantial part of the increase and not actually population growth. When an adjustment was performed to take the improvement in the address list into account, the actual growth of the population was estimated to be in the range of 440,000 persons or around 6 percent. Further, when the change in population is adjusted for the improvement in coverage between 1990 and 2000, the impact of net migration changed from positive to negative.<sup>16</sup>

This residual method is especially difficult to use in small areas, because of differences in coverage over time and by characteristics of the population, such as race and Hispanic origin. This approach does not provide any idea of gross flows of population. These limitations often preclude the methods usefulness, especially in small areas or with populations suspected of having coverage issues. Most important of all, the residual method does not provide a standard way to separate net domestic migration from net international migration.

#### **IV. What We Heard, Where We Are and Where We Need to Go**

The current measurement of migration is handicapped by the absence of real-time data on the movement of people, or of systems capable of determining gross flows in circumstances that would make such knowledge extremely valuable. As climate change continues to trigger events such as sea level rise, extreme weather, and conditions leading to increases in wildfires, migration data should offer a view of the repercussions for populations, in the form of

---

<sup>16</sup> When the natural increase for the decade is subtracted from the change in population, the original population count results in positive net migration over the decade (686,000-584,000=102,000) but when change in adjusted for increased coverage, this results in an actual loss associated with net migration (440,000-584,000 or – 144,000). See <https://www1.nyc.gov/site/planning/planning-level/nyc-population/census-summary-2000.page>.

dislocations due to housing loss and economic shocks. Knowing about shifts in population can help local governments and organizations be proactive, in their planning for the distribution of water resources, for encouraging new businesses, and for the promotion of new housing that is consistent with population shifts; but to be effective, such plans need to be interactive with sources that provide current data. Knowledge of who is coming to an area and who is leaving, for example on a quarterly basis, is fundamental to inform these decisions.

Thus, two general pathways for research emerge:

- Expansion and integration of existing data sources relating to content, compilation and analysis
- the identification of additional sources that can be used to monitor population movements and integrated with existing data, to create opportunities to model migration

## **V. Expansion and Integration of Existing Data Sources**

### **1. *Expansion of ACS sample and content***

The question about residence one year ago is recorded for every person aged one year and over, along with information on the location of the previous residence for those one year of age and over who moved. This level of detail represents a missed opportunity because the sample size of the ACS is insufficient to provide gross flows to and from areas of the country that would be most meaningful to small-area data users. Optimally, the availability of user-defined flow data, identified interactively, would help address important local issues. Finally, with increases in sample size and/or oversampling of small areas would come the ability to look at annual trends in migration in a meaningful way.

Imagine a business leader trying to evaluate a policy that aims to offer opportunities for local talent to discourage job-related out movement. The current five-year period estimates are insufficient as a tool for this kind of assessment. Annual flow data delivered in a timely fashion for a customized geographic area would be better than what is currently available. The increase in sample size is no panacea, however. The problems cited earlier about the ACS data collection yielding omitted or incomplete information in the migration question need to be addressed.

One task to consider is to use administrative data to fill gaps when information on a move is lacking. Changes of address on IRS tax returns, discussed earlier, and permanent changes of address from USPS records – cited later – would be two good sources of information to supplement the ACS, filling the void left by missing or incomplete migration information. Finally, the addition of well-tested content on reasons for moving from other surveys would further enhance the ability of local

decision-makers to assess local policies and plans to retain talent.<sup>17</sup> In a world that has become increasingly untethered from single employers and a function of shorter-term GIG employment, it is important for local leaders to have a barometer that can come from an enhanced ACS.

Finally, the Census Bureau needs to revisit the geography used to identify migration flows. At present, priority is given to selected geographic summary levels (i.e., metro areas, counties and Minor Civil Divisions) that omit important origin-destination pairs that are important (e.g., flows between a number of large cities). Thus, in some states, small exchanges of population are reported that are not statistically meaningful and, likely, press the envelop on disclosure, while in other states large, important geographic areas are not well represented in the data.

## ***2. Integration of ACS sample with information from LEHD***

The Longitudinal Employer-Household Dynamics (LEHD), part of the Center for Economic Studies at the Census Bureau, contains several data products used to examine the workforce. The program is part of the Local Employment Dynamics (LED) partnership in an attempt to supply state and local governments with more economic information. Data for LEHD is obtained from federal, state and local administrative sources and Census Bureau data on employers and employees. The LEHD program then combines these to produce data on employment and workers and statistics on “employment, earnings and job flows at detailed levels of geography for different demographic groups”.<sup>18</sup>

Goetz (2017) raises the possibility of combining the ACS and Longitudinal Employer-Household Dynamics (LEHD) Data. The possible linkage of ACS respondents with respondents in the LEHD dataset has the potential to add a key economic framework to migration, one that could be studied longitudinally. Specifically, the extensive data on migrants in the ACS can be combined with information on job-to-job moves in LEHD,<sup>19</sup> yielding a greater understanding of migration flows at granular geographic levels. The main limitation is that the LEHD data cannot capture migration that does not involve a job change. However, a combined ACS-LEHD dataset opens the possibility of studying migration, labor and residential mobility, as well as unemployment outcomes for job-related moves.

---

<sup>17</sup> The content of the ACS is determined through a process whereby federal agencies need to make the case for new content, based upon specific agency business that requires such data. Moreover, the ACS is a self-response instrument, which is not the case with the CPS or AHS, both of which are interviewer-administered. See:

<https://www.census.gov/content/dam/Census/library/visualizations/2017/comm/acs-questions.pdf>

<sup>18</sup> <https://lehd.ces.census.gov/>

<sup>19</sup> <https://lehd.ces.census.gov/data/>



## VI. Additional Sources of Data on Geographic Mobility

It is important to acknowledge at the outset that the following sources of data are not meant to directly translate into the number of persons who move and/or migrate. As discussed earlier, the Census Bureau currently uses IRS data to construct *rates* of gross domestic migration, not as a direct indicator of the actual *number* of persons who move. The data below can be used as indicators of movement, with substantial variation in the link between the specific source and what can be *inferred* about movements. These sources vary in content and geographic scope, which affects the generalizability of any conclusions. However, taken together, and combined with current data and methods, these new data may offer a more timely and geographically relevant insight on patterns of migration, and the status of migration as a component of population change. With the move towards model-based and synthetic estimation methods, these sources may become especially useful as *symptomatic* of migration patterns.

### 1. *Postal Change of Address Data*

The US Postal Service makes a dataset available of approximately 160 million addresses where individuals and businesses have filed change-of-address requests. These are deemed permanent changes-of-address and the data are provided on a regular basis to companies that have obtained licenses from the USPS.<sup>20</sup> (The Census Bureau actually uses data from the Postal Service as part of its geographic frame updates.) Since addresses can be geo-coded to reflect movement across county boundaries, these data can be used as a gauge of migration, or at finer levels of geography to depict local mobility. The National Association of Realtors, for example, examined changes of addresses from January to June of 2021. Like with the CPS, these data revealed a decline in geographic mobility, compared to the same period a year earlier, with movement continuing away from larger urban centers to small towns and rural areas (National Association of Realtors Research Group, 2021). Data are available at the ZIP code level and have been used by some localities to measure movements for monthly periods over the course of the pandemic (New York City Office of the Comptroller, 2021). Because data on postal changes of address are national, they allow for a standardized source of data that can be used at very granular geographic levels in combination with information from IRS and other sources.

### 2. *Private-sector sources*

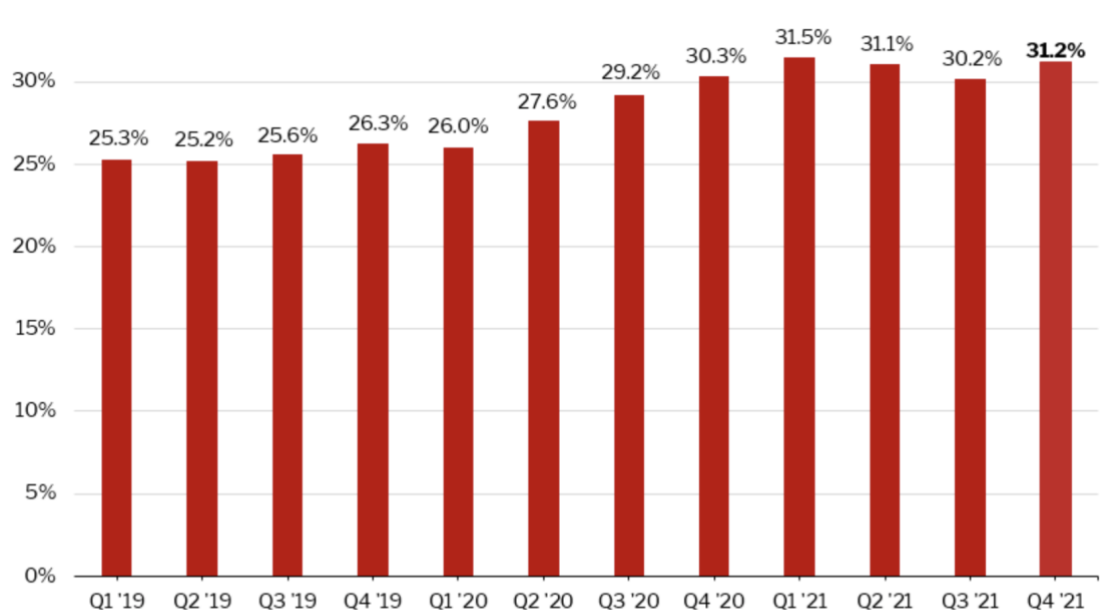
Data being harnessed by the private sector can be very instructive for migration analysis, given that a number of organizations have been leveraging a combination of more traditional data sets (e.g., loan applications, home sales) with newer sources from the Internet and social media (e.g., home searches). What characterizes many of these data sources is the timely nature of information. For example, Redfin – a real estate company – uses Internet searches to show the latest trends in the desire

---

<sup>20</sup><https://postalpro.usps.com/mailling-and-shipping-services/NCOALink>

to relocate. Using criteria to separate out more elaborate searches by users of their platform, Redfin can help identify larger migration patterns by origin and destination. These data are available quarterly, back to 2017, with the latest data for the last quarter of 2021. They show that interest in relocation to another metropolitan area – as gauged by the number of “serious” searches -- has not waned, a pattern that is different from the official data cited earlier, and as shown below (see Figure 2).<sup>21</sup>

Figure 2. Share of House Hunters Looking to Relocate Is Near Pandemic Peak  
*Percent of Redfin.com users searching for homes outside their home metro area.*



Source: Redfin user search data

**REDFIN**

One reason given for this pattern is the continuation of opportunities for remote work. There is no doubt that these data may not enjoy the level of representativeness found in official surveys like the ACS and CPS, but they can play a role when combined with other information. For example, records from CoreLogic are available on applications for home mortgage loans, a better gauge of a commitment to relocate. These data show that migration has waned in 2021, given expectations of increasing interest rates and a return to the office for many.<sup>22</sup> Moreover, data are available for origins and destinations of loan applicants and can be used to identify those areas experiencing net gains in migrants and those experiencing net losses. Figure 3 shows those areas experiencing net gains.

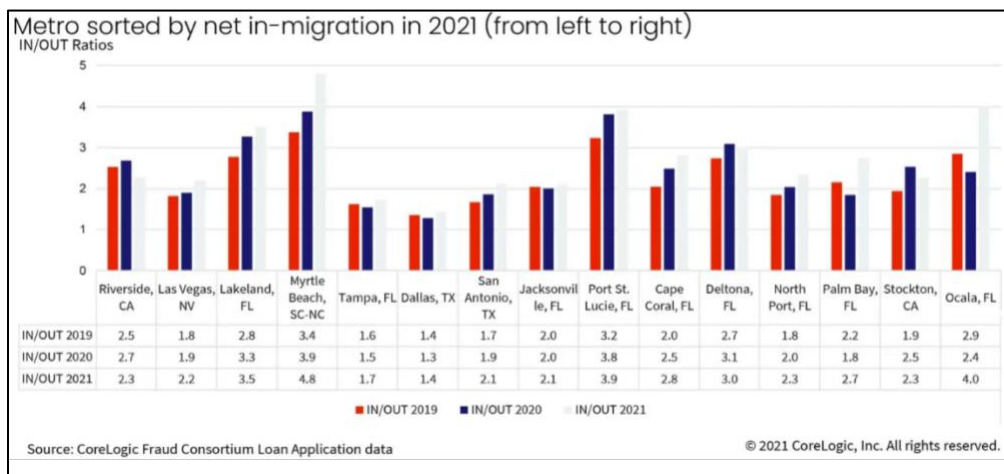
<sup>21</sup>The latest migration analysis is based on a sample of about 2 million Redfin.com users who searched for homes across 111 metro areas in the fourth quarter, excluding searches unlikely to precede an actual relocation or home purchase. To be included in this dataset, a Redfin.com user must have viewed at least 10 homes in a particular metro area, and homes in that area must have made up at least 80% of the user's searches. See:

<https://www.redfin.com/news/q4-2021-housing-migration-trends/>

<sup>22</sup> <https://www.corelogic.com/intelligence/how-homebuyer-migration-changed-in-2021/>

The salient point is that these data are current and tell a story to a much greater degree than most official sources. For example, with interest rates expected to rise throughout 2022 and the latest variant of the pandemic waning, will the Redfin platform reveal a drop in serious searches for homes in other metropolitan areas? Also, discussions can be pursued with these companies to discuss what it would take to get information at finer geographic levels? It is likely that cooperative agreements can be established to acquire the data for modeling migration. One of the goals of the curated data effort would be to facilitate such activities and encourage cooperative agreements to enhance the modeling of migration.

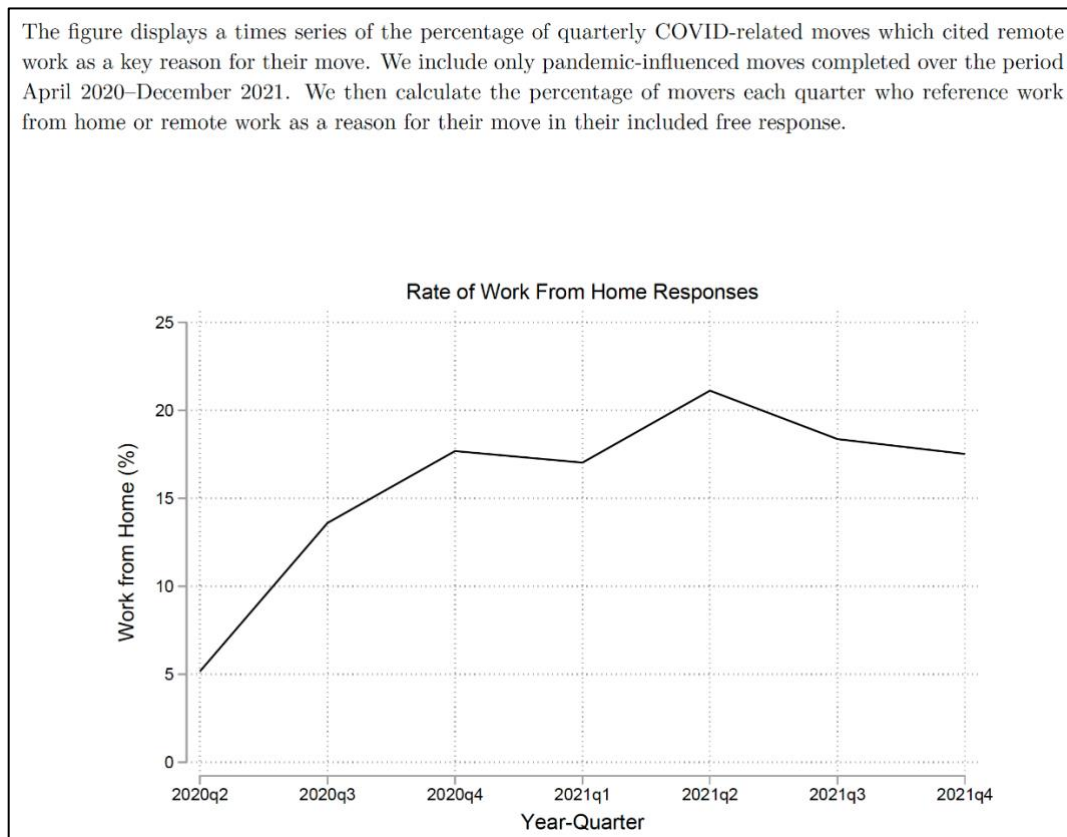
Figure 3. In/Out Ratios for Top 15 Metros with Highest Inbound Loan Applications  
(Note that each metro consists of three bars, with the rightmost bar lightly shaded.)



Another potential source of information on migration is the administrative records of companies that move households/persons. Again, these data are not a direct representation of all movers, but taken together with other sources, may be symptomatic of larger trends. One of the most interesting illustrations of such an approach can be found in the work of Haslag and Weagley (2022). They combine data for more than 300,000 interstate moves from private sources with data from the CPS. The first dataset is from UniGroup, the parent of two major moving companies, for 2017-2021 by ZIP code at origin and destination. The authors estimate that this sample represents about 5 percent of all interstate moves in the nation during the four-year period. (In addition, a subsample of movers completed a survey with information on income, age, household size, along with responses to questions on how the pandemic affected their decision to move.) These largely administrative data were then combined with information from the CPS ASEC to gauge representativeness and identify biases in the moving company administrative records, especially regarding the characteristics of interstate migrants, such as the reasons for the move. One big limitation of this comparison is that CPS data were not available for the latest year of the analysis – 2021, while these administrative records were available through May of 2021. Finally, data from the USPS on changes of address were used as a further check on the coverage of the mover data.

The results showed that the changes to everyday life brought about by the pandemic was the main factor motivating many to move. Much of this involved the change to remote work and the location flexibility that this provides, especially in higher income white collar occupations (Figure 4 below). Again, the most important point of this research is that it is not only relevant but timely, with data calculated quarterly. Moreover, while administrative data are hardly devoid of problems, the issue of incomplete responses to migration questions on some census surveys makes the integration of administrative records into migration analysis even more salient.

Figure 4. Time-Series for Work from Home-Related COVID-19 Moves



Source: Haslag and Weagley, 2022

### 3. *Digital Trace Data from Cellphones*

#### Overview

Digital Trace Data exist in many forms, from social media and from other technologies that provide data points people create as part of their daily lives. Arguably, the most useful of the digital trace sources comes from research being done using cellphone technology, which can

drive the creation of models to describe migration behavior over time and space (Florio et al. 2021). The records are generally referred to as *Call Detail Records* (CDRs). When compiled, billions of these data points are created and provide myriad opportunities for mining important information on population movements. Even though experiments with these data have been occurring for the past 15 years, advances in technology, data capture, and the synthesis of very large data sets have only recently resulted in efforts to develop standards and methods for working with these data. A number of efforts – domestic and international – have attempted to provide best practices for measuring migration with these data (Florio et.al., 2021; Lai et.al., 2019). At least some of this research has come about as a result of internal displacement within nations and across national borders, due to natural disasters and other events that result in dislocation (Luca, 2021:5-6; Bengtsson, 2011; European Commission, 2016).

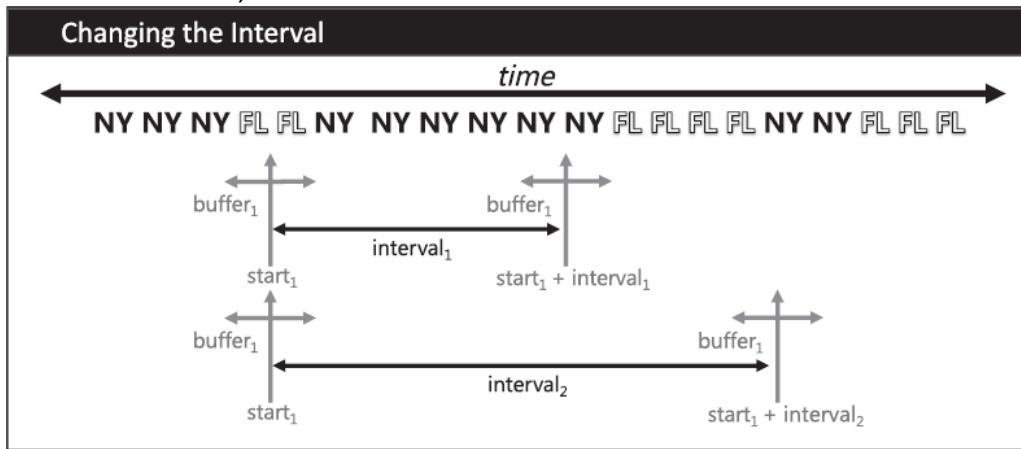
### Methods

Digital Trace Data from cellphones are different from traditional sources, in that they have the capability of allowing data users to monitor movement as it is happening. Cellphone data points have three key elements in common for the determination of movement: an anonymized individual ID, a timestamp, and a geographic location. These data allow researchers to create a range of observations around temporal point 1, called a “buffer,” or series of short-term moves, and then determine which sequence of observations at time 2 are indicative of a *transition* associated with a move across a specified boundary (Florio et.al., 2021). Since most movement is very local, the task really becomes how to define the buffers so that shorter-term transitions are eliminated from consideration, leaving transitions of longer duration that are indicative of migration to and from areas of the nation.

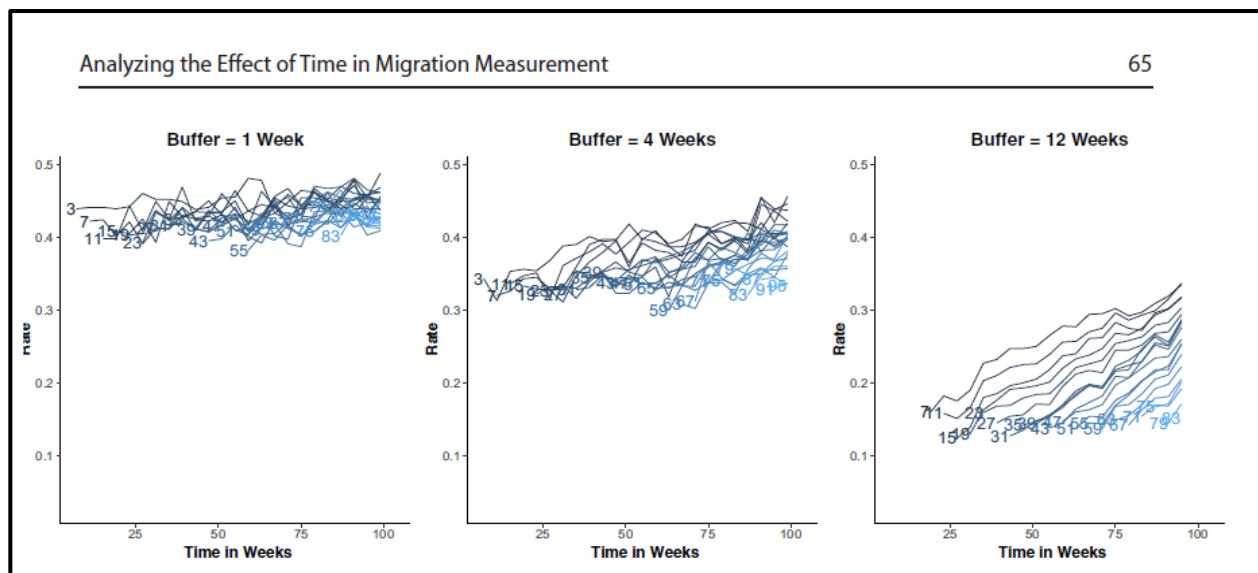
The model’s dimensions are time and distance – longer intervals between observations over greater distances between geographic points are associated with higher probabilities of being correctly identified as migration. The key is to create enough of a time duration to engender a better sense of movement that actually constitutes migration or a move across a county boundary. If the buffers are of sufficient size to eliminate the effect of short-term “noise” consisting of very local movements, then increasing time should be associated with more data points that reveal migration between New York and Florida, for example (see Figure 5). And indeed, simulations by the authors demonstrated this finding (Florio et al. 2021:65). Moreover, the authors applied this model using CDRs for the 14 regions of Senegal to create rates of inter-regional migration for 2013-14. Even though the data were just for one year, the authors were able to change start dates and observe CDR “bumps” in the time interval that were indicative of seasonal migration flows. This analysis would be especially relevant for evaluating domestic migration in the U.S., for example between New York and Florida.



**Figure 5.** This figure changes the interval and holds the seasonality (start) and residency criteria (buffer) constant. It presents the cumulative exposure to migration risk (interval). Source: Florio et al., 2021



**Figure 6.** Migration rates are derived from simulated data. Each line tracks the rate of individuals classified as migrants as the interval increases from a specific reference point, while the buffer is held fixed. As the time buffer increases, the simulated migration rate declines. These estimates provide a range of estimates to assess data quality. Source: Florio et al., 2021



An interesting question concerns how migration rates based on CDRs compare with those based on traditional surveys and censuses. Analysis by Lai et al. (2019) using CDRs to measure migration between regions of Namibia for 2011-2012 showed a solid relationship between data from their national census and their CDR-based model ( $r=0.91$ ). Similarly, inter-regional rates

using CDRs for Portugal in 2006-2007 also correlated highly (.0.94) with the migration picture from their census (Hankaew, 2019).

#### Key Issues

1. *The automated nature of “responses.”* Greater consistency in reporting is possible, compared with self-reports and missing data on surveys, but CDRs do not necessarily correspond with the same single individual over time (Tjaden, 2021).
2. *Coverage* - While the size of CDR data sets is impressive by any measure, it is difficult to determine how representative these data may be of a population. The ubiquitous nature of cellphone use may lead researchers to believe that the scope of coverage is not an issue. However, the geographic distribution of cell towers, which are the intermediary in the transmission of location data, may not allow for good coverage of the population using cellphones and the cellphone users themselves may not be representative of the general population (Luca, 2021).
3. *Confidentiality* – Although research has shown that cellphone data can be anonymized, the specter of surveillance is a major issue and one that may give rise to limitations on the availability of data. Efforts by cellphone users, government agencies, public utilities, and private corporations may greatly restrict access to these data. At the very least, user agreements will likely become more elaborate, which may compromise the scope and form of the data that can be made available for research (International Organization for Migration, 2018: Luca, 2021:8-9).
4. *Other issues* – There are many issues that will need to be addressed to use cellphone data. These include establishing partnerships to obtain these data; ethical dimensions with respect to privacy and protection of human subjects; and introducing approaches such as Bayesian methods to combine different migration data with other data (e.g., demographics) in a consistent way (Hughes et al. 2016).

## VII. Summary and Implications for the Curated Data Enterprise

*Content Curation* usually refers to a process where data are selected, organized, and integrated as a means of adding value to an understanding of what the data represent. Such an enterprise relies on expertise in understanding the needs of the data user community and consensus on the data that are most useful towards understanding the interplay between migration and local labor markets, given changes in the location of work (e.g., the ability to work from home), especially during the pandemic.

Local planners have considerable interest in how changes in the larger economy and the world of work are reflected in migration patterns and population change in their localities. Yet current

measures are incapable of keeping pace with these changes. The pandemic has made real-time migration data more important to monitor trends as has the impact of natural disasters, especially because of extreme events related to climate change. Areas of the country that have been affected by wildfires, extreme weather, flooding, and droughts associated with longer-term watershed issues, have all experienced changes in their patterns of migration. These events precipitate movements that need to be monitored in something close to real-time, so local planners can react quickly. State and local governments require population data for revenue projections, for the distribution of resources, the identification of needs, and the establishment of priorities and strategies to address those needs.

What is clear from this evaluation of migration data, is that no single source of information will suffice to describe internal migration patterns in the U.S. Most current measures of migration conceptualize it as permanent movement across a county boundary, focused on a sequence at two fixed points in time – an origin point and one at destination. In the case of surveys, respondents are asked about their location at an earlier timepoint, as with the ACS and CPS. Survey response, in general, has declined over the past two decades and, in the case of data on migration, the problem of missing and/or incomplete responses is particularly acute, which is an especially difficult problem, given the relatively small subset of sample cases associated with migratory moves. Administrative records, such as tax returns, reveal these origin and destination points as addresses at two points in time and can help fill the missing data void in survey information; but, here too, these data were not designed for the purposes of statistical compilation and issues involving missing data and the scope of population coverage are important limitations. Both survey and administrative records provide snapshots of the population and in both cases, these are collected retrospectively with a substantial time lag for results to be published. As the data users have made clear, increased access to timelier data on migration needs to be a priority and that existing data are insufficient along those lines, especially given the impact of the pandemic on small places. Some data sources have the potential to measure migration, such as utility data (i.e., water, electricity, and gas) but are not readily available from public utilities due to confidentiality concerns and the challenge of obtaining data from each utility.

Other timelier sources are now available on various platforms to identify and study migration, like records of persons who have indicated changes of address to the postal service, applied for mortgage loans, sought real-estate in other parts of the country, or moved using commercial movers. These data can be reported for sub-annual periods and tracked to identify trends. Moreover, digital trace data represent the next level or frontier in the study of migration, with cellphone CDRs being at the heart of digital data-mining efforts. While experiments have been conducted using data from the ever-evolving digital world, it is only with more recent advances in technology, data capture, and the synthesis of large data sets that standards and methods are being developed for the compilation and analysis of these data. Such data are the next frontier and need to be incorporated into models that draw more accurate and timelier pictures of migration behavior.

All of this takes place in an environment where access to the details of collection and processing of data may be restricted, due to respondent confidentiality or the proprietary nature of certain data products by the private and/or non-profit sectors. Therefore, in addition to the technical issues that need to be addressed, any research roadmap needs to identify ways of cultivating partnerships that will result in access to the information and data required for a proper assessment. But first it is important to identify the datasets that are most relevant for the task at hand by getting a handle on the products the CDE seeks to produce.

#### **a. Starting with the End Products**

The goal of the CDE is to produce useful products for the decision-makers who are formulating policies, designing programs, and implementing strategies for local applications related to climate change, watershed issues, emergency preparedness and mitigation, housing policies, and business and economic development. One goal of the CDE is to offer information on migration behavior to inform those decisions. Using information from the *Listening Sessions*, it is instructive to ask questions about the applications of migration data:

- A local planning agency is trying to get a handle on the impact of the pandemic on their local labor market. They want to better understand the movement of people into and out of their jurisdiction. What data can be provided for insight into these trends?
- A local non-profit wants to develop a picture of population change for a custom geographic area for more efficient delivery of services. The components of population change are an important part of this assessment, along with the reasons why people move – how are those reasons related to the services they provide?
- A watershed issue related to climate change has triggered a need to project water requirements based on population forecasts. A critical part of this assessment involves an analysis of migration patterns in the area served by the watershed and places that may become part of the area with given new development proposals. What products may be available to understand current migration as a component of population change?
- An emergency management agency is trying to cope with flooding after an extreme weather event. Many residents of the municipality have been forced to relocate to other places and local town officials are trying to get some idea of the size of these outflows, along with the destinations of those who have moved to another county/state. They may also want to know who is likely to return? What kind of data product could be provided to help with these local officials?

## **b. Needed Research: Ten Paths to the Creation of Relevant Products**

With these questions in hand, a path for research can now be identified – one that will lead to the development of data that can serve as a foundation for decision making. Key research questions now emerge:

1. What level of geographic detail and/or subpopulations are most relevant?
2. What survey data will be of use for framing the current state of migration for decision-makers at state and local levels?
3. What mechanisms exist for monitoring and projecting migration patterns by reason for migration and could potentially serve as the basis for looking forward and building a strategic plan?
4. The samples used for many of the surveys described in this use case are national in scope, but many have substantial detail. What kind of detail is most useful for local area applications, and where can it be found?
5. What kinds of data are available for small areas – counties, cities, towns, and other entities where decision-making occurs?
6. What actions need to occur to obtain access to Cellphone Call Detail Records (CDRs) for use by the Census Bureau?
7. How do confidentiality requirements, especially those surrounding the Census Bureau’s current Disclosure Avoidance System (DAS), affect the capacity of the CDE to gain access to data that can be shared with the data user community?
8. How can the CDE develop the interactive partnerships necessary to gain access to the data necessary for content curation in a manner that brings value to **all** parties involved?
9. How best can the creation of models integrating survey, administrative and digital data with data from the four Frames<sup>23</sup> be achieved? For example, how can data on postal changes of address from the geographic frame be best utilized for those studying migration as part of the demographic frame?
10. How can data users be given access to systems that enable them to use the data interactively for user-defined geographic areas?

---

<sup>23</sup> Through the Census Bureau’s Enterprise Frames Program, four key statistical frames are being linked to create a new universal statistical frame (Ratcliffe, 2021, U.S. Census Bureau, 2021b, Keller et al., 2022). Their goal is to link together using geocodes and other keys from the Business Register of establishments, the Master Address File of housing units, the Longitudinal Employer-Household Dynamics (LEHD) jobs frame, and a demographic frame of individuals into a unified resource. This universal statistical frame will form the foundation, or scaffold, for organizing data and information within the Curated Data Enterprise (CDE).



## REFERENCES

- Anderson, Margo, C. Citro and J. Salvo, eds. (2012) *Encyclopedia of the U.S. Census* (2<sup>nd</sup> Edition), Thousand Oaks CA: CQ Press.
- Bengtsson, L, X Lu, A. Thorson, R. Gardfield, and J. Von Schreen. (2011). "Improved Response to Disasters and Outbreaks by Tracing Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti." *PLoS Medicine*, 8(8), e1001083.
- DeWaard, Jack, M. Hauer, E. Fussell, K.J.Curtis, S.D. Whitaker, K. McConnell, K. Price, D. Egan-Robertson, M. Soto, and C. Anampa Castro. (2021). "User Beware: Concerning Findings from Recent U.S. Internal Revenue Service Migration Data," *Population Research and Policy Review*, <https://doi.org/10.1007/s11113-021-09663-6>
- Edmonston, Barry (2012). "Internal Migration," in *Encyclopedia of the U.S. Census* (2<sup>nd</sup> Edition), Thousand Oaks CA: CQ Press. Pp. 284-289.
- Ellen, Ingrid Gould and O'Regan, Katherine M., How Low Income Neighborhoods Change: Entry, Exit and Enhancement (September 1, 2010). US Census Bureau Center for Economic Studies Paper No. CES-WP- 10-19, Available at SSRN: <https://ssrn.com/abstract=1687759> or <http://dx.doi.org/10.2139/ssrn.1687759> .
- Florio, Lee, Emilio Zagheni, Guy Abel, Johnathan Hill, Gabriel Pestre. Emmanuel Letouzé, and Jixuan Cai (2021) "Analyzing the Effect of Time in Migration Measurement: Using Georeferenced Digital Trace Data." *Demography* (2021) 58(1):51–74
- Foster, Thomas B., M. Ellis, and L. Fiorio (2018) "The Opportunities and Challenges of Linked IRS Administrative and Census Survey Records in the Study of Migration." CARRA Working Paper 2018-06, Center for Administrative Records Research and Applications U.S. Census Bureau Washington, D.C.
- Goetz, Christopher F. (2017). "The Potential for Using Combined Survey and Administrative Data Sources to Study Internal Labor Migration," *Working Papers* 17-55, Center for Economic Studies, U.S. Census Bureau.
- Golding, Shaun A. and R. L. Winkler. (2020). "Tracking Urbanization and Exurbs: Migration Across the Rural–Urban Continuum, 1990–2016." *Population Research and Policy Review* (2020) 39:835–859, <https://doi.org/10.1007/s11113-020-09611-w>
- Hankaew, Soranan, S. Phithakkitnukoon, M.G. Demissie, L. Kattan, Z.Smoreda, and C. Ratti. (2019). "Inferring and Modeling Migration Flows Using Mobile Phone Network Data." *IEEE Access*, Digital Object Identifier 10.1109/ACCESS.2019.2952911 .

Haslag, Peter H. and Weagley, Daniel, From L.A. to Boise: How Migration Has Changed During the COVID-19 Pandemic (March 18, 2022). Available at SSRN: <https://ssrn.com/abstract=3808326> or <http://dx.doi.org/10.2139/ssrn.3808326>

Hughes, Christina, E. Zagheni, G.J. Abel, A. Wiśniewski, A. Sorichetta, I. Weber, and A.J. Tatem. 2016. "Inferring Migrations: Traditional Methods and New Approaches based on Mobile Phone, Social Media, and other Big Data" Report prepared for the European Commission project #VT/2014/093. Programme for Employment and Social Innovation. <https://data.europa.eu/doi/10.2767/61617>.

International Organization for Migration (IOM) (2018). "Big Data and Migration: How data innovation can serve migration policymaking." *Data Bulletin*, Issue No. 5, ISSN 2523-5060.

Keller, S., Prewitt, K., Thompson, J., Jost, S., Barrett, C., Nusser, S., Salvo, J., Shipp, S. (2022). A 21st Century Census Curated Data Enterprise. A Bold New Approach to Create Official Statistics. Proceedings of the Biocomplexity Institute. Technical Report 2022-1115. <https://doi.org/10.18130/r174-yk24>

Lai Shengjie, E. zu Erbach-Schoenberg, C. Pezzulo, N.W. Rukanonchai, A. Sorichetta, J. Stelle, T. Li, C.A. Dooley, and A.J. Tatem. (2019) "Exploring the use of mobile phone data for national migration statistics." *Palgrave Communications*. <https://doi.org/10.1057/s41599-019-0242-9> | [www.nature.com/palcomms](http://www.nature.com/palcomms).

Luca, Massimiliano, G. Barlacchi, N. Oliver, B. Lepri (2021). "Leveraging Mobile Phone Data for Migration Flows." arXiv:2105.14956v1 [cs.CY].

National Association of Realtors Research Group (2021). Migration Trends Report: 2021 January to June, Washington, DC. [www.nar.realtor/research-and-statistics](http://www.nar.realtor/research-and-statistics).

New York City Office of the Comptroller (2021). "The Pandemic's Impact on NYC Migration Patterns." November. <https://comptroller.nyc.gov/reports/the-pandemics-impact-on-nyc-migration-patterns/>.

Ratcliffe, M. (2021). Frames Program Update Census Scientific Advisory Committee Fall Virtual Meeting: September 23-24, 2021 <https://www2.census.gov/about/partners/cac/sac/meetings/2021-09/presentation-frames-program-update.pdf>

Sanchez, Thomas W. & Casey J. Dawkins (2001) Distinguishing city and Suburban Movers: Evidence from the American housing survey, *Housing Policy Debate*, 12:3, <https://doi.org/10.1080/10511482.2001.9521420>.

Tjaden, Jasper. (2021) "Measuring Migration 2.0: A Review of Digital Data Sources." *Comparative Migration Studies*. <https://doi.org/10.1186/s40878-021-00273-x>

U.S. Census Bureau (2021a). *Methodology for the United States Population Estimates: Vintage 2020*.

Washington DC, May (Version 2). <https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2010-2020/methods-statement-v2020-final.pdf>

U.S. Census Bureau. (2021b). U.S. Census Bureau Budget, Fiscal Year 2022. U.S. Department of Commerce, Economics and Statistics Administration.

[https://www.commerce.gov/sites/default/files/2021-06/fy2022\\_census\\_congressional\\_budget\\_justification.pdf](https://www.commerce.gov/sites/default/files/2021-06/fy2022_census_congressional_budget_justification.pdf)