Data and Models for Alaskan Migration

Eddie Hunsinger

Alaska Department of Labor and Workforce Development
eddie.hunsinger@alaska.gov

November 28, 2018

This file includes some edits and consolidation - EddieH 12042018

[THIS FILE INCLUDES SOME EDITS AND CONSOLIDATION - EDDIEH 12042018]

Thank you for attending this talk about Alaska migration data and models. I'll first give some background and overview info, about what I work on and and Alaska population for context, then explain more about the PFD data, how we use it, and some of its strengths and weaknesses in comparison with two other sources for Alaska migration information: the American Community Survey (ACS) and US Internal Revenue Service (IRS) tax statistics. I'll then present some comparisons of the PFD and ACS data on migration by age, and challenges in reconciling them. Finally, I'll review some different formats of Alaska migration by age profiles, and characteristic changes in them for data from 2000 to 2005, 2005 to 2010, and 2010 to 2015.

Images from laborstats.alaska.gov

Most US states have some staff who work on population data (projections, and many states make their own estimates as well), and partnerships programs to help ensure accurate and useful data from the US Census Bureau. I lead a three-person team (Eric Sandberg, Liz Brooks, and me; very grateful for their expertise) that does these things in Alaska.

In Alaska we actually have and get to use a terrific administrative data resource: the Alaska Permanent Fund Dividend (PFD) applicant database. The Alaska Permanent Fund is an investment fund that's seeded by oil revenue, and it's currently worth about 60 billion USD. Each year, every resident of Alaska (meeting certain eligibility requirements) may apply for a dividend check from the fund, and a dividend check will typically be something like a couple thousand dollars. Applicants represent about 90 percent of the state's 740 thousand residents.

Where other states have to use things like tabulations of vehicle registrations and school enrollment to make independent estimates, we are lucky to have access to microdata with terrific coverage.

We geocode it (a huge task that makes all its use possible, and all thanks, along with the migration tabulations, to Eric Sandberg) and use it extensively each year to produce population and migration statistics (and colleagues use it for many employment stats too, but Ill focus on population, and especially migration data). And I'll talk more about the PFD data, and comparison to a couple other sources, in a bit.

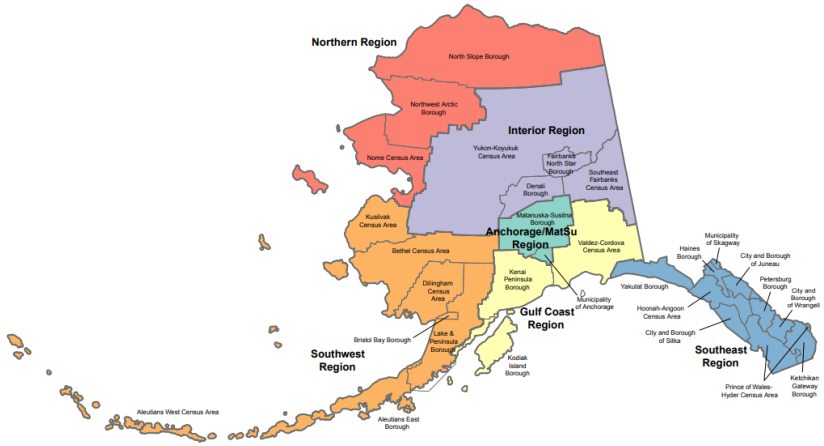


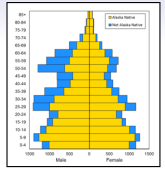
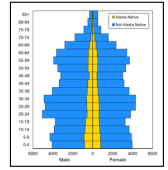
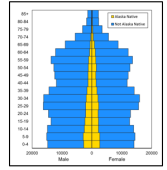
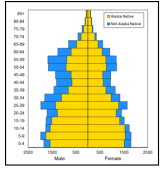
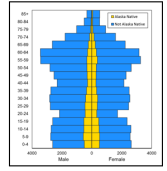
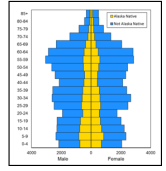
Image from laborstats.alaska.gov

This is, of course, a map of Alaska. As you probably know, it is a very large US state - it covers more than 1.7 million square kilometers of land - and it has a small population for its size just 737,000 people as of our latest estimates.

Alaska has 29 "boroughs and census areas," and for research we also refer to six "economic regions" - you can see those divided by different colors here. Connecting road systems go through just these seven boroughs and census areas in the Interior, Anchorage/Mat-Su, and Gulf Coast regions; and a into a limited parts of Yukon-Koyukuk Census Area and to oil fields in North Slope Borough; and then Haines, Skagway, and a little community called Hyder in Southeast, but there's also a state-run ferry system that connects many southern coastal communities, and airfields all across the state.

And there are more than 300 communities across the state, including more than 150 in the completely remote Southwest and Northern regions, and Yukon-Koyukuk Census Area. Today if I say "western" or "northern" Alaska, I am referring to remote communities - if you go west or north of here [].

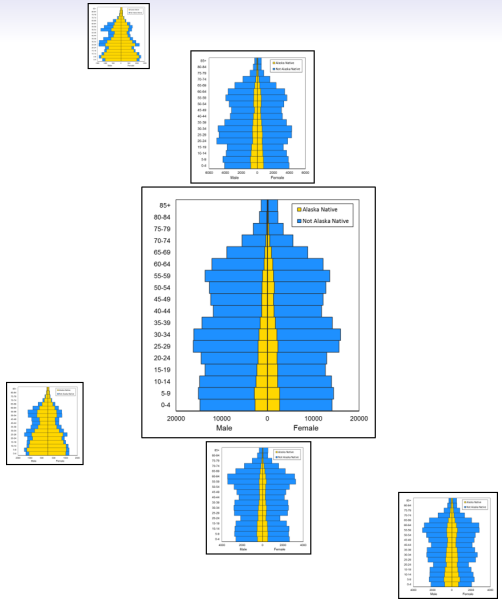
The western and northern communities are younger, and typically majority Alaska Native, the Southeast and Gulf Coast regions are typically older, and the largest settlements are in the Anchorage/Mat-Su and Interior regions, around the cities of Anchorage and Fairbanks.

Data from laborstats.alaska.gov

To give a better idea of the demographic characteristics of the six regions, here are population pyramids showing the Alaska Native (in yellow) and non-Native populations. As you can see - I tried to place them according to their location on the map.

You can see the Southwest/western and Northern regions are younger (this helps population growth, with relatively fewer people at highest ages), and majority Alaska Native, and the Southeast and Gulf Coast regions are older. If you look closely you can also see some skewed sex ratios for working ages, especially due to military in the Interior region, and mining and oil, and seafood industries in Northern and Southwest Alaska. A statewide pyramid would look most similar to the Anchorage/Mat-Su region.

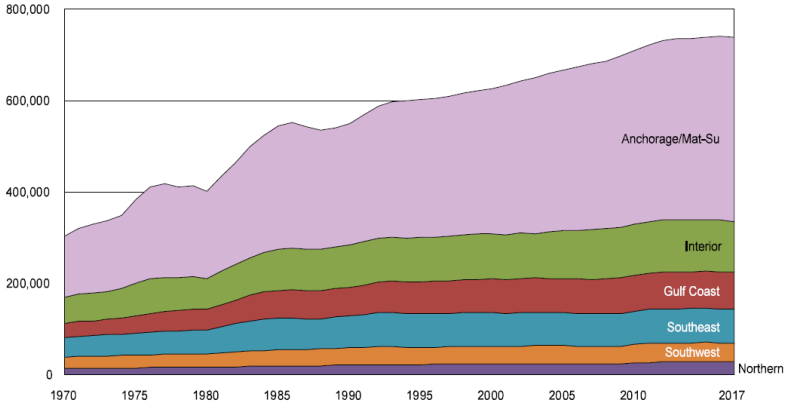


Data from laborstats.alaska.gov

And these are adjusted to give an idea of their relative sizes.

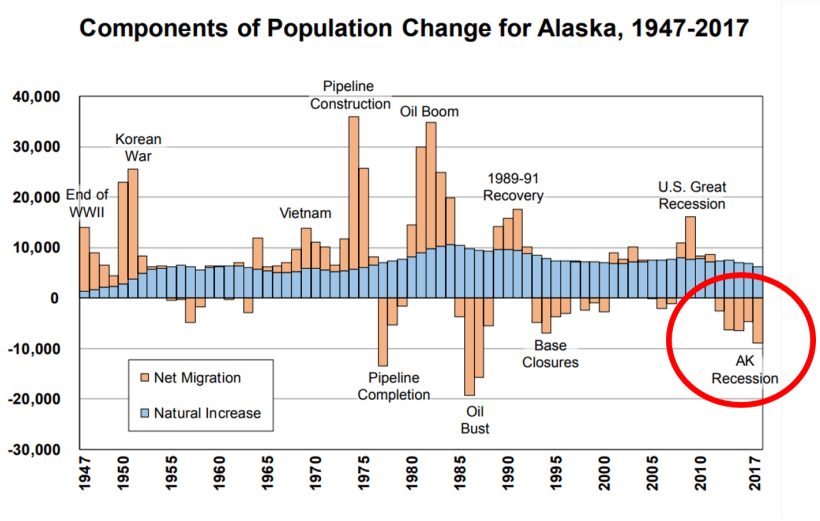
Still though, there are hundreds of disconnected/discrete communities across the less populated regions.

Population by Alaska Region, 1970 to 2017



From laborstats.alaska.gov

This slide is population change by region over time - 1970 to present. You can see the last few years have been pretty flat - actually remarkably flat, and this is the longest period of no-growth on record.



From laborstats.alaska.gov

Looking at the statewide components of change, you can see there have been big swings, and the latest downturn is at the same time as a statewide recession (deemed a recession based on employment growth rather than GDP), and this is the longest period of net migration losses we've ever seen. The recession is mostly connected with oil prices though (oil, and oil prices, is a crucial Alaska's economy), and the big drop in oil prices we saw a few years back actually happened after we were seeing significant net migration losses.

I also always like to highlight that overall migration flows are much larger and less volatile than net migration. About 5 to 7 percent of Alaskans arrive in the state in a given year, and about 5 to 7 percent leave in a given year. Due to military and our mix of industries, this is somewhat bit higher than other other states, which are more like 3 to 5 percent to and from, but still very steady. Even in a very bad year economically, more than 30,000 new people arrive in the state. At the borough and census area level, with in-state migration added, the rates are about 5 to 10 or 12 percent to and from each year.

Alaska Migration and U.S. Recessions

How the Lower 48 economy affects moves to and from Alaska

By **NEAL FRIED**

In 2009, as the nation was in a deep recession, we published an article in the June issue of *Trends* asking whether that recession, like others in the past, would affect Alaska's migration patterns.

Historically, fewer people leave Alaska during a national recession and more move north seeking job opportunities. The reverse is also true; when the national economy is flourishing and Alaska has a downturn, Alaska is more likely to register net migration losses.

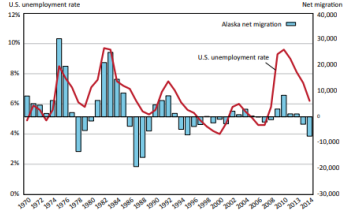
Now that the U.S. recession has played itself out, it's clear it was no exception — Alaska gained population each year from 2009 to 2012 through net migration, or migration in minus migration out. (See Exhibit 1.)

In, out migration usually cancel out

Alaska has significant population turnover every year. During the past decade, typically between 40,000 and 50,000 people moved to and from Alaska annually. These common, relatively large migration flows aren't always tied to economic conditions, as people move for a range of reasons including military rotations, school calendars, retirement, or changing family responsibilities.

1 Alaska Migration, U.S. Jobless Rate Track Together

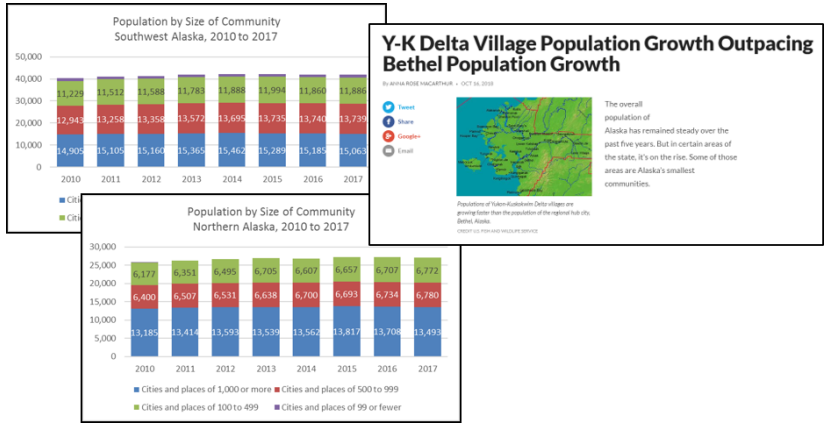
1970 to 2014



Source: Alaska Department of Labor and Workforce Development, Research and Analysis Section

From laborstats.alaska.gov

I wanted to note briefly - has been of interest this week - this is an article by Neal Fried from a few years back (his second on the topic, actually) that discusses an apparent relationship of Alaska net migration and the national unemployment rate - it is interesting how they may connect, and learning how the connection might be tested.



From laborstats.alaska.gov, and kyuk.org

And just an interesting side note from recent years: The western and northern regions have been some of the state's fastest growing areas in recent years, and indeed, we've actually been seeing stronger growth in small villages in western and northern Alaska than in the larger "hub" communities.

Images from pfd.alaska.gov, census.gov, and irs.gov

So turning to comparison of data sources: There are three sources of migration data for Alaska: Alaska Permanent Fund Dividend applications, IRS Tax Statistics, and the American Community Survey.

Alaska's statewide Permanent Fund Dividend (PFD) applicant database covers around 90 percent of the estimated usual resident population. Applicants each have Unique IDs in the system and SSNs are provided, which we can use to link them from year to year and to some other data sets. Fields include physical address as well as mailing address, date of birth, place of birth, sex, uniform military... Sounds so ideal!

But there are many drawbacks for sure. Applicants must reside in the state for at least one calendar year to be eligible, creating a lag on new migrants to the state that must be reconciled - typically by using averages of multiple years or by calculating in-migration as a residual. Also, the residence definition is different than the Census Bureau's concepts of "usual residence" (used for the Census and annual population estimates" and "current residence" (used for American Community Survey data collection). And special populations are often not as well-covered - these include military and workers who also have residences out of state. Also, there are problems of address reliability, and ambiguity in residence for many Alaskans, and challenges of geocoding rural addresses. And people who just don't apply in a given year, or apply with a different address, may be counted inaccurately as migrants. And because it's not intended as a migration data source, many options to improve it (such as changing questions) are not available.

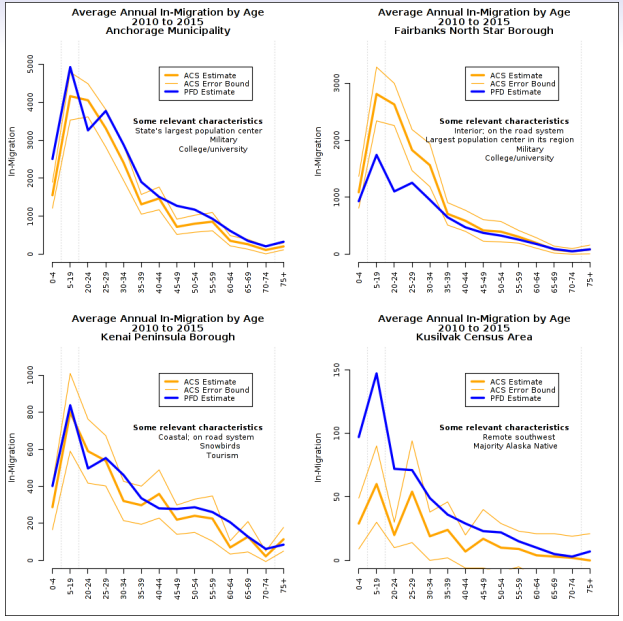
Images from pfd.alaska.gov, census.gov, and irs.gov

Really, the PFD data has a very different objective than population or migration estimation, and misses a lot of people, and some groups missed more than others a big challenge.

But we still use it in three different ways that we feel comfortable with for population and migration data production: (1) Where we can benchmark it to a reliable complete count underscoring the importance of a complete census for the most used or valuable variables. (2) Where we can effectively supplement and improve (not just substitute) information and not harm information. (3) Where a more raw form of it is useful to some end-users (no doubt fewer, but some), and we can describe for users readily, plainly, and specifically what the data is.

The ACS data is great because it is intended as a migration data source, but it is based on a small sample of the population, and its definitions are a bit tricky, including especially that data collection is based on "current residence" but weighted to estimates based on "usual residence" - this creates some issues I think for some regions of Alaska.

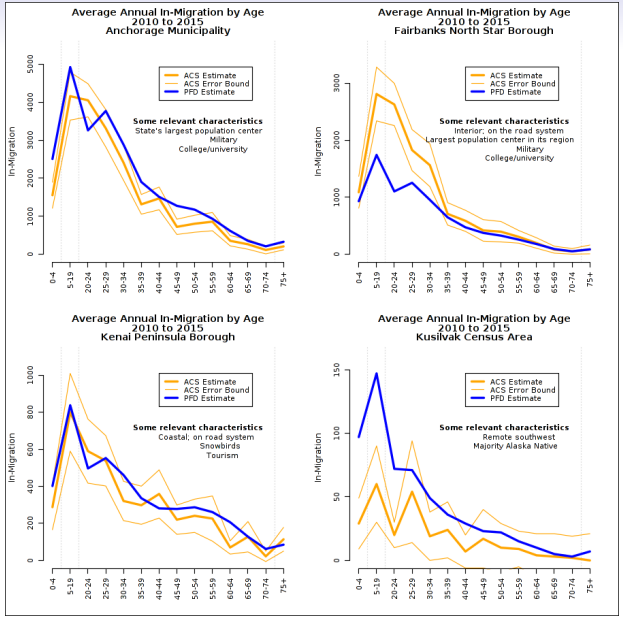
IRS-based data, which I don't look at today, are much more limited for use by state agencies like ours. The Census Bureau can use IRS microdata to make great estimates down to the county (or borough in Alaska) level, but the only public migration tabulations are for total flows (and not by demographic characteristics). Additionally, their coverage for low-income regions, where many households don't need to file, is worse.



From shiny.demog.berkeley.edu/eddieh/AKMigrationAgePFDtoACSCompare

Looking at the PFD and ACS data, these graphs are in-migration by age groups (note 5 to 19 is combined) to selected areas, from a special migration tabulation that we produce every five years (blue is PFD-based, orange is ACS-based), and for which we do (for the out-mover side) make a very simple (and I especially hope clear-for-users) adjustment for deaths by age. The graph to the top-left covers Anchorage, the largest borough in the state - about 300,000 people. I provide some info on characteristics for each area. Fairbanks North Star Borough (top-right) (total population 100,000) is a major population center in Interior Alaska its also home to significant military and student populations Kenai Peninsula (bottom-left) (population 65,000) is south of Anchorage and has larger seasonal populations and a few population centers snowbirds and generally a higher median age (42 for usual residents) Kusilvak Census Area (bottom-right) (population 8,500) covers 13 remote communities in southwest Alaska, with no large population center population there is the youngest in the state (median age 24), over 90 percent Alaska Native, and relies significantly on subsistence.

One thing that stands out to me is how consistent the PFD-based data/shapes are compared to the ACS-based data.

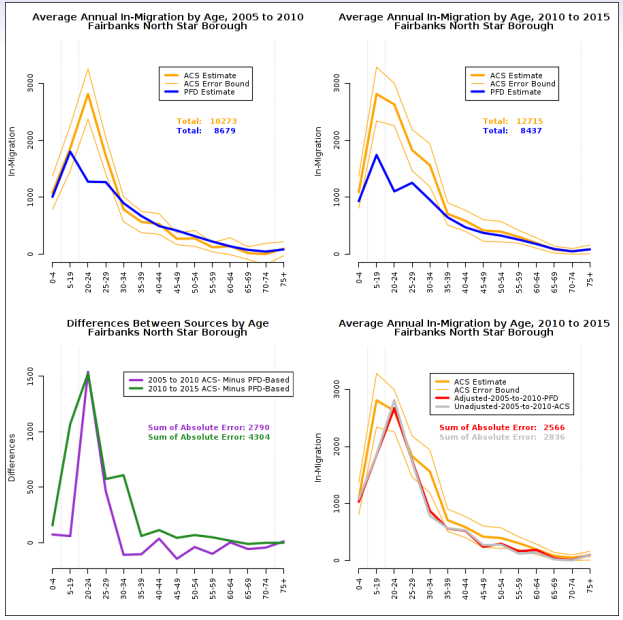


From shiny.demog.berkeley.edu/eddieh/AKMigrationAgePFDtoACSCompare

Some important issues with the PFD-based data: From earlier comments, there are features of the PFD data that can push/bias them higher than actual migration, such as people not applying in one year (or applying in a different place) even though they didnt move. And there are features that will bias them to be lower than actual migration, such as people dont reside in the state long enough (even up to two years) to be eligible - I believe can see here it for the Fairbanks populations (military, students and young workers). These biases must cancel each other out some, but its not readily known where or how much. The different factors causing errors affect some areas more than others, and maybe some of these could be guessed at or estimated by an expert, but not readily with precision. And so important, I think: they dont in their basis aim to represent fair/equitable coverage for different areas. The goal of the PFD applicant database isnt to track migration, so again, some valuable options to improve the data are missing. And the data arent useful for everyone some users may need a more conclusive (less-biased) estimate, or an estimate with adjustments for more of the biases.

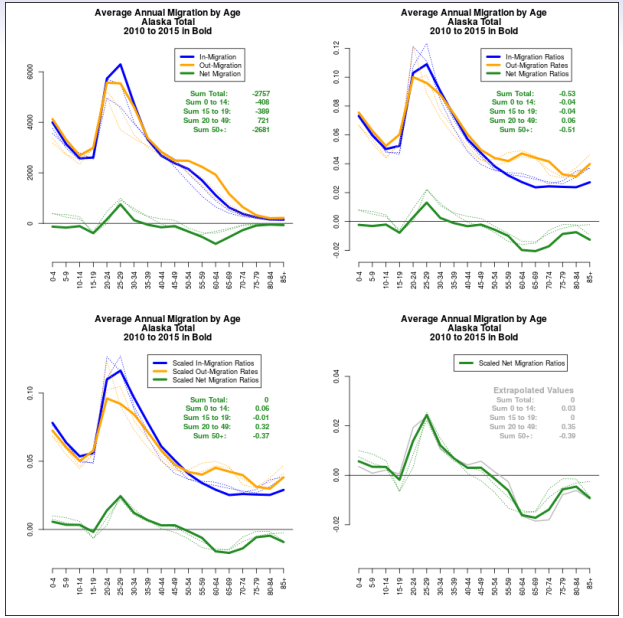
The ACS data must have their own non-sampling error too (in addition to the sampling error): Here (as big as it is for Kenai Peninsula and Kusilvak Census Area) we see just the sampling error (which would be inconsequential for the PFD-based sample size). And the challenges of residence definitions and low response or poor response.

Like the graphs by age here, comparing PFD-based and ACS-based total migration between boroughs and census areas (county-to-county migration) also showed many differences that were bigger than the ACS sampling error.



From shiny.demog.berkeley.edu/eddieh/AKMigrationAgePFDtoACSReview

This is a closer look at just Fairbanks North Star Borough - from a "Shiny" application I posted online - and trying to reconcile differences for the PFD and ACS data. The top shows two different periods, the bottom-left shows differences between the sources, and the bottom right shows adjustment applied to to try and predict ACS with PFD. Unfortunately, from looking at areas across the state, this simple way doesn't seem to do better than simply re-using the earlier-period ACS data. I think it's very interesting though, because (1) as we flip through the different areas we see a lot of instability in the ACS data between the two different periods - much more than in the PFD-based data and much more than expected, and also that (2) info from the PFD-based data doesn't seem to help. Maybe this is just a quirk of the two periods of ACS data - ACS was new in the mid-2000s - but it stands out for sure.

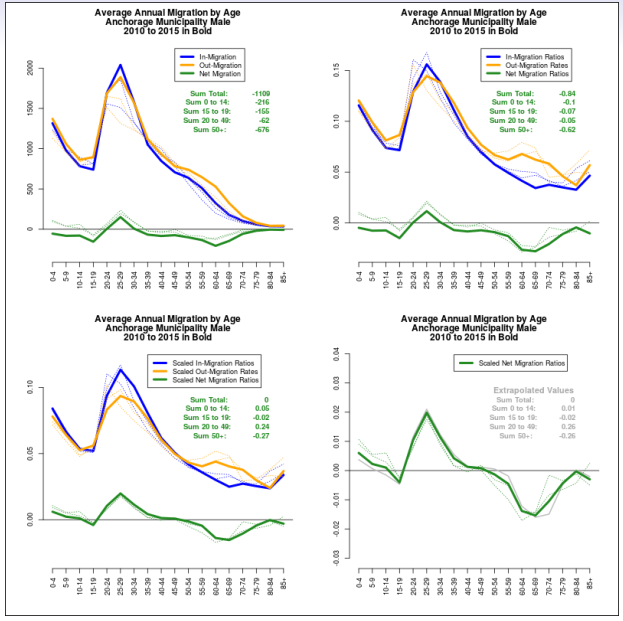


From shiny.demog.berkeley.edu/eddieh/AKPFDMigrationReview

Here is another interface, this one to look more closely at the PFD based data, and the age groups are more consistent so these look like typical migration profiles. The top right shows age-specific rates, and I give values for gains and losses by age groups. We see net losses for post-high school, and for ages 50+ in Alaska. I sum age specific rates, much like TFRs to use as indices. I'm using 50+ because when people talk about migration of seniors often 50+ migration is especially useful and relevant, since by 65+ much of the migration story is completed - and Rogers-Castro retirement migration often starts at 50.

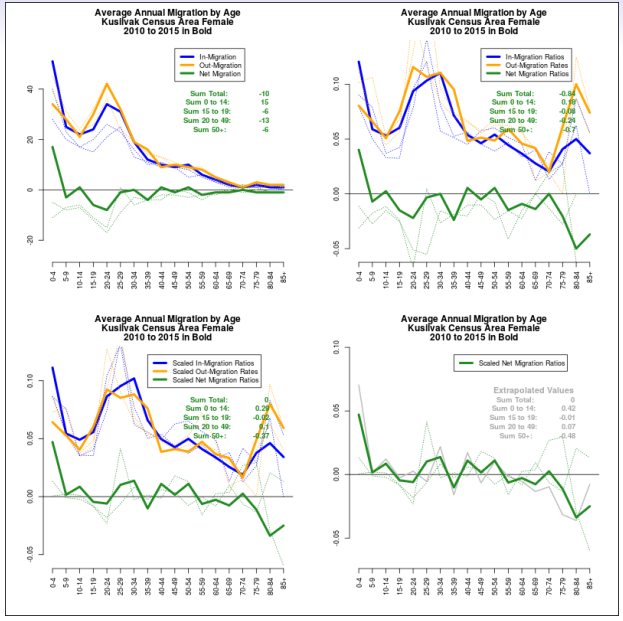
The graphs in the lower panels are adjusted for the levels - just like the first step of Rogers-Castro is to scale. And I think this gives a better idea of how the profile, and particularly the net migration profile (which we're often most interested in for population estimates and projections) is changing over time. There is discussion of losses (or gains) of seniors, and also of "brain drain," and there was also some question about net losses of kids, but we see here - including age as well as level adjustments -that things seem fairly consistent, but there's been some shift rightward over the three periods we've looked at. Highlighting this: In grey on the right are extrapolated values (these are simply from average change between the previous three periods).

Of course the overall zero-net-migration level used here is just to control things for comparisons, and there of course can be consistent positive or negative net migration over time. The net migration profile can be shifted up and down with weights from the combined in- and out-migration profiles to control things - we see this in many published stochastic population projection works, and I do this with employment-based and stochastic population projections exercises for Alaska that are available on the web and I can talk more about in follow-up.



From shiny.demog.berkeley.edu/eddieh/AKPFDMigrationReview

This is just for Anchorage males, and we see its approximately consistent with the statewide shapes and levels. And controlled for level...Here still we see some rightward shift, like statewide, that I think is interesting. I think the index values are useful for assessment, and the shapes - perhaps using numbers or crude rates rather than age-specific rates - are stable enough that if we project population out over the long term - with them shifted up or down to the appropriate levels as needed - we get a stable age structure that I feel would be useful for assessment.



From shiny.demog.berkeley.edu/eddieh/AKPFDMigrationReview

For Kusilvak females though - a much smaller population - even with the large coverage of PFD, residual net migration is very shaky, unstable - some of this may be from our data development choices (such as imputation of 0-year-old migration) and could be improved upon, but still it seems shaky. Notably, for projections we use crude migration rates, and the profiles are used just to share migration out - we don't use age specific migration rates or migrant pooling, as these seem to give wonky results.

On the bottom is Kusilvak controlled for level. One more notable item on this: There's been concern that women are leaving western and northern communities at a higher rate than men, but looking at the data just in this sort of accounting way, it doesn't seem overwhelming. Still, it may be useful to investigate more.

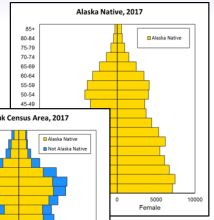
Thank You

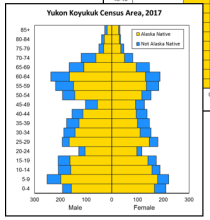
Contact information
Email: eddie.hunsinger@alaska.gov
Phone: (00-1) 907-269-4960

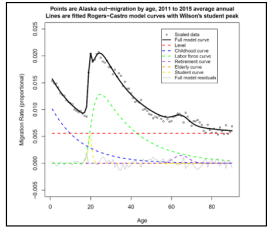
Alaska population and migration data on the web
Population estimates: live.laborstats.alaska.gov/pop
Population projections: live.laborstats.alaska.gov/pop/projections.cfm
Migration data: live.laborstats.alaska.gov/pop/migration.cfm

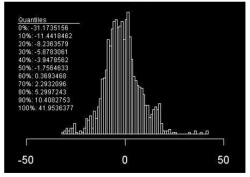
Alaska migration data viewers used today:
shiny.demog.berkeley.edu/eddieh/AKPFDMigrationReview
shiny.demog.berkeley.edu/eddieh/AKMigrationAgePFDtoACSCompare
shiny.demog.berkeley.edu/eddieh/AKMigrationAgePFDtoACSReview

Thank you, and thanks especially to Eric Sandberg and also David Howell for their development of the PFD-based migration data, and also to Liz Brooks and all my colleagues at AKDOL Research and Analysis.









Data and images from laborstats.alaska.gov, github.com, and demog.berkeley.edu/ eddieh

Some additional topics I'm interested to discuss while I'm here: Yukon-Koyukuk Census Area's distinctive age structure and the history behind it; Rogers-Castro fitting and use; empirical prediction errors database; integration of the Applied Demography Toolbox with Github.