Whose Income is Hump Shaped?

Jonathan Page, PhD
University of Hawaii at Manoa
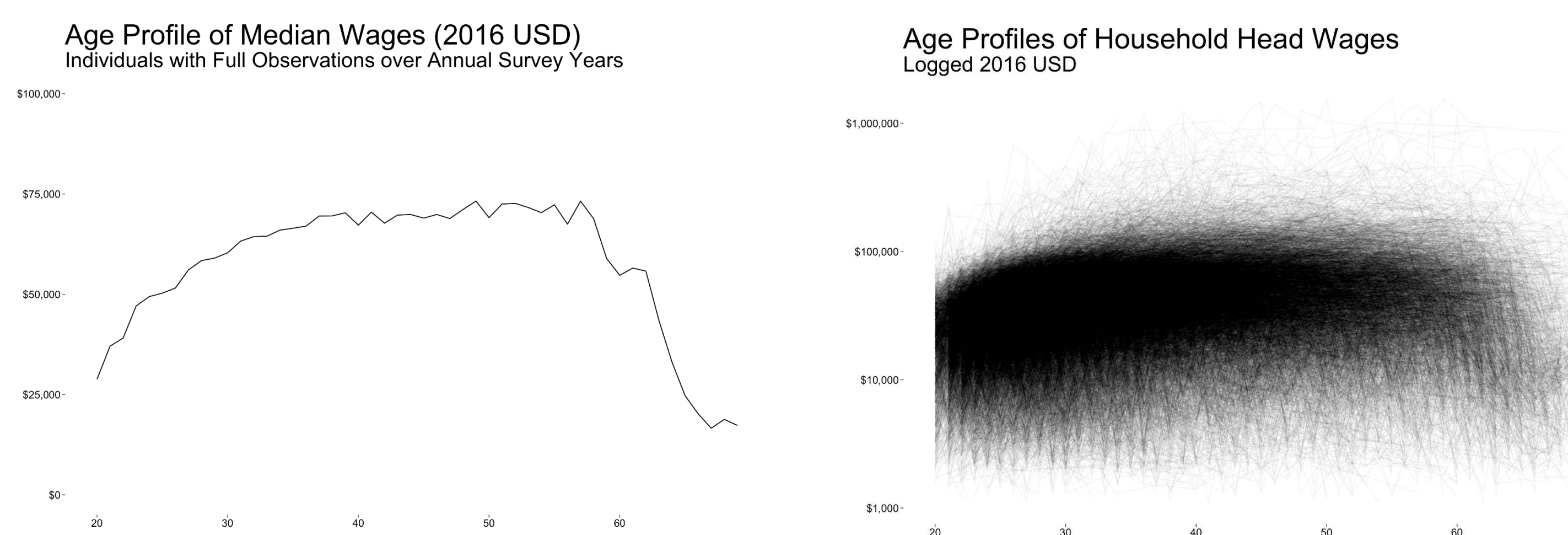University of Hawaii Economic Research Organization (UHERO)

Typical Profile vs. Profile of Typical

The aggregate age profile of wages is hump shaped, but this age profile does not belong to any individual. Each year is treated independently and the resulting profile is constructed from the typical observation over all individuals one age at a time. This research project presents methods for identifying the typical age profile of wages, where previous research and theory has largely focused on the age profile of typical wages. I use the medoid to identify a typical member of a set. The key distinction to keep in mind is as follows:

Typical age-profile of wages
vs.
Age-profile of **typical** wages

Age Profile of Median Wages (2016 USD)
Individuals with Full Observations over Annual Survey Years

Age Profiles of Household Head Wages
Logged 2016 USD

Age Profiles of Individual Wages
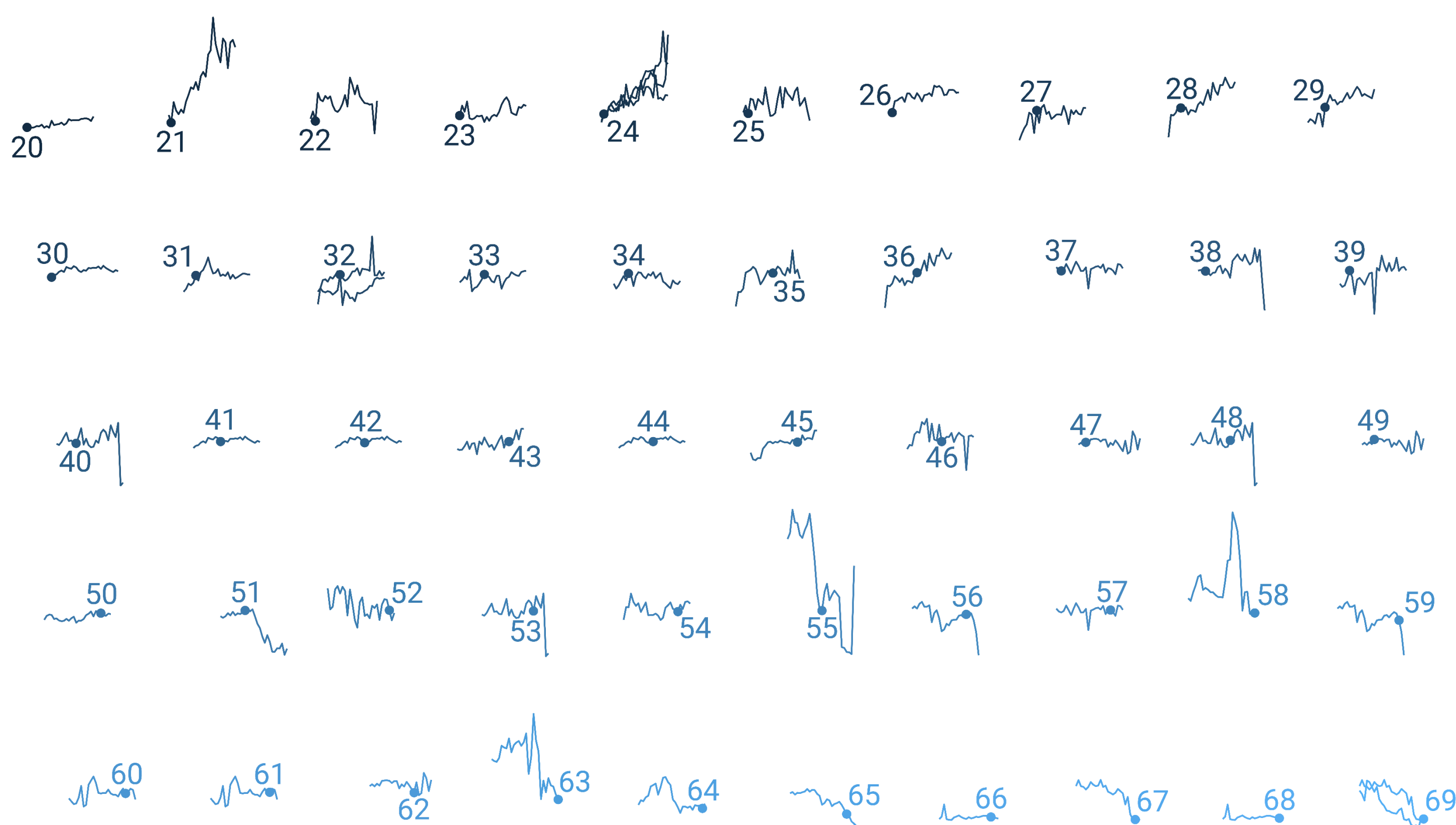Individuals with the Median Wage for a Given Age

Figure 1. The hump-shape, deconstructed

Typical Profile (Medoid)

The age profile of average wages is usually drawn as a single line, but each point may live on a different individual's age profile (see figure 1 above).

The medoid is the observation that is closest on average to all others in its set. Since this definition depends only on pairwise distances, we can identify the medoid for a set in any space with a distance defined.

I focus on dynamic time warping (DTW) distance to illustrate the procedure and point to a larger body of work[1, 2]. Profiles are also referred to in the time-series clustering literature as trajectories.

PSID Household Head Wages

The data used for this analysis is the Panel Study of Income Dynamics (PSID). I focus on head of household wages. This measure allows me to present the core methods of identifying a medoid profile and of time-series clustering while sidestepping issues of equivalence in adjusting family income.

Beginning with the 1970 survey (covering wages earned in 1969), PSID participants provided the wage of the household head. Following the 1996 survey (1995 wages), surveys occur biennially. For simplicity, the rest of this analysis is restricted to even age responses. The latest year of available data is the 2015 survey (2014 wages). I limit my sample to household heads whose wages were above $1,000, less than $1,000,000, and neither imputed nor top-coded. I use the CPI-U[3] to adjust wages for inflation.

Typical Age-Profiles of Wages

Dynamic Time Warping allows for the time dimension to be warped allowing for flexible alignment of spikes/troughs occurring on neighboring ages. I calculate the medoids in this distance space for even age observations for three cohorts: ages 20-48, 30-58, and 40-68 (see figure 2).

Typical Age Profiles of Wages in DTW-Space
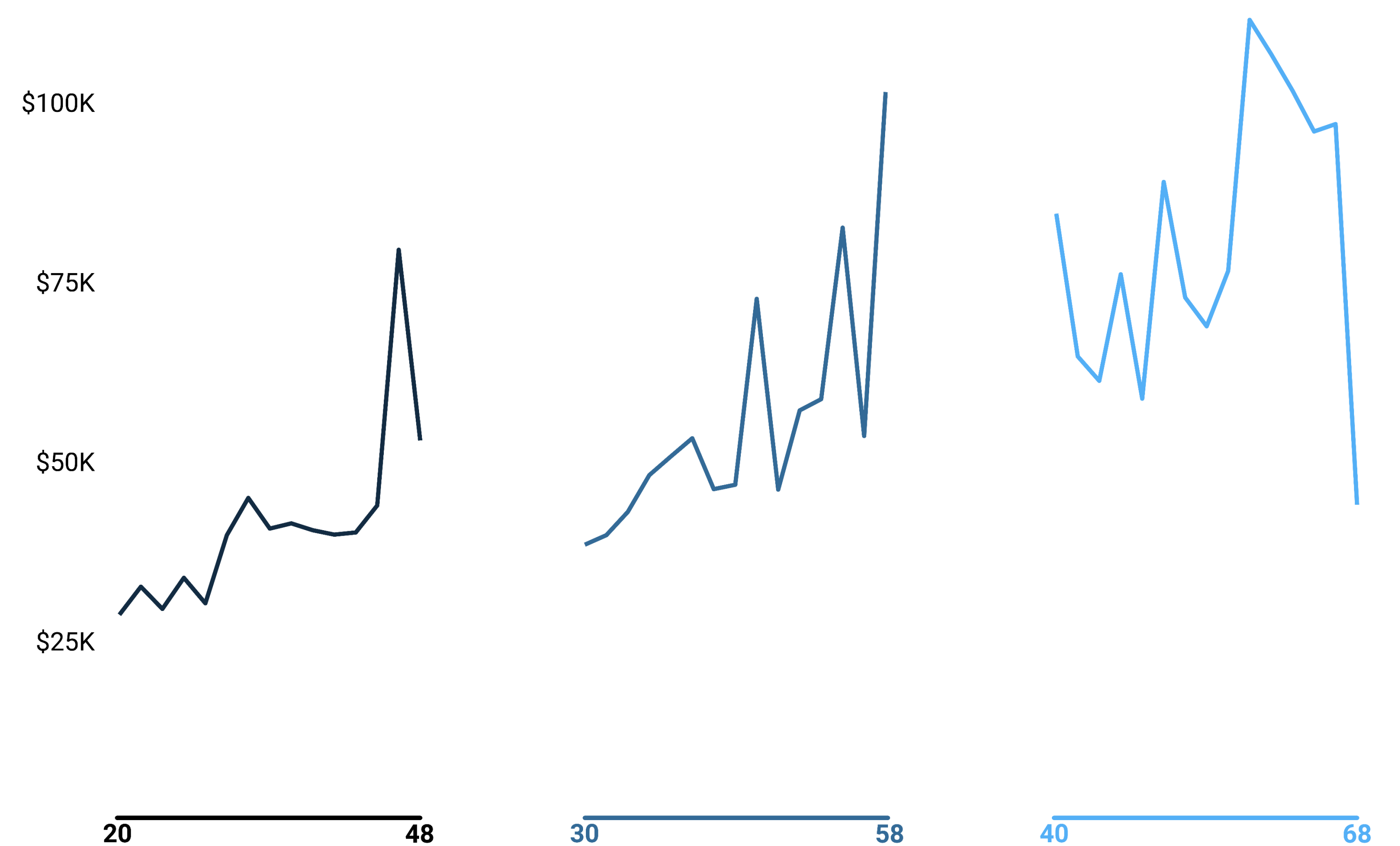Biennial Household Head Wages (2016 USD)

Figure 2. Med profiles, three ways

Profile Resolution Through Clusters

The time-series clustering literature provides a rich toolkit for increasing the resolution in the summary of a bag of profiles. Here I split each cohort of profiles into four clusters using partitioning around medoids, an iterative procedure of assigning individuals to their nearest medoid and calculating group medoids for the next round (see figure 3).

Typical Age–Profile of Wages within Clusters
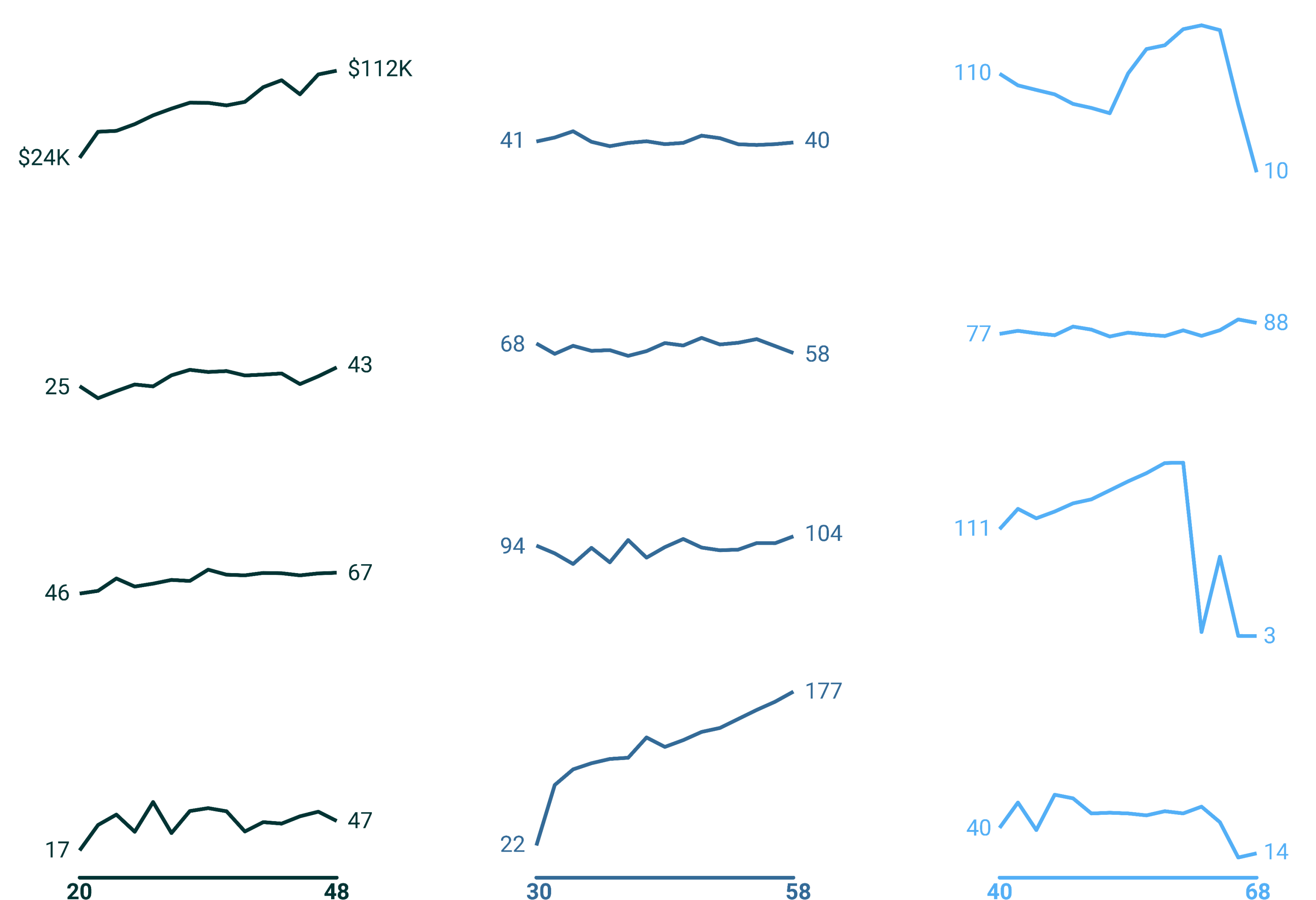Clustering in DTW-Space (Thousands of 2016 USD)

Figure 3. Clustering summary

Next Steps

Using publicly available survey data has its limitations. Administrative data will provide a consistent cross-individual notion of income and would reduce non-response attrition in the data. To avoid privacy issues surrounding identifying a medoid individual, multiple representatives will be identified as a medoid group through a bootstrap sampling procedure.

Contact

Jonathan Page
University of Hawaii Economic Research Organization
Email: jrpage@hawaii.edu
Phone: (808) 956-7065

References

1. Aghabozorgi, S.; Shirkhorshidi, A. S. & Wah, T. Y. (2015) Time-series clustering - A decade review. Information Systems, 53, 16-38.
2. Fu, T.-C. (2011). A review on time series data mining. Engineering Applications of Artificial Intelligence, 24(1):164-181.
3. U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCSL, November 24, 2017