Retention Trends Within Arkansas' WBL-to-Workforce Pipeline

Authors
Nagaraj Bettadapura, ADWS/TANF, Arkansas
Sam Keathley, KYSTATS, Kentucky
Eric Schroer, LAHSA, California
Serpil Tokdemir, Arkansas DMS, Arkansas
Team Lead: Nishav Mainali, ARDATA, Arkansas

Introduction

We believe that prospective employers and employees or apprenticeship enrollees may both benefit from an evaluation on how to get staff to complete apprenticeship programs. But that isn't necessarily our focus. Our focus is to understand the relationship between post apprenticeship retention and employer and employee benefit. We believe these matters because it enables us to assess the value that both employers and employees derive from the completion of work-based learning programs. And it gives us the ability to compare those who stay in their industry after their apprenticeship and those who leave their industry.

For our research methods, we had a literature review looking at both employer and employee benefits. While Employers benefit from better understanding if the employees they invest in will stay, our question and literature review focus on post-apprenticeship completion.

Existing literature demonstrated a focus on retention and attrition among apprenticeship participants, prior to or regardless of apprenticeship program completers. A lot of the research on attrition and retention was by Biligisoy. In 2003 they looked into attrition and retention in the construction agency. They used what used the RAPIDS (previously called AIMS) data to understand the difference between programs funded jointly (employer and union) and unilaterally (employer only). They used the Cox Regression model (some results shown on the next slide) and found that retention rates were higher for apprentices in joint programs. They found lower rates of completion for women than men and for racial and ethnic minorities than for whites. They also found that apprenticeship duration rose with the unemployment rate.

In 2018, they looked at the great recession's impact on retention and attrition. They found some differences in attrition with the great recession (compared with other downturns). The quit hazard increased for trainees during the great recession and decreased for unilateral trainees. The eventual probability of cancellation was lower among the union-employer joint programmes than in the employer-only programmes, although the difference narrowed down during the Great Recession.

In terms of the employer benefits, we did see specific to a position apprenticeship program in New England that there were sort of Ancillary retention benefits where it improved workload, and there was some Anecdotal data where they believe that increased retention. But that program also demonstrated statistically significant performance benefits where this preventative health clinic was able to administer a higher volume of services than in prior years when they didn't have the apprenticeship program completers. We also noted that for a CVS program for pharmacy technicians and store managers in Michigan, South Carolina, and Georgia, they also stated that the program helped reduce the turnover rate by roughly half.

In terms of the literature review on the apprentice or on the employee side benefits, we found a research study from Mathematica that lists the variety of long-term and short-term earnings benefits and their effectiveness. It is stated that, specifically in the 9th year following their apprenticeship program enrollment, those participants earned an average of nearly $6,000 more than similar non-participants. And this is considerably durable given that it's nine years out. We also found a separate study focusing on how in roughly nine to twelve quarters after completion, apprenticeship completers had roughly $3,500 additional earnings per quarter, and that same cohort had roughly 10% higher employment rate. One other benefit we wanted to mention was that the literature review found that women who participated in apprenticeship programs from Qualitative and Anecdotal data really valued them. They did suggest

that there could be additional support for more women in these programs, such as childcare and assistance related to harassment prevention.

Research Question

Our specific research question is what are the retention trends within Arkansas work-based learning workforce pipeline exclusively for those who complete work-based learning apprenticeships? Additional related questions we explored include:
- How does employer level retention compare to industry retention,
- How does retention relate to wages?
- How does retention differ by demographic groups, particularly geographic groups?

Data/Model

The literature review described in the introduction section above provided much of the context and inputs to develop the framework for the research question and the structure of the project. In addition to this, the team worked closely with the Coleridge and the Arkansas data administration team(s) to learn the out-of-box available database schema, tables, the description of the various columns and values constituting the tables. The initially available database table structure presented some challenges with establishing linkages across domains, as well as transforming the data to perform exploratory analysis. Subsequent iterations were made available to the database administration team that was not only more longitudinal in nature but also much more suitable for data analytics. With these base tables in place, the team had an opportunity to finalize additional external / 3rd party data that needed to be specifically asked for to be imported and made available for the data analysis.

The entry point for most of the project work began with deep-dive analysis of the Registered Apprenticeship Sponsor Information Database (RAPIDS) client level database table. This table provided us with the data elements that helped formulate the cohort to be studied for this project.

The RAPIDS database provided us information of the apprenticeship outcomes like completers and non-completers, their demographics, and the various apprenticeship program providers..

The Unemployment Insurance Quarterly Wage Data and the Quarterly Census of Employment and WAGES (including NAICS Industry and Employer Data) was another very critical datasets for the project. Data from these tables were used especially to explore, analyze and visualize the outcomes of the cohort. Reading from these tables was a two step process. These tables provided information on wages, employer, and industry for the cohort after their apprenticeship exits. It provided us information about not only industry and occupational information, but also wage progression over the years.

As the project question was formulated, the team decided to include additional data sources to complete the analysis for the project. This required the additional requests to bring in the Inflation Adjustment Data as well as the Labor Market Information on High Demand Occupations (Arkansas Employment Projections). These data sets helped the project to explore the wages in the broader context of inflationary impact on wages but also understand the apprenticeship related occupations vis-a-vis LMI and in-demand and projected occupations. These data points could be used to present opportunities for the apprenticeship program administrators to compare existing programs vis-a-vis LMI trends in the state.

All client and outcome datasets were appropriately filtered for the State of Arkansas.

Framing of the research question, followed by exploratory analysis of the data available and the various characteristics of the available datasets, the cohort was narrowed down to study apprenticeship participants who completed between 2012 through 2018.

Since the question of retention required analyzing the quarters of employment after completion, the timeline was designed to give us at least 1 full year pre-Covid, i.e., 2018 completers and exits could be analyzed for the 2019 year.

Additionally, considering the motivation for this research project, data sets were selected to provide information about geography, demography, inflation adjustment and type of employment during and post-apprenticeship, as well as wages earned.

Project Goal(s) and Quantification of Outcomes

The project focuses on analyzing the apprenticeship completer characteristics (biographic and demographics) and their outcomes with regards to employer retention, industry and / or occupations. Success of an apprenticeship program benefits many stakeholders either directly or indirectly. Potential stakeholders who could benefit from this project include local employers, education and training providers, jobseekers (first-time or those seeking career changes and/or growth), workforce program administrators and WBL/apprenticeship program designers. The project could also contribute to a better understanding of the value proposition derived from WBL programs. Analyzing the trends of retention and complementing that with wage trends could also provide valuable data points on what could be the elements of a successful apprenticeship program.

Analysis Methods and Procedures

The project adopted many of the standard best practices required for a typical data analytical project, including but not limited to:

Project scoping and the applicability of privacy and confidentiality. The project team members came in with varied backgrounds and skill sets that spanned technical and technology skills but also with a business and programmatic background. The common thread though amongst the team members was an agile mindset. This helped us iterate and evolve not just the framing of the research question but was also extremely helpful in terms of discovering, studying-up and sharing prior literature reviews, and which of those could provide us a frame of reference to move forward on the current project.

The actual conducting of the data analysis roughly followed the course work (weekly lectures, course pre-work, sample notebooks) that was laid out over the many weeks. This combined with the team members' own experiences with working on technical and business projects were leveraged through all phases of the project execution.

Typical methods of exploratory data analysis included understanding the schema, tables that were made available, the uniqueness, constraints, quality of the data itself, and inference methods were employed.

Understanding the ADRF usage terms and conditions, tools like jupyter notebook, R Studio and Open Libre were some of the tools that were widely used to conduct our data analysis.

It was also extremely valuable listening in to the weekly lectures to understand before performing the required data linkages within and across (sub)-domains followed by conducting cross-sectional analysis. Each of these activities were regularly discussed and validated within the team, as well as receive clarification from the team lead and / or during the weekly coursework.

The team also briefly experimented with feature engineering to identify the most valuable data elements that could be considered to define the cohort and potentially help frame the research question itself. This was more experimental and rudimentary but was helpful in finalizing how to view the input data from the RAPIDS table.

In addition to the default data sets, and the database administrator having to work through the ETL processes to bring in additional datasets, the team used basic techniques using SQLs, spreadsheet, python's statistical packages (numpy, pandas, scipy) and R to conduct its own transformations to study the data.

Part of this processing involved transforming available countywide demographics into a generalized delta/urban/semi-urban recategorization for better presentation and contextualization of analysis

Measures developed specific to his project

Of all the various dimensions analyzed, the following were considered most relevant to the project presentation, and therefore visualized / highlighted for public consumption:

- Trends of apprenticeship completers who chose to remain with their apprenticeship employer vis-à-vis those who changed employers
- Trends of completers who remained in the same industry versus whose who changed industries
- Industries that were relatively more successful in retaining completers
- Median weekly wage trends of completers who remained with their apprenticeship employer versus those who changed employers
- Longer term (2 year plus) analysis of mean quarterly wages after completion
- Geographical comparisons between urban and semi-urban/rural and delta vs non-delta on the ability of apprenticeship employers' ability to retain apprenticeship completers over 4 years
- Contrasting apprenticeship retention across workforce areas, and the average median wages difference (if any) between retained employees vs non-retained employees

Key Findings

Listed below are the key findings from the project:

- 3 years out from the completion of apprenticeship program completers are likely to stay employed with their apprenticeship employers
- Completers earned $100 more per week than non-completers
- Specialty trade companies retained close to 70% of their apprentices two years after completion
- Employers in the rural parts of the Arkansas Delta region are more likely to retain apprenticeship completers

Implications of the findings and potential inputs to the State's policy makers

The motivation of this project was not only to learn the tools, techniques, methods and general awareness of the value of data and analytical methods but also where possible explore the transferability of data analysis to business decisions and / or contribute to sound policy making. Below is a list of potential considerations for the State Of Arkansas' various stakeholders:
- Given the higher retention rates of apprenticeship completers with the employer, program expansion could be a beneficial strategy for existing and / or new employers adopting the WBL model
- In addition to increased job opportunities and job security, apprenticeships have demonstrated to increase wages for completers. This could be an attractive data point to attract more Arkansans to join apprenticeship/WBL programs
- The success of employer retention in the Arkansas Delta region could indicate that expanding apprenticeship programs in this region could improve the social and economic well-being of residents in this region without having to consider relocation

To further elaborate the findings and the rationale for these inputs to policy makers, we have listed a set of charts in the Appendix.

Caveats

As with any project comes limitations, constraints and / or assumptions. The current project is no different, and the caveats are accordingly listed below:

- Data distribution across certain demographic dimensions was limited, example there was a significantly higher percentage of white and male apprenticeship participants which limited the ability of the research to review data from a racial or gender perspective for example
- Analyzing non-completer reasons or non-retained employees or the lack of qualitative data from employers could provide valuable insights and a more contextualized understanding of retention and completion characteristics
- Limited record linkage with regards to the lack of onet code in the WAGE data makes it difficult to determine if the apprentice continued in the same occupation and / or understanding if there was further progression or regression down the line.

Possible Extensions and Opportunities for further research

Given the scope, the constraints and the findings from this study, there is an opportunity for further areas of study. The project team believes that incorporating research from the below list could add additional value for apprenticeship program administrators and policy makers:

- Factor cost-of-living / normalization of outcome data - with rising inflation, and the concept of a living wage (that varies from region to region) could be more relevant to realizing the true effect of the **real value of** wages received post completion.
- Supplement with qualitative data from employers - this information is currently not captured in any of the data sets. However, understanding data characteristics of apprentices and the employers' themselves could provide more value with understanding the pieces that go into building and operating a successful apprenticeship program.
- Further analysis of Completer retention vs Program retention - the study encountered many non-completers during the course of designing the cohort. Understanding the reason(s) why the program itself could not see an apprentice through completion, or for that matter an apprentice actually moving on to a different job prior to completion could offer valuable insights and a more comprehensive picture of additional retention dimensions.

Appendix

Apprenticeship Completers that Remain With their Apprenticeship Employer vs. Completers who Change Employers
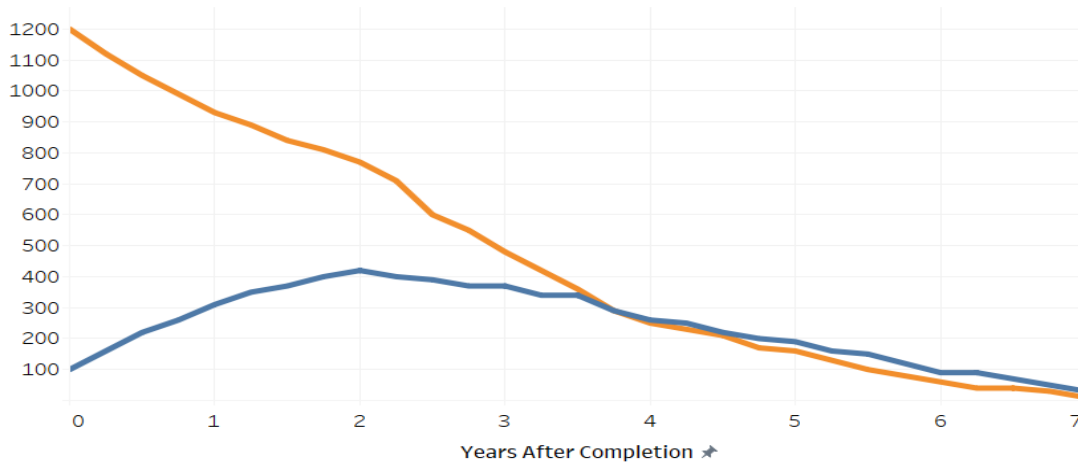

Chart 1:Trends with employer retention over years. There appears to be an inflection point at around the 4th year

Non-Retained Apprenticeship Completers that Remain in the Same Industry vs. Completers who Change Industries
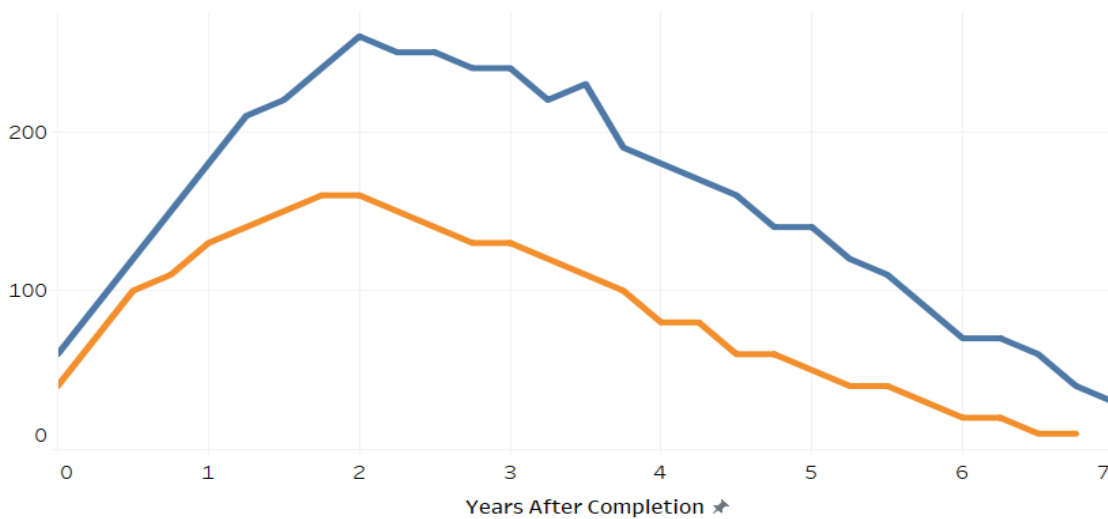

Chart 2: Industry retention over the same 7 year period shows the industry retention diversion amongst the completers. Around year 2 appears to be a peak with the change-in-industry trend, though overall there appears (overall) more completers changing industry versus staying within. A potential follow-up investigation that combines the findings from this chart with the employer retention trends could be illustrative

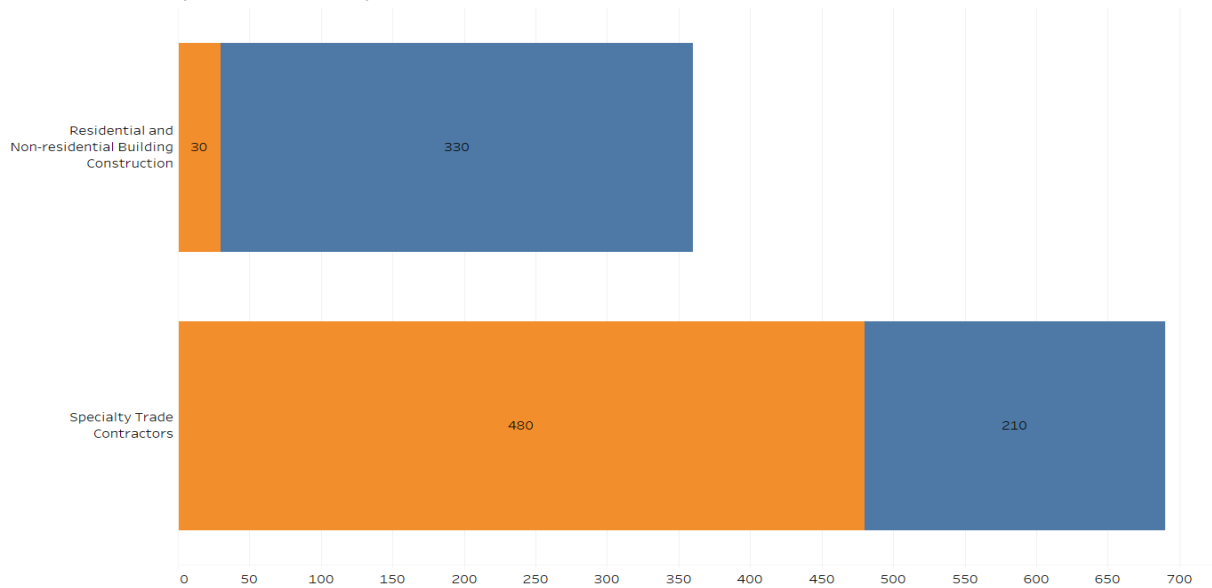**Which Industries (Three-digit NAICS) Retain Their Apprenticeship Completers?**

Residential and Non-residential Building Construction	30 / 330
Specialty Trade Contractors	480 / 210

Chart 3: Speciality Trade Contractors have a higher success rate with retention

Median Weekly Wages of Completers that Remain With their Apprenticeship Employer vs. Completers who Change Employers
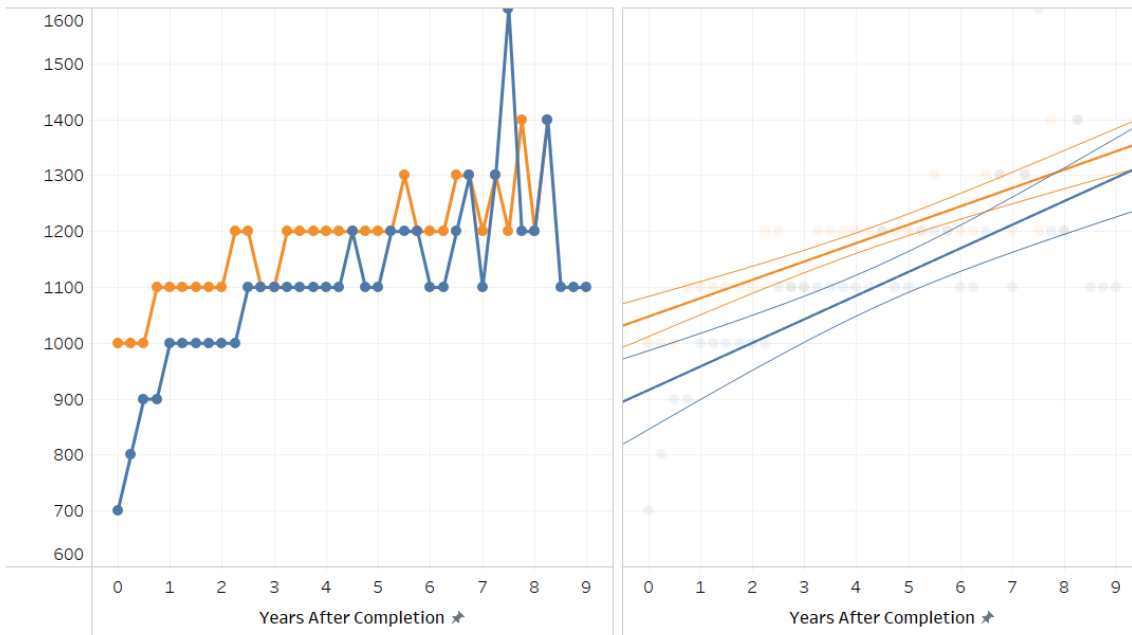
Chart 4: Median wages appear to be higher among employer retention candidates overall (but for a spike in the far out years of 7 and around 9
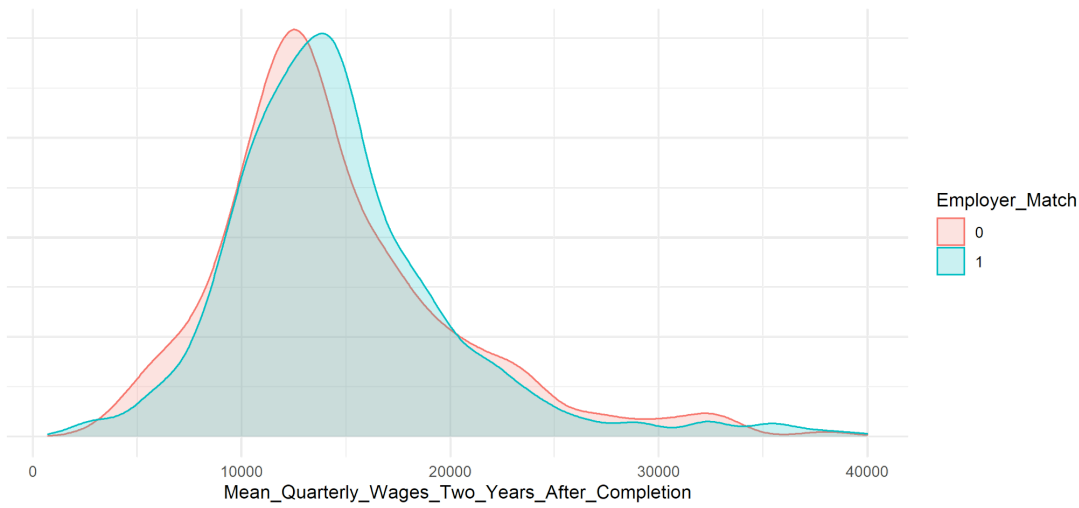
Chart 5: Distribution of generally a higher wage employment among employer retained apprenticeship completers
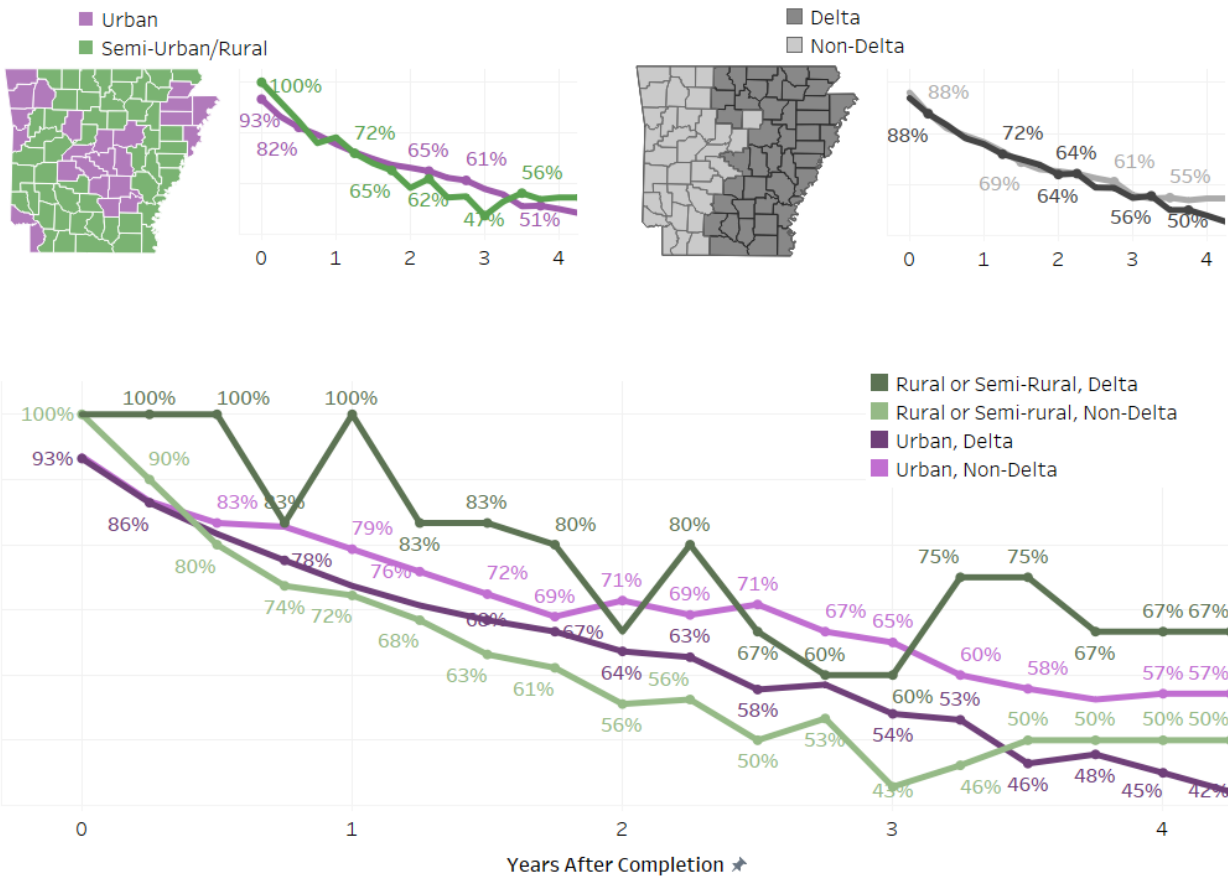

Chart 6: Rural / Semi-Rural Delta shows the highest retention characteristic in the short, medium and long term