Analytical Demonstration

Understanding and Strengthening Postsecondary STEM Pathways from Enrollment to Employment: Evidence from Arkansas

Task 3

STEM Attrition Report: Finalize Technical Report documenting analytical approaches and findings (Deliverable 8)
STEM Subgroup Report: Finalize Technical Report documenting analytical approaches & findings (Deliverable 10)

for

Coleridge

Project No. ADC-STEM-24-N8

Ahu Yildirmaz, Ph.D.

20 F Street NW, 7th Floor

Washington, District of Columbia, 20001 US

ahu.yildirmaz@coleridge.us

First Submission: 06/19/2026

Final Submission: 07/27/2026

Table of Contents

Project Team **3**
Data Partners & Contributing Agencies **3**
Funding Acknowledgment **3**
Introduction **4**
Research Questions **6**
Postsecondary STEM Pathway 6
Timing and Determinants of STEM Persistence and Attrition 6
Workforce Outcomes 7
Data and Methods **7**
Data Sources 7
Data Linkage and Coverage 8
Analytical Sample and Cohort Construction 8
Definition of Key Measures 10
Findings and Discussion **12**
Descriptive Overview of Postsecondary STEM Pathways 12
Enrollment Patterns and Institutional Context 14
Observation Windows, Exit Conditions, and Censoring 15
Timing of STEM Persistence and Attrition 16
Variation in Attrition by Degree and Cohort 16
Hazard Model Results 18
Labor Market Outcomes by Postsecondary STEM Pathway 20
Subgroup and Contextual Analysis of STEM Pathways 23
Reproducibility & Future Work **26**
References **28**
Appendices **31**

Project Team

The following individuals contributed to the design, implementation, analysis, and reporting of the project at various stages. Roles and dates reflect the period during which each team member was actively involved.

Coleridge

- Justin Baer, Ph.D., Technical Point of Contact and Project Lead
  January 16, 2026 – May 22, 2026
- Nate Barrett, Ph.D., Technical Point of Contact and Project Lead
  August 21, 2024 – January 15, 2026
- Natsuko Nicholls, Ph.D., Project Manager and Research Coordinator
  November 25, 2024 – April 30, 2026
- Chip O'Connell, Ph.D., Data Analyst
  January 26, 2026 – August 1, 2026
- Allison Nunez, Ph.D., Data Analyst
  August 21, 2024 – December 15, 2025

State of Arkansas

- Robert McGough, State Project Partnership Lead
  August 21, 2025 – August 1, 2026
- Jake Walker, Ph.D., Data Steward & Technical Data Advisor
  August 21, 2025 – August 1, 2026
- Tina Moore, E.D., Policy Advisor
  August 21, 2025 – August 1, 2026

Data Partners & Contributing Agencies

- Arkansas Office of State Technology, Department of Shared Administrative Services
- Arkansas Division of Higher Education, Arkansas Department of Education
- Arkansas Workforce Connections, Arkansas Department of Commerce

Funding Acknowledgment

This work was supported by funding provided under the National Secure Data Service (NSDS) Demonstration project under management by the National Center for Science and Engineering Statistics (NCSES) within the U.S. National Science Foundation (NSF) through the America's Data Hub Consortium (ADC) program, managed by Advanced Technology International (ATI), under ADC Base Agreement No. 2022-373 and Project Agreement ADC-STEM-24-N8: *Data Integration to Estimate Science, Technology, Engineering and Mathematics (STEM) Attrition and Workforce Supply: A Pilot Approach.* Coleridge served as the Project Agreement Holder for this effort.

Introduction

The STEM workforce is central to the nation's capacity for scientific advancement, economic competitiveness, and technological innovation (Carnevale, Smith, & Melton, 2011; Langdon et al., 2011; National Science Foundation, 2024). National labor market projections indicating a growth rate in STEM employment double the growth rate in general employment by 2034 underscore the continuing need to strengthen pathways that attract, prepare, and retain individuals in STEM fields (National Science Foundation, 2026). Yet persistent challenges within the education-to-workforce pipeline constrain the nation's ability to meet projected demand (Agarwal, 2026). Attrition from STEM majors during postsecondary education remains substantial, with nearly half of students who initially declare a STEM field leaving the major by their fourth year (Chen, 2013; Riegle-Crumb, King, & Irizarry, 2019; Speer, 2023). Even among STEM degree completers, only a minority ultimately enter STEM occupations (Buffington et al., 2016; Jelks & Crain, 2020; Kim, Jung, & Mlambo, 2021). These patterns raise critical questions about where attrition occurs, which students are most affected, and under what conditions persistence and successful transitions are most likely.

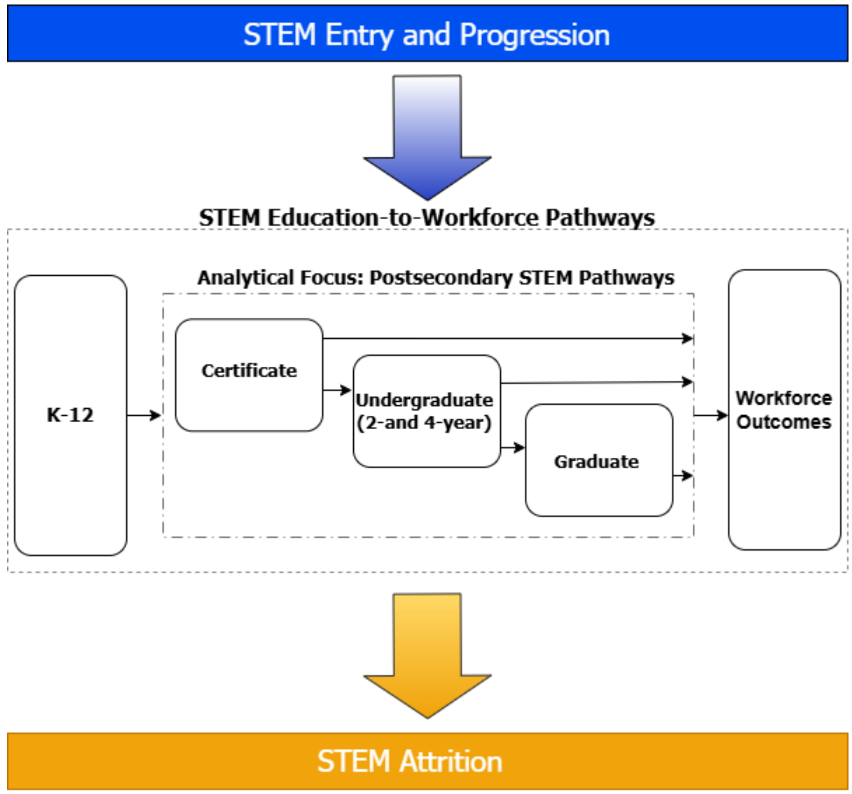
This report presents a comprehensive analysis of postsecondary STEM pathways in Arkansas, drawing on the state's integrated administrative data infrastructure and the Coleridge Initiative's Administrative Data Research Facility (ADRF). Building on earlier phases of this work, the analysis incorporates methodological refinements that enhance analytic precision, including improved cohort construction, an individual–institution–term longitudinal data structure, and the estimation of discrete-time hazard models with logistic specifications. These enhancements allow for a more accurate representation of student trajectories across institutions, degree levels, and time, and enable explicit examination of when students are most likely to persist in or exit STEM pathways.

Conceptually, the study is grounded in complementary perspectives widely used in the STEM pipeline literature. One perspective emphasizes attraction and attrition, focusing on the conditions under which individuals enter STEM pathways and the points at which they exit. The other centers on retention and persistence, examining continued enrollment, progression toward degree attainment, and subsequent transitions into STEM-related employment. These perspectives motivate an analytical approach that treats STEM participation as a dynamic, longitudinal process rather than a static outcome (Engberg & Wolniak, 2013; Stevens, Armstrong, & Arum, 2008; Xie, Fang, & Shauman, 2015). At a broader level, the STEM pipeline can be understood as a multi-stage ecosystem spanning academic preparation, postsecondary education, and workforce entry, with multiple opportunities for persistence, redirection, or exit along the way. Prior research has documented broad patterns within this ecosystem, but many studies focus on isolated transitions, rely on cross-sectional comparisons, or lack the longitudinal resolution needed to examine when and how attrition occurs (Chen, 2013; Green & Sanderson, 2018; Lee & Ferrare, 2019).[1]

---

[1] The complete literature review was compiled and submitted to NSF as part of Task 2 and Deliverable 4 on March 17, 2025.

Figure 1. Analytical Framework for the Arkansas STEM Pathways Study



As shown in Figure 1, this study narrows that broader framework to a specific and policy-relevant segment of the STEM ecosystem: postsecondary persistence and degree attainment among students who begin their postsecondary careers in STEM fields across certificate, undergraduate, and graduate pathways, and their subsequent transition into the workforce.[2] By focusing on this segment, the study moves beyond static outcome measures and instead examines STEM pathways as processes unfolding over time. This framing makes it possible to ask not only who ultimately completes but also when attrition is most likely to occur, how persistence differs across institutional contexts and student groups, and how postsecondary STEM pathways relate to workforce outcomes. The narrow framing does have the drawback of excluding the direct pathway from K-12 to STEM workforce, but it allows for a full focus on postsecondary pathways with K-12 indicators as a set of pathway determinants.

Arkansas' statewide longitudinal administrative data system is particularly well suited to this approach. This study builds on Arkansas' broader investments in linked administrative data infrastructure, including prior NSF-supported work by ARData on record linkage fidelity, transparency, and linkage bias in statewide administrative data (Gudipati et al., 2025).The availability of term-level postsecondary enrollment records, degree information, linked workforce

[2] Hereafter, for simplicity, the term "degree attainment" is used broadly to include both certificate/credential attainment and degree attainment, unless otherwise noted.

outcomes, and selected K-12 indicators makes it possible to follow individuals across institutions and over time, observe gaps and re-enrollment behavior, and model persistence and attrition as time-varying processes rather than single endpoints. By integrating statewide longitudinal administrative data with refined modeling strategies, this analysis offers an empirically grounded account of STEM pathway dynamics in Arkansas. The findings provide evidence relevant to policymakers, educators, and workforce leaders seeking to strengthen STEM talent development and address inequities across the pipeline.

Research Questions

This study examines postsecondary STEM pathway dynamics in Arkansas by focusing on three related domains: the structure of postsecondary STEM pathways, the timing of persistence and attrition, and workforce outcomes associated with different educational trajectories. The questions build on the project's original analytical plan while incorporating adaptations enabled by Arkansas' administrative data infrastructure, including the availability of selected K-12 indicators of academic preparation and performance.

Postsecondary STEM Pathway

1. **What pathways do students follow after entering postsecondary education in a STEM field?**

   This question examines the overall structure of postsecondary STEM pathways and the distribution of educational outcomes among students who begin their postsecondary careers in STEM. Specifically, the analysis addresses:

   1.1. What proportion of students who begin postsecondary education in a STEM field complete a STEM degree, complete a non-STEM degree, or exit postsecondary education without earning a degree?

   1.2. How do these pathway outcomes vary across institutional contexts, including students who begin at two-year versus four-year institutions?

Timing and Determinants of STEM Persistence and Attrition

2. **When do students exit STEM pathways during postsecondary education?**

   This question focuses on the timing and determinants of persistence and attrition from STEM fields during postsecondary education, including the role of pre-college preparation, institutional context, and variation across student groups. Specifically, the analysis asks:

   2.1. At what points during the postsecondary trajectory are students most likely to exit STEM pathways?

2.2. How do indicators of academic preparation prior to postsecondary entry, including measures derived from K-12 assessment data, relate to subsequent persistence and attrition in STEM fields?

2.3. How do patterns of persistence and attrition vary across student subgroups, including gender, race and ethnicity, age, geographic origin, and institutional entry point?

Workforce Outcomes

3. How do postsecondary STEM pathways translate into STEM workforce outcomes?

The final question examines how variation in postsecondary educational trajectories translates into differences in workforce outcomes. In particular, the analysis addresses:

3.1. How do short-, mid-, and long-term labor market outcomes differ across STEM starters who complete STEM degrees, complete non-STEM degrees, or exit postsecondary education without a degree?

3.2. To what extent are STEM degree completers more likely to enter STEM-related industries, as identified using employer North American Industry Classification System (NAICS) industry classifications?

Data and Methods

Data Sources

This study uses longitudinal administrative data from the Arkansas statewide longitudinal data system that houses data from multiple state agencies. Data from the Arkansas Department of Education, Division of Higher Education (ADHE) were used to examine postsecondary STEM persistence, attrition, and degree attainment, while data from the Arkansas Department of Commerce, Arkansas Workforce Connections (AWC) were used to examine labor market outcomes. Finally, data to construct K-12 indicators comes from the Arkansas Department of Education, Division of Elementary and Secondary Education.

The postsecondary data includes student-level enrollment and degree records from public and private institutions in Arkansas spanning academic years 2011 through 2023. Enrollment records are observed at the term level and include information on institution attended, declared major, enrollment intensity, and degree-seeking status. Degree records provide information on degree level, field of study, and year of completion.

To support consistent classification of STEM fields, the analysis incorporates Arkansas' STEM productivity reference tables, which map instructional program codes (CIP codes) to state-defined STEM categories. Additional reference tables are used to standardize institutional identifiers, term structures, and degree classifications across years. These data allow individuals to be followed across institutions and over time, capturing enrollment continuity, gaps in enrollment, degree attainment, and field-switching behavior.

Where relevant for downstream analyses, postsecondary records are linked to workforce data to assess early labor market outcomes. These linked data support analyses of education-to-workforce transitions but are not the primary focus of the persistence and hazard analyses presented here.

Data Linkage and Coverage

This study draws on a longitudinal, linked administrative data infrastructure that integrates records across multiple domains, including K-12 education, postsecondary enrollment, and workforce outcomes. Data from the State of Arkansas were securely ingested into the Administrative Data Research Facility (ADRF), with Social Security Numbers (SSNs) and other individual-level identifiers transformed using privacy-preserving hashing procedures prior to linkage. These procedures ensure that individual identities are protected while preserving the ability to construct longitudinal records across data systems. Arkansas's cross-sector linkage capacity also builds on prior NCSES-supported work by ARData to assess record linkage fidelity, transparency, and bias in statewide administrative data.[3]

Data linkage was conducted within the secure ADRF environment, enabling individual-level records to be connected across systems while maintaining strict confidentiality protections. The resulting integrated dataset includes:

- K-12 indicators of academic preparation and performance
- Postsecondary enrollment and degree records from the Arkansas Division of Higher Education
- Workforce outcomes derived from Arkansas Department of Commerce Unemployment Insurance (UI) wage records

Linkage across these data sources allows individuals to be followed longitudinally from pre-college preparation through postsecondary education and into the labor market within the state of Arkansas. The analytical sample includes all individuals observed in the Arkansas postsecondary data who meet the study's cohort definition, regardless of whether they can be linked to K-12 or workforce records. The workforce data only allows us to track these individuals into employment within the state of Arkansas, while K-12 data are available for a subset of individuals who attended high school in Arkansas and subsequently enrolled in postsecondary institutions within the state. Analyses incorporating K-12 and workforce indicators are therefore conducted on a linked subset of the full analytical sample.

[3] Prior ARData work conducted under the America's DataHub Consortium focused on record linkage fidelity for the foreign-born science and engineering workforce, while also producing broader methodological guidance relevant to statewide administrative data linkage.

Analytical Sample and Cohort Construction

The base analytical unit is the individual. The analytical sample includes individuals whose postsecondary careers are defined as beginning with their first non-post-baccalaureate enrollment record observed in the ADHE data. Individuals must have at least one enrollment record that is not classified as pre-baccalaureate or continuing education. Students whose observed postsecondary activity consists exclusively of post-baccalaureate enrollment are excluded from the analytic cohort. As a result, individuals observed in master's or doctoral pathways in this study must also have an earlier certificate or undergraduate enrollment record in the Arkansas postsecondary data.

Students are classified as STEM starters or non-STEM starters based on whether they declare a STEM major during any of the first four observed terms in the student enrollment data, using Arkansas's STEM productivity definitions to classify majors as STEM or non-STEM. Individuals who are enrolled in a STEM major in at least one of these first four terms are classified as STEM starters; all others are classified as non-STEM starters. This approach is intended to capture entry into STEM pathways during the initial stage of postsecondary enrollment while allowing for some early movement across majors. This initial classification anchors all subsequent analyses of persistence, field switching, and degree attainment.

To support longitudinal analysis of postsecondary trajectories, a person-institution-term analytical file is constructed for all individuals in the base sample. Enrollment records are ordered temporally, and a single observation per individual per term is retained to establish a consistent time scale for analysis.

The Arkansas higher education data begins in 2012. The analytical sample is further restricted to individuals who began their postsecondary careers in academic year 2018 or earlier to ensure sufficient follow-up time for observing degree attainment or exit behavior. Individuals are followed longitudinally through the academic year of 2023[i], allowing for observation of enrollment continuity, gaps in enrollment, degree attainment, and exit from postsecondary education.

When individuals are enrolled at multiple institutions within the same term, enrollment records are consolidated to a single term-level observation using a consistent prioritization rule. Additional data integrity procedures are applied to ensure temporal consistency within individuals' enrollment histories. Overlapping or duplicate records within institutions are resolved through consolidation or removal, and cases of simultaneous enrollment across institutions are handled using this prioritization approach. These steps produce a coherent person-term structure, ensuring a uniform time scale for analysis while preserving the ability to capture complex enrollment patterns, including multi-institution attendance and transitions over time.

Figure 2. Analytical Cohort Construction and Observation Framework

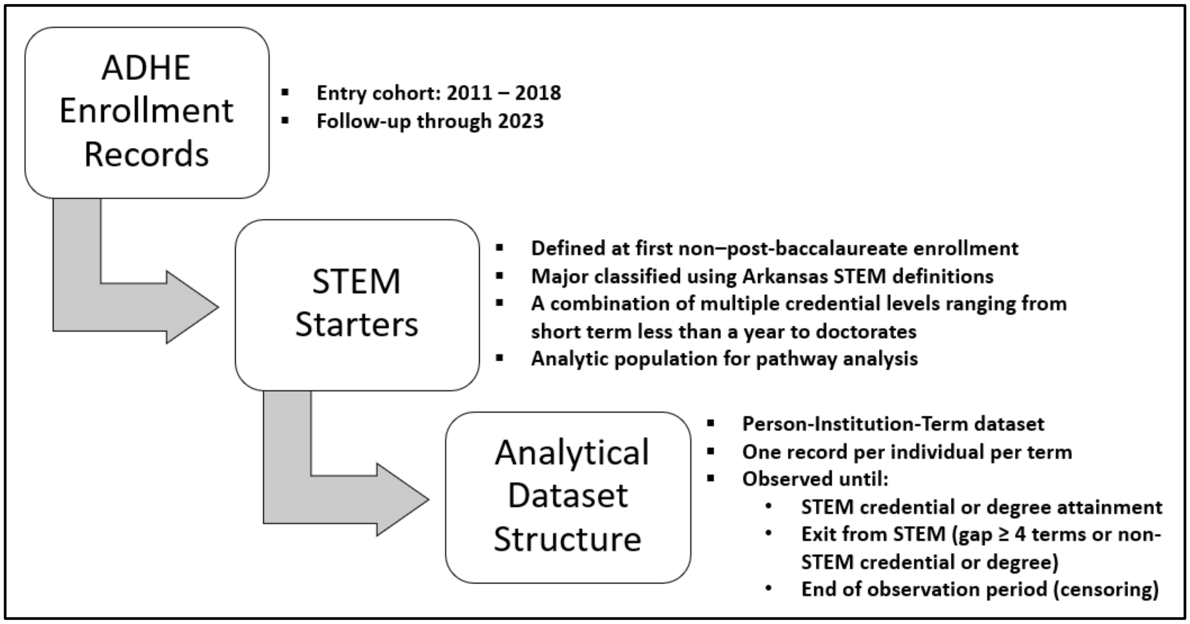


Figure 2 illustrates the selection of the analytic sample and the longitudinal structure of the data used in the analysis. Individuals enter the cohort at their first non-post-baccalaureate enrollment between academic years 2011 and 2018 and are followed at the term level through 2023. The figure also depicts the construction of the person-term dataset and the observation window used to capture enrollment continuity, degree attainment, and exit from postsecondary pathways. This longitudinal structure provides the foundation for the discrete-time hazard models used to examine persistence and attrition.

Definition of Key Measures

This section defines the key measures used to operationalize postsecondary STEM pathways, including STEM classification, postsecondary outcomes, workforce outcomes, and indicators of academic preparation. These measures are constructed to reflect the longitudinal progression of individuals through STEM pathways, from initial postsecondary entry to early labor market outcomes; they are applied consistently across descriptive and model-based analyses.

STEM Classification

Students are classified as STEM starters or non-STEM starters based on whether they are observed with a declared STEM major in any of the first four observed academic terms in the student enrollment data. Major fields are classified using Arkansas's STEM productivity reference tables, which map instructional program codes (CIP codes) to state-defined STEM categories. Individuals who are enrolled in a STEM major in at least one of these first four terms are classified as STEM starters; all others are classified as non-STEM starters[ii]. This definition is intended to capture entry into STEM pathways during the initial stage of postsecondary enrollment while allowing for early movement across majors. This initial classification serves as the baseline condition for all subsequent analyses. Individuals classified as STEM starters form the primary analytic population for examining persistence, attrition, and degree attainment within STEM pathways.

Postsecondary Outcomes

The analysis focuses on movement through postsecondary STEM pathways, with particular attention to persistence, attrition, and degree attainment. Outcomes are defined at the individual level and evaluated across degree levels, including associate, bachelor’s, master’s, and doctoral degrees, with distinctions made between STEM and non-STEM fields of study.

- Persistence is defined as continued enrollment in a STEM field without meeting exit criteria.
- Attrition from STEM is defined as a transition out of the STEM pathway and may occur through two mechanisms:
    - completion of a non-STEM degree
    - prolonged interruption in postsecondary enrollment
- Degree attainment is treated as a terminal outcome at the degree level of interest. Individuals who complete a STEM degree are considered to have successfully persisted through the STEM pathway for that degree level, while those completing non-STEM degrees are classified as having exited the STEM pathway prior to completion.

Workforce Outcomes

Workforce outcomes are measured using quarterly earnings data derived from Unemployment Insurance (UI) wage records. These data provide information on employment and earnings for jobs covered under the UI system within Arkansas.

Two primary workforce measures are constructed:

- Employment status, based on the presence of UI-covered earnings in a given period
- Earnings, measured as annual wages constructed from summed quarterly earnings and summarized using average wages[4]

Due to known biases in administrative data reporting (Gudipati, et al., 2025), any calculations regarding overall rates of employment were avoided and wage analyses focused solely on wages that were present in the data.

In addition, employment in STEM-related industries is identified using employer industry classifications based on the North American Industry Classification System (NAICS). STEM-related industries include computer system design, engineering services, scientific research and development, telecommunications, and advanced technology manufacturing (refer to Appendix B for a list of industries included). Because NAICS codes classify industries rather than

[4] Annual wages are calculated by summing quarterly wages within each calendar year. For the overall analysis, wages across all employers are aggregated to the individual level. For the NAICS-specific analysis, annual wages reflect the sum of quarterly wages earned within the specified industry.

occupations, this measure captures employment within a STEM-related industry rather than identifying specific STEM occupations. Further caution should be applied when interpreting rates of industry employment, due to limitations on the workforce data as well as the biases mentioned above.

K-12 indicators

Measures of academic preparation prior to postsecondary entry are derived from linked K-12 administrative records. Measures of academic preparation prior to postsecondary entry are derived from linked K-12 administrative records. In the current analysis, these include high school GPA, participation in STEM coursework, and participation in career and technical education (CTE). These indicators are used to provide additional context for understanding how pre-college preparation may be associated with subsequent persistence, attrition, and degree attainment within postsecondary STEM pathways.

Findings and Discussion

We begin our discussion of our findings and results with a descriptive overview of the STEM-starter analytical sample, before moving on to hazard and subgroup analyses.

Descriptive Overview of Postsecondary STEM Pathways

Postsecondary STEM pathways in Arkansas exhibit substantial variation in structure, timing, and enrollment continuity. Students who begin their postsecondary careers in STEM fields enter the pipeline through both two-year and four-year institutions, with relatively modest differences in initial STEM participation by institution type. Differences in outcomes emerge over time rather than at entry, underscoring the importance of examining persistence and attrition as dynamic processes rather than static endpoints (Chen, 2013; Rask, 2010; Speer, 2023).

To establish the overall structure of these pathways, we begin by examining the relationship between students' initial field of study and their eventual degree outcomes.

Figure 3. Postsecondary Pathways by STEM Starter Status and Exit Outcomes in Arkansas

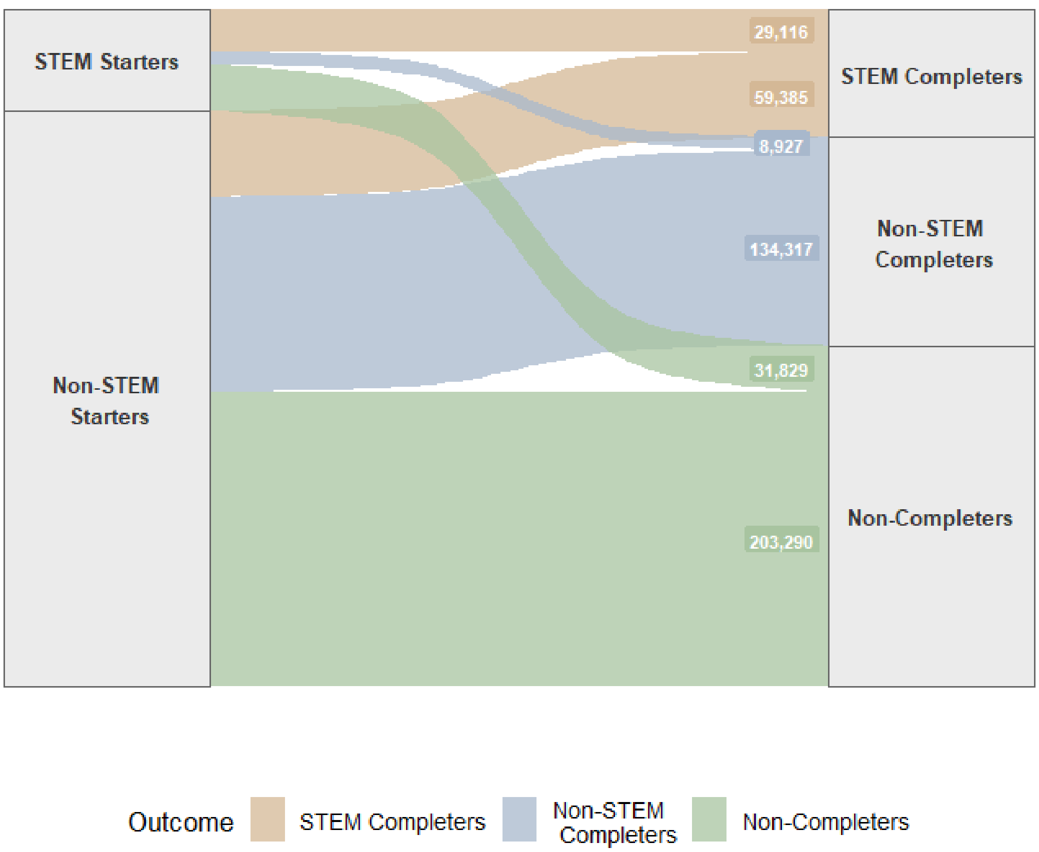


Figure 3 illustrates the distribution of postsecondary pathways from initial field of study to eventual degree outcomes in Arkansas. Among students who begin their postsecondary careers in STEM fields, a substantial share complete a STEM degree (29,116, or about 42 percent of the STEM starter sample), while others transition to non-STEM fields (13 percent) or exit postsecondary education without earning a degree (45 percent). At the same time, the figure highlights that STEM degree production is not limited to STEM starters: a notable number of students who begin outside of STEM ultimately complete a STEM degree (59,385, or about 15 percent of the non-STEM starters), reflecting movement into STEM fields after initial enrollment.

These patterns reveal an important asymmetry in postsecondary pathways. While some students enter STEM and subsequently leave, others enter non-STEM fields and later transition into STEM, indicating that the STEM pipeline is not strictly linear. However, the overall flow also shows that non-completion remains a dominant outcome across both groups, underscoring the extent to which initial entry into STEM does not guarantee persistence through degree completion.

These descriptive patterns highlight the importance of examining not only final outcomes, but also the timing and mechanisms through which students persist in or exit STEM pathways. In the sections that follow, we build on this descriptive overview by examining when attrition from STEM is most likely to occur and how these patterns vary across student and institutional characteristics.

Enrollment Patterns and Institutional Context

Degree attainment provides an important lens for understanding movement through STEM pathways, as completion of a degree, whether in STEM or non-STEM fields, represents a terminal outcome that reflects either persistence within or departure from STEM.

Descriptive analyses indicate that overall degree attainment rates are broadly similar between individuals who begin postsecondary education in STEM fields and those who begin outside of STEM. However, differences emerge in the direction of field movement. Among STEM starters, the rate of completing a non-STEM degree (about 13 percent) is slightly lower than the rate at which non-STEM starters transition into and complete STEM degrees (about 15 percent), highlighting the importance of pathways into STEM that occur after initial postsecondary entry.

An institutional entry point plays a significant role in shaping degree outcomes. Students who begin at four-year institutions are more likely to attain a degree overall compared to those who begin at two-year institutions. Notably, there are only minor differences in initial STEM participation by institution type; students entering two-year and four-year institutions exhibit similar rates of STEM entry, even under broader definitions of STEM fields. However, outcomes diverge over time, with four-year starters more likely to complete STEM degrees, suggesting that differences in persistence and attainment are driven more by postsecondary experiences and institutional context, or perhaps individual-level considerations such as opportunity cost.

To better understand longitudinal enrollment behavior, we examine patterns of institutional attendance and enrollment continuity. Most individuals attend a single institution throughout their observed postsecondary careers; however, enrollment gaps remain present within these trajectories, indicating that discontinuities are not limited to students who transfer across institutions. At the same time, there is a clear positive association between institutional mobility and enrollment gaps. Individuals who attend multiple institutions—approximately 20 percent of undergraduate students and 10 percent of graduate students—are more likely to experience interruptions in enrollment, reflecting more complex and less continuous postsecondary trajectories.

These patterns have important implications for modeling persistence and attrition, as they highlight the need to distinguish between short-term interruptions and longer-term exits from postsecondary education or STEM pathways. Taken together, these descriptive findings underscore the importance of institutional context, enrollment continuity, and pathway flexibility in shaping postsecondary STEM outcomes and motivate the use of hazard-based modeling approaches that explicitly account for the timing and sequencing of enrollment and exit behavior.

Observation Windows, Exit Conditions, and Censoring

Building on these descriptive patterns, we next examine when students are most likely to exit STEM pathways and how attrition risk evolves over time using discrete-time hazard models. The descriptive patterns summarized above document substantial heterogeneity in postsecondary STEM pathways, particularly with respect to degree attainment, time to completion, and enrollment continuity. While these results provide important context, they do not capture when students are most likely to exit STEM pathways or how attrition risk evolves over time. To address these questions, we examine term-to-term persistence and exit using discrete-time hazard models.

Observation windows hazard analyses are defined empirically based on observed average durations to degree attainment by degree level. Observation windows were informed by observed average time to degree attainment across the analytic sample. As expected, average time to attainment increased with degree level and was generally longer for bachelor's degree attainment among students who began at two-year institutions than among those who began at four-year institutions, reflecting transfer pathways and enrollment sequencing.[5]

For the hazard curve analyses presented below, individuals enter the risk set at the time of their initial postsecondary enrollment and are followed longitudinally until one of the following occurs:

- STEM degree attainment (treated as a terminal outcome)
- Exit from the STEM pathway, defined as:
    - completion of a non-STEM degree of the same or higher level, or
    - a prolonged enrollment gap, defined as an interruption in enrollment of four or more consecutive academic terms
- End of the observation window, at which point individuals are treated as right-censored

For each degree type, the observation window is set at 150 percent of the average observed time to degree.

For the hazard regression models, all degree types were included in the analytical sample. The exit condition was set to completion of a non-STEM Bachelor's degree (or an interruption in enrollment of four or more consecutive academic terms), and the observation window was set to 19 terms, or 150 percent of the average observed time to a Bachelor's degree. This approach allows progression into advanced degrees to be modeled, while limiting the influence of students who progress through multiple degrees on the final model results.

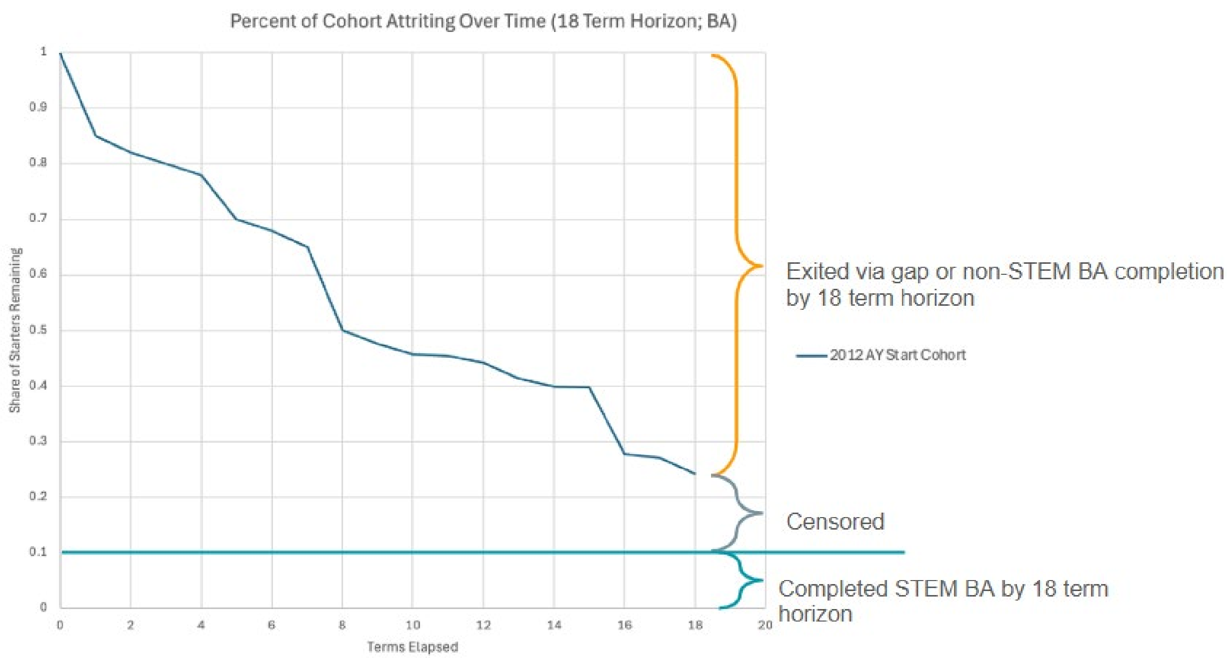
[5] Average time to attainment was approximately 3-4 years for associate degrees, 3-5 years for bachelor's degrees depending on starting institution, about 6 years for master's degrees, and roughly 7-9 years for doctoral degrees.

Throughout the analysis, “exit” refers specifically to exit from the STEM pathway rather than exit from postsecondary education more broadly. Depending on the analytic context, this may occur through degree completion in a non-STEM field or departure from postsecondary enrollment.

Timing of STEM Persistence and Attrition

Figure 4 presents an example of cohort survival along the pathway to a STEM bachelor’s degree, illustrating how persistence and attrition unfold across academic terms. Individuals exit the risk set upon attainment of a non-STEM degree or upon experiencing a prolonged enrollment gap, while those who remain enrolled without meeting exit criteria by the end of the observation window are treated as right-censored.

Figure 4. Example survival curve illustrating persistence and exit along the pathway to a STEM bachelor’s degree

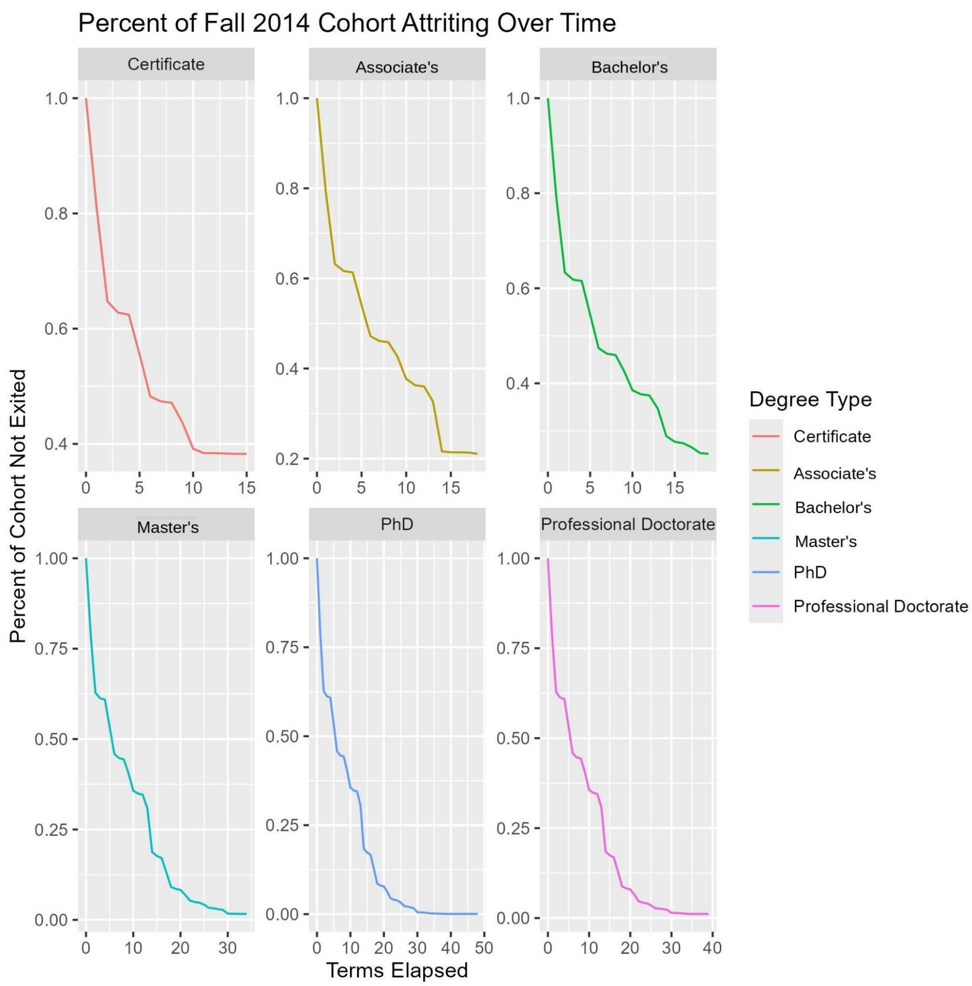


The survival curve shows that attrition risk is concentrated in the earliest terms of postsecondary enrollment, with the probability of persistence declining most sharply during the initial stages of the pathway. After this initial period, persistence stabilizes substantially among students who remain enrolled, with comparatively lower rates of exit in later terms.

Variation in Attrition by Degree and Cohort

To examine how these patterns vary across degree pathways and student cohorts, Figures 5 and 6 present survival curves by degree type and entry cohort, respectively.

Figure 5. Survival curves by degree type illustrating persistence and attrition in STEM pathways



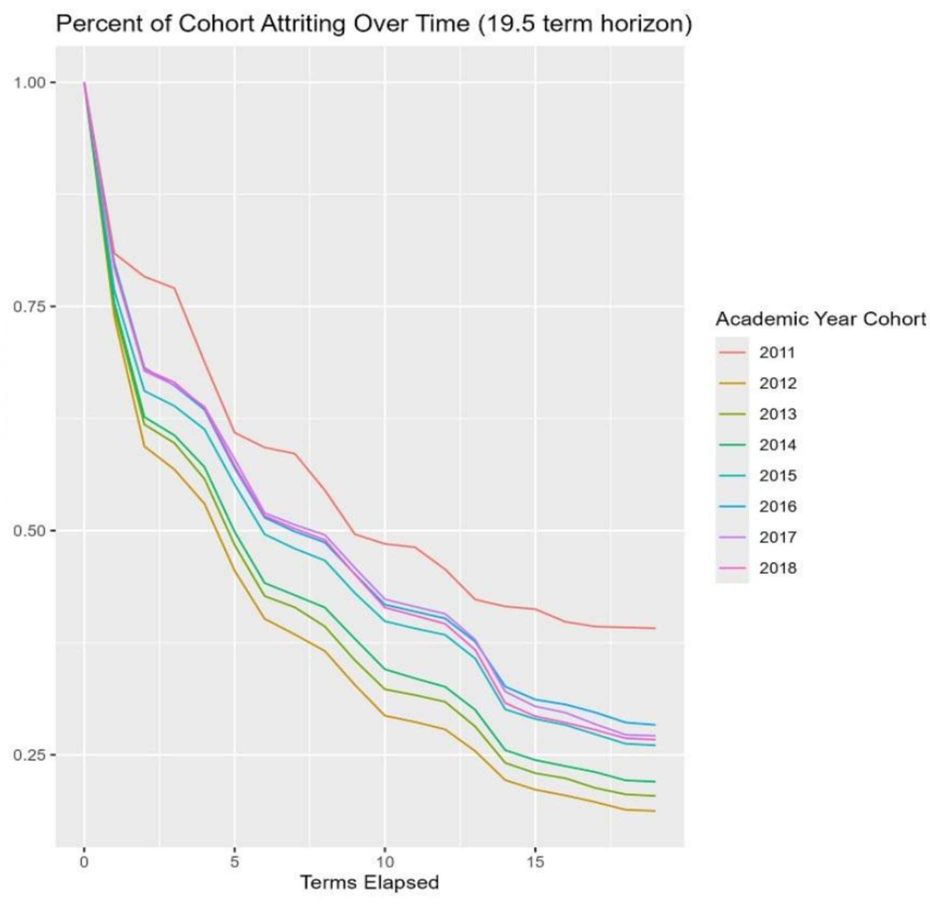
Across all degree levels, the proportion of individuals remaining on the survival curve over time closely corresponds to the share of STEM starters who ultimately complete that degree in STEM. As degree levels increase, this proportion declines substantially, reflecting cumulative attrition along the pathway and the fact that only a subset of students persist to more advanced credentials.

Survival patterns for advanced degrees, including master's and doctoral programs, are broadly similar across cohorts. This reflects the relatively small number of students who progress to these levels, as most attrition occurs earlier in the postsecondary trajectory, particularly around the stages associated with associate and bachelor's degree completion. At the same time, the persistence of a subset of students into graduate-level STEM pathways underscores the existence of longer-term trajectories within the pipeline.

Figure 6 further illustrates cohort-specific survival patterns along the pathway to a STEM Bachelor’s degree. Although modest differences are evident across entry cohorts, particularly in long-run persistence, the overall shape of the survival curves is remarkably consistent. Attrition is concentrated in the earliest terms across all cohorts, with the steepest declines occurring within the first four to six academic terms. Beyond this point, persistence stabilizes and exit rates decline substantially among students who remain enrolled.

For more recent cohorts, survival curves for advanced degrees decline more sharply and, in some cases, reach zero within the observation window. This pattern reflects right-censoring due to limited follow-up time rather than true absence of completion, as these cohorts have not yet had sufficient time to reach advanced degree milestones.

Figure 6. Survival curves for the STEM Bachelor’s degree pathway by cohort (2011–2018)



Hazard Model Results

To formally estimate the relationship between elapsed enrollment time and the probability of exit from STEM pathways, we employ discrete-time hazard models using logistic regression. This approach is well suited to the term-level structure of the administrative data and allows the conditional probability of exit to be estimated at each enrollment period, given continued enrollment up to that point. A full technical specification of the hazard function and estimation approach is provided in Appendix A.

Across specifications, hazard model results indicate that attrition risk is highest during the earliest terms of postsecondary enrollment and declines over time. The estimated hazard of exit in the second term is approximately 80 percent of the first-term hazard and declines further by the fourth term. These patterns align closely with the descriptive survival analyses and indicate that the risk of STEM attrition is most pronounced early in postsecondary enrollment, before stabilizing in later terms. These findings are consistent with prior research on STEM persistence and major switching (Chen, 2013; Astorne-Figari & Speer, 2019; Speer, 2023) and reinforce the importance of early postsecondary experiences in shaping long-term STEM outcomes. Full model estimates, including regression coefficients and standard errors, are reported in Appendix Table A1.

Variation in attrition across demographic, geographic, and programmatic factors is examined in the Subgroup section.

Implications

The findings from the descriptive and hazard analyses consistently indicate that attrition from postsecondary STEM pathways is concentrated in the early stages of enrollment. Differences in outcomes are not driven primarily by variation in initial entry into STEM but instead emerge as students progress through postsecondary education, particularly during the first several academic terms.

These results underscore the importance of early postsecondary experiences, especially for students beginning at two-year institutions, where early momentum and access to institutional supports are strongly associated with longer-term outcomes (Bahr et al., 2017; Fink et al., 2024; Wang, 2015). More broadly, the findings highlight the value of viewing STEM pathways as dynamic processes shaped by timing, sequencing, and institutional context rather than as static endpoints.

Methodologically, the alignment between descriptive survival patterns and hazard model estimates provides confidence in the analytical framework and supports its use for identifying periods of elevated attrition risk. At the same time, the analysis illustrates the importance of clearly defined observation windows, exit conditions, and censoring assumptions when working with longitudinal administrative data.

Finally, while the present analysis focuses on postsecondary persistence and degree attainment, the framework is designed to extend to education-to-workforce transitions. Linking postsecondary pathways to labor market outcomes provides an opportunity to assess how differences in persistence and completion translate into variation in employment and earnings, and to evaluate the alignment between STEM education and workforce demand (Buffington et al., 2016; Jelks & Crain, 2020).

Labor Market Outcomes by Postsecondary STEM Pathway

Building on the analysis of postsecondary persistence and degree attainment, we examine how variation in STEM pathways translates into labor market outcomes. By linking postsecondary records to Unemployment Insurance (UI) wage data, the analysis provides insight into employment and earnings among STEM starters following postsecondary enrollment.

Workforce outcomes are measured using UI wage records, which capture quarterly earnings for employment covered within the Arkansas UI system. These data allow individuals to be followed from postsecondary education into the labor market, providing a consistent basis for examining employment and earnings outcomes associated with different postsecondary trajectories. However, these measures reflect in-state, UI-covered employment only and do not capture out-of-state employment or work not covered by the UI system. A substantial majority of STEM starters are observed in UI-covered employment at least once following postsecondary enrollment, indicating strong overall labor market attachment within the observed system.

To get a complete picture of the workforce pathways for STEM starters, STEM completers, and non-STEM completers, we take two different approaches. First, to see how STEM starters fare in specific industries, we base the industry wage analysis on the sample from the hazard analyses above. However, because this sample was restricted to students whose first observation was pre-graduate and students were only observed for the relevant observation windows, to get a more complete picture of how STEM completers' wages compare to non-STEM completers, we examine wages for the entire graduating class of 2014 for the overall wage analyses presented below.

Wage outcomes differ markedly across postsecondary pathways. STEM degree completers exhibit higher average wages compared to both non-STEM completers and individuals who exit without earning a degree. Non-STEM completers also demonstrate higher earnings than non-completers, though the gap between STEM and non-STEM degree holders remains substantial. These patterns indicate that both degree attainment and field of study are important predictors of early labor market outcomes.

Table 1. Average Wages by Postsecondary STEM Pathway, 2014 Graduating Class

Postsecondary Pathway	Post-3-year	Post-5-year

STEM Completers (all)	$29,856.31	$35,526.42
– Associate	$28,852.06	$31,835.80
– Bachelor’s	$31,224.41	$40,405.11
– Master’s	$50,174.19	$55,768.51
– Doctoral	$72,304.97	$70,316.94[6]
Non-STEM Completers (all)	$22,363.82	$28,277.75
– Associate	$17,282.46	$22,926.312
– Bachelor’s	24689.8641	$31,744.54
– Master’s	$37,882.47	$42,316.93
– Doctoral	$53,194.50	$86,866.69

Note: Post-3-year and post-5-year refer to time since graduation.

Table 1 reports average wages for STEM starters of all cohorts at three and five years following graduation, disaggregated by degree attainment. These results provide a concise summary of how differences in postsecondary pathways translate into variation in early labor market outcomes.

Wage outcomes vary substantially across postsecondary pathways and degree levels. STEM completers exhibit 34 percent higher earnings than non-STEM completers, with additional variation by degree level. In particular, individuals who complete bachelor’s and graduate-level STEM degrees tend to have higher earnings at both the three- and five-year marks, reflecting the labor market returns to both degree attainment and field of study.

In addition to differences in earnings levels, variation across pathways reflects differences in the types of industries in which individuals are employed.

Table 2 examines differences in wage outcomes across major industry sectors using NAICS-based classifications, comparing STEM and non-STEM completers, as well as the percentage of each group that was employed in the industry. Table 2 examines students at any time after graduation.[7] These results provide additional insight into how differences in postsecondary pathways translate into variation in employment sectors and associated earnings. STEM degree completers are more likely to be employed in industries with higher concentrations of STEM-related activity, having a higher concentration in all but one sector (Information).

[6] Note the decrease in wages for STEM-completer’s who earned doctorate degrees. This seems to be an anomaly in the data, as this category sees an increase in wages in every other set of years.
[7] The average wages in Table 2 include annual wages for all individuals in these industries for any year after graduation. It is possible that students worked in multiple industries, so an individual’s wages could be counted twice and an individual could appear in multiple industry employment rates.

Table 2. Average Wages and Employment Concentration by Industry and STEM Completion Status (NAICS-Based), STEM Starters

STEM-related Industry (NAICS)	Non-STEM Completers - Avg Wage (% of Group)	STEM Completers - Avg Wage (% of Group)
Manufacturing - 325412, 325414 - 334111, 334220, 334413, 334511, 334512, 336411, 339112, 339113, 339114, 339115, 339116	- $28,161.99 (0.34) - $34,882.52 (3.21)	- $49,469.68 (0.41) - $46,259.60 (3.42)
Information 511210, 517311, 517312, 517919, 518210, 519130	- $103,988.47 (6.74)	- $141,465.72 (6.01)
Professional & Technical 541330, 541380, 541511, 541512, 541513, 541519, 541620, 541690, 541714, 541715	- $53,906.06 (14.7)	- $112,543.06 (15.6)
Education 611310, 611519	- $63,387.49 (14.5)	- $124,240.49 (19)

Wage outcomes within industries also differ between groups. In several STEM-related industries, STEM completers exhibit higher average wages than their non-STEM counterparts, reflecting the role of specialized skills in shaping earnings. Furthermore, there are a wide variety of occupations represented in these industries, both STEM-related and non-STEM occupations. It is likely that the non-STEM completers are being sorted into non-STEM occupations within STEM-related industries, and this sorting contributes to the wage differential. At the same time, variation across industries indicates that labor market outcomes are influenced not only by degree attainment but also by the sectors in which individuals are employed.

These findings suggest that differences in labor market outcomes across postsecondary pathways are driven in part by differential sorting into industries, with STEM completers more likely to enter higher-paying, STEM-intensive sectors. However, these results should be interpreted with caution. In the absence of occupation-level data (SOC), the analysis relies on industry classifications (NAICS) as a proxy for STEM employment. Because industries include a mix of STEM and non-STEM occupations, this approach does not allow for a direct assessment of the alignment between STEM degree attainment and employment in STEM occupations. Further, additional caution should be used when interpreting these findings, as we only have employment information for the state of Arkansas, so these rates are best interpreted as the rate of employment in an industry within Arkansas.

Differences in labor market outcomes are also evident in wage trajectories over time. Individuals who complete STEM degrees tend to enter the labor market with higher initial earnings and experience more sustained wage growth in subsequent periods. In contrast, individuals who complete non-STEM degrees exhibit lower and more variable earnings trajectories, reflecting greater instability in labor market outcomes. Differences across time points also suggest that earnings gaps widen as individuals progress further into the labor market. For example, STEM completers had an average wage of $35,526.42 five years after graduation, while non-STEM completers had an average wage of $28,277.75, for a difference of $7,248.67. Ten years after graduation, this difference had grown to $9,057.68, with STEM completers earning an average of $55,887.93 to non-STEM completers' $46,830.25 These differences underscore the cumulative effect of postsecondary persistence and degree attainment on longer-term economic outcomes.

Figure 7. Wage Trajectories, 2014 Graduating Class

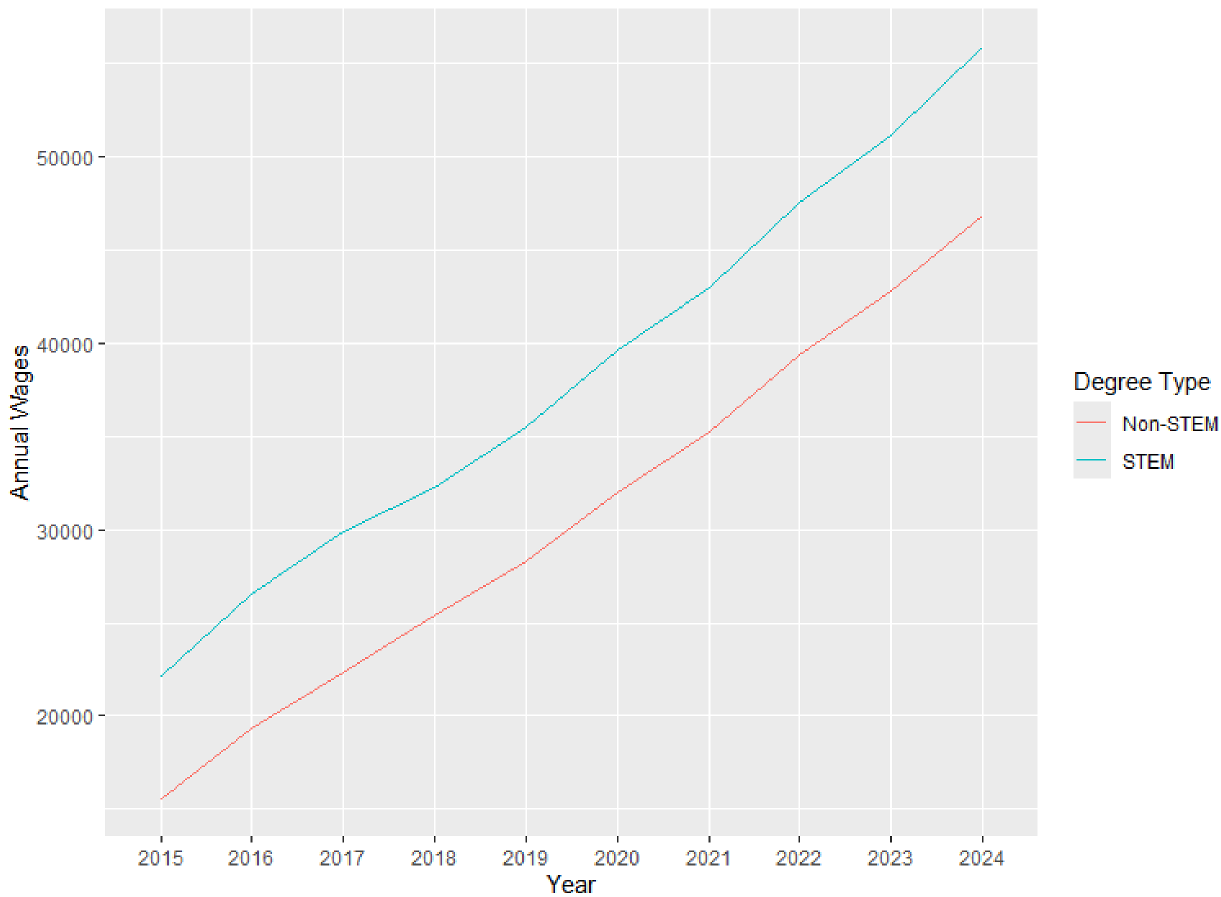


Figure 7 illustrates these patterns by showing wage trajectories over time across postsecondary pathways.

Subgroup and Contextual Analysis of STEM Pathways

Distribution of STEM Starters and Completers Across Subgroups

To provide context for subgroup analyses, we begin by describing the demographic and geographic composition of students who enter postsecondary STEM pathways.[8] These

[8] The absence of direct measures of socioeconomic status in the available data (e.g., PELL, financial aid) limited the ability to assess the role of financial resources in shaping persistence and thus represents an important area for future analysis.

distributions establish the baseline population from which persistence and attrition patterns are examined.

The population of STEM starters reflects substantial diversity across demographic characteristics, including gender, race, and ethnicity. Notably, STEM starters are distributed relatively evenly by gender, with female and male students represented in approximately equal proportions (34,851 female students to 35,111 male students)[9]. This pattern differs from commonly observed national trends, where male students have historically been overrepresented in STEM fields (Wang, 2015; Riegle-Crumb & King, 2010). In the Arkansas context, this balance persists through degree attainment, as STEM completers are also roughly evenly divided by gender (14,629 female completers compared to 14,538 male completers).

The racial and ethnic composition of STEM starters reflects both concentration and diversity within the Arkansas postsecondary system. White students constitute the majority of STEM starters, accounting for approximately 65 percent of the population. Black or African American students represent a substantial share, comprising nearly 19 percent of STEM starters, while Hispanic students account for a smaller but meaningful proportion (5 percent). We can compare these proportions to the overall Arkansas population, which is 68 percent White, 15 percent Black, and 10 percent Hispanic[10].

The racial and ethnic composition of STEM completers reflects patterns broadly similar to those observed among STEM starters, with some differences in representation across groups. White students constitute the majority of STEM completers, accounting for approximately 70% of the population. Black or African American students represent a smaller share of completers (14%) relative to their representation among STEM starters, while Hispanic students comprise a modest proportion of degree recipients. Asian students and those identifying with two or more races represent smaller but consistent shares of completers.

Differences in representation between STEM starters and completers highlight the importance of examining persistence and attrition patterns across racial and ethnic groups, while recognizing that subgroup sizes vary and may influence the stability of estimates for smaller populations.

In particular, the relative decrease in representation among Black students between STEM entry and completion underscores the importance of examining differential persistence across groups.

Differences in Persistence and Attrition Across Subgroups

Building on the descriptive distributions presented above, we examine how persistence and attrition vary across subgroups using discrete-time hazard models. These analyses provide

[9] Note that overall demographic totals will differ slightly from the overall completion totals from Figure 3 because demographics characteristics are self-reported at institution entry and therefore a single individual may be associated with multiple demographic characteristics.
[10] https://www.census.gov/quickfacts/fact/table/AR/PST045225

insight into how demographic, socioeconomic, geographic, and institutional characteristics are associated with the likelihood of exiting STEM pathways over time.

When considered alongside the descriptive distributions of STEM starters and completers, these findings suggest that differences in persistence contribute to observed variation in degree attainment across groups. However, these estimates reflect conditional associations rather than causal effects and should be interpreted in the context of observed covariates included in the model or captured in the data.

Despite relatively balanced representation among both STEM starters and completers, hazard model estimates indicate modest differences in attrition risk by gender. Conditional on observed characteristics, male students exhibit slightly higher probabilities of exiting STEM pathways at certain stages of the postsecondary trajectory. This pattern suggests that similar overall rates of entry and completion may mask differences in how students persist through STEM pathways.

Geographic context is similarly associated with variation in STEM persistence. Students originating from the South Census region exhibit lower attrition risk compared to those from other regions, holding other characteristics constant. Within Arkansas, county-level variation further highlights the importance of local context, including differences in educational resources, institutional access, and support networks. These findings are consistent with research emphasizing the role of place-based factors in shaping postsecondary trajectories (Koricich, Chen, & Hughes, 2018; Qian & Koedel, 2024).

Institutional and enrollment contexts also play a significant role in shaping attrition risk. Students who begin at four-year institutions exhibit lower probabilities of exit compared to those who begin at two-year institutions, consistent with earlier descriptive findings on degree attainment. In addition, students who enter in non-traditional terms, such as the spring semester, are more likely to exit STEM pathways than fall entrants. These patterns align with evidence of greater enrollment instability and more complex trajectories among non-traditional entry cohorts.

Finally, variation in attrition risk across STEM fields suggests that programmatic context also influences persistence. Students in certain fields, such as computer science, exhibit lower attrition risk, while those in agriculture-related programs show comparatively higher exit probabilities.

These findings highlight that variation in STEM persistence reflects the interaction of individual characteristics, institutional pathways, and broader contextual factors, reinforcing the importance of targeted strategies to support persistence among diverse student populations.

K-12 Academic Preparation and Early STEM Indicators

Patterns in K-12 Academic Preparation among STEM Starters

To provide additional context on pre-college academic preparation, we examine a subset of STEM starters whose postsecondary records can be linked to K-12 administrative data from Arkansas. Approximately one-third of STEM starters in the analytic sample are matched to K-12 records. This match rate is consistent with expectations, as it reflects the proportion of students who complete high school in Arkansas and subsequently enroll in postsecondary institutions within the state.

Among matched students, K-12 indicators, including high school GPA, participation in STEM coursework, and engagement in career and technical education (CTE), provide insight into variation in academic preparation prior to postsecondary entry. Descriptive patterns suggest that students who persist in and complete STEM pathways tend to have stronger academic preparation, as reflected in higher GPAs (3.51 for STEM completers compared to 3.39 for those who completed a non-STEM degree and 2.86 for those who did not complete a degree) and greater exposure to STEM-related coursework with an institution of higher learning during high school (72% of STEM completers compared to 67% of non-completers). Participation in CTE pathways is also widespread, with more than 93% of matched students having taken at least one CTE course. This may partly reflect Arkansas' statewide emphasis on access to CTE courses, which has increased exposure to applied coursework across a broad range of students.[11]

These findings should be interpreted as descriptive rather than representative of the full STEM starter population, given that K-12 data are available only for students who remain within Arkansas across educational stages.

Incorporating K-12 Indicators in Models of STEM Persistence and Attrition

As an extension of the baseline hazard models, we incorporated selected K-12 indicators of academic preparation, including high school GPA, participation in STEM coursework with an institution of higher learning while in high school, and participation in CTE courses. Results indicate that academic performance prior to postsecondary entry is strongly associated with persistence in STEM pathways. In particular, higher high school GPA is associated with a substantially lower likelihood of exiting STEM, underscoring the importance of academic preparation in shaping postsecondary outcomes.

In contrast, indicators of STEM course-taking in high school show a modest positive association with attrition risk after controlling for other factors. Because these estimates are conditional on entry into STEM pathways, they should not be interpreted as reflecting the overall effect of STEM course-taking on STEM participation; they may instead capture differences in course

[11] See for example: https://access.ade.arkansas.gov/

intensity or selection into more demanding fields of study. Similar results hold for participation in CTE courses.

These results should be interpreted as preliminary, as K-12 data are only available for a subset of the sample[12]. However, they demonstrate the analytical value of incorporating pre-college indicators into models of STEM persistence and attrition.

Reproducibility & Future Work

This study is designed to support transparency and reproducibility across geographies, including regions or states, as well as other populations of interest within the constraints of secure administrative data environments. All analyses were conducted within the Administrative Data Research Facility (ADRF), where access to individual-level data is restricted and outputs are subject to disclosure review. As a result, while the underlying data cannot be shared directly, the analytical approach is documented in sufficient detail to support replication in comparable settings.

The construction of the analytic dataset follows a longitudinal person–term structure, linking individuals across K-12 (where available), postsecondary, and workforce records (when matched). Key variables, including indicators of STEM participation, enrollment continuity, degree attainment, and workforce outcomes, are defined using consistent rules that are documented in the technical appendix. The hazard modeling framework is similarly specified, including the definition of exit events, censoring rules, and model covariates, to facilitate reproducibility of results in other administrative data contexts.

At the same time, the analysis reflects several limitations inherent to administrative data. K-12 indicators are available for a subset of the population, and workforce outcomes are measured using Unemployment Insurance (UI) wage records that capture in-state, UI-covered employment only. In addition, STEM employment is assessed using industry classifications (NAICS), which do not provide direct information on occupation. These constraints limit the ability to fully capture education-to-workforce transitions, particularly for individuals who migrate out of state or work in non-covered employment.

Future work will build on this foundation in several directions. First, ongoing efforts will refine the integration of K-12 indicators into models of postsecondary persistence, including expanded measures of academic preparation and coursework. Second, additional work will further examine variation across subgroups and institutional contexts, including interactions between academic preparation and postsecondary experiences. In the current analysis, measures of socioeconomic status, such as financial aid or Pell eligibility, were not available in the analytic dataset, limiting the ability to fully assess the role of economic factors in shaping STEM persistence. Incorporating these measures would provide a more complete understanding of equity and access in STEM pathways.

[12] Due to the limited sample, some variables were removed from the K-12 model for disclosure reasons.

Third, while the analysis includes geographic indicators, more detailed investigation of regional variation, particularly at finer geographic levels such as counties or local labor markets, would further illuminate place-based differences in STEM pathways and outcomes. Finally, improvements in workforce data, particularly the incorporation of occupation-level information, would enable more precise analysis of alignment between STEM education and employment outcomes.

To support reproducibility and scalability, the project team is also developing detailed data documentation and analytic code that can be exported from the ADRF environment, subject to disclosure review. These materials are intended to facilitate the application of this analytical framework to non-STEM fields within Arkansas, as well as to comparable linked administrative data systems in other states. More broadly, this approach supports the extension of longitudinal pathway analysis across domains and jurisdictions.

References

Agarwal, K. (2026). America's STEM Talent Pipelines Are Under Strain. Association of American Universities. https://www.aau.edu/newsroom/leading-research-universities-report/americas-stem-talent-pipelines-are-under-strain

Astorne-Figari, C., & Speer, J. D. (2019). Are changes of major major changes? The roles of grades, gender, and preferences in college major switching. Economics of Education Review, 70, 75–93. https://doi.org/10.1016/j.econedurev.2019.03.005

Bahr, P. R., Jackson, G., McNaughtan, J., Oster, M., & Gross, J. (2017). Unrealized potential: Community college pathways to STEM baccalaureate degrees. The Journal of Higher Education, 88(3), 430–478. https://doi.org/10.1080/00221546.2016.1257313

Bettencourt, G. M., Manly, C. A., Kimball, E., & Wells, R. S. (2020). STEM degree completion and first-generation college students: A cumulative disadvantage approach to the outcomes gap. The Review of Higher Education, 43(3), 753–779. https://doi.org/10.1353/rhe.2020.0006

Bhatti, H. (2025). One Million More: Assessing a Decade of Progress in Undergraduate STEM Education. Journal of Microbiology & Biology Education, 26(3), https://journals.asm.org/doi/10.1128/jmbe.00155-25

Buffington, C., Cerf, B., Jones, C., & Weinberg, B. A. (2016). STEM training and early career outcomes of female and male graduate students: Evidence from UMETRICS data linked to the 2010 census. American Economic Review, 106(5), 333–338. https://doi.org/10.1257/aer.p20161124

Carnevale, A. P., Smith, N., & Melton, M. (2011). STEM: Science, technology, engineering, mathematics. Georgetown University Center on Education and the Workforce.

Chen, R. (2012). Institutional characteristics and college student dropout risks: A multilevel event history analysis. Research in Higher Education, 53(5), 487–505. https://doi.org/10.1007/s11162-011-9241-4

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM fields (NCES 2014-001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Chow, M. & Goldschlag, N. (2023). High-Tech Industries Are Concentrated in Five Coastal Metro Areas. US Census Bureau. https://www.census.gov/library/stories/2023/02/where-in-the-united-states-are-the-high-tech-jobs.html

Ehrenberg, R. G., Jakubson, G. H., Groen, J. A., So, E., & Price, J. (2007). Inside the black box of doctoral education: What program characteristics influence doctoral students' attrition and graduation probabilities? Educational Evaluation and Policy Analysis, 29(2), 134–150. https://doi.org/10.3102/0162373707301707

Engberg, M. E., & Wolniak, G. C. (2013). College student pathways to the STEM disciplines. Teachers College Record, 115(1), 1–27. https://doi.org/10.1177/016146811311500102

Fink, J., Myers, T., Sparks, D., & Smith Jaggars, S. (2024). Toward a practical set of STEM transfer program momentum metrics. Research in Higher Education, 65(2), 259–282. https://doi.org/10.1007/s11162-023-09759-2

Gudipati, V., Bejugam, P., Saco, H., & McGough, R. (2025). Roadmap for improving the fidelity of administrative data linkage in support of evidence-based policy and practice on the foreign-born science and engineering workforce. Prepared under the Foreign-Born Scientists and Engineers and the US Workforce project of the America's DataHub Consortium.

Green, A., & Sanderson, D. (2018). The roots of STEM achievement: An analysis of persistence and attainment in STEM majors. The American Economist, 63(1), 79–93. https://doi.org/10.1177/0569434517721770

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? Economics of Education Review, 29(6), 911–922. https://doi.org/10.1016/j.econedurev.2010.06.010

Jelks, S. M. R., & Crain, A. M. (2020). Sticking with STEM: Understanding STEM career persistence among STEM bachelor's degree holders. The Journal of Higher Education, 91(5), 805–831. https://doi.org/10.1080/00221546.2019.1700477

Kim, D., Jung, J., & Mlambo, Y. A. (2021). Institutional selectivity and occupational outcomes for STEM graduates: A generational comparison. The Journal of Higher Education, 92(3), 435–464. https://doi.org/10.1080/00221546.2020.1819945

Kim, Y. K., & Sax, L. J. (2018). The effect of positive faculty support on mathematical self-concept for male and female students in STEM majors. Research in Higher Education, 59(8), 1074–1104. https://doi.org/10.1007/s11162-018-9500-8

King, B. (2016). Does postsecondary persistence in STEM vary by gender? AERA Open, 2(4). https://doi.org/10.1177/2332858416669709

Koricich, A., Chen, X., & Hughes, R. P. (2018). Understanding the effects of rurality and socioeconomic status on college attendance and institutional choice in the United States. The Review of Higher Education, 41(2), 281–305. https://doi.org/10.1353/rhe.2018.0004

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). STEM: Good jobs now and for the future. U.S. Department of Commerce, Economics and Statistics Administration.

Lee, Y. G., & Ferrare, J. J. (2019). Finding one's place or losing the race? The consequences of STEM departure for college dropout and degree completion. The Review of Higher Education, 43(1), 221–261. https://doi.org/10.1353/rhe.2019.0095

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. Science Education, 95(5), 877–907. https://doi.org/10.1002/sce.20441

Mendoza, P., Villarreal, P., & Gunderson, A. (2014). Within-year retention among Ph.D. students: The effect of debt, assistantships, and fellowships. Research in Higher Education, 55, 650–685. https://doi.org/10.1007/s11162-014-9327-x

National Science Board, National Science Foundation. 2026. STEM Talent: Education, Training, and Workforce. Science and Engineering Indicators 2026. NSB-2026-1. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/nsb20261/.

National Science Foundation. (2024). The state of U.S. science and engineering 2024. National Science Board. https://ncses.nsf.gov/pubs/nsb20243

Perez-Felkner, L., Thomas, K., Nix, S., Hopkins, J., & D'Sa, M. (2018). Are 2-year colleges the key? Institutional variation and the gender gap in undergraduate STEM degrees. The Journal of Higher Education, 90(2), 181–209. https://doi.org/10.1080/00221546.2018.1486641

Qian, C., & Koedel, C. (2024). Could shifting the margin between community college and university enrollment expand and diversify university degree production in STEM fields? Research in Higher Education, 65(1), 42–69. https://doi.org/10.1007/s11162-023-09750-x

Rask, K. N. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. Economics of Education Review, 29(6), 892–900. https://doi.org/10.1016/j.econedurev.2010.06.013

Riegle-Crumb, C., & King, B. (2010). Questioning a white male advantage in STEM: Examining disparities in college major by gender and race/ethnicity. Educational Researcher, 39(9), 656–664. https://doi.org/10.3102/0013189X10391657

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. Educational Researcher, 48(3), 133–144. https://doi.org/10.3102/0013189X19831006

Roberts, B & Wolf, M. (2018). High-tech industries: an analysis of employment, wages, and output. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-7/high-tech-industries-an-analysis-of-employment-wages-and-output.htm

Speer, J. D. (2023). Bye bye Ms. American Sci: Women and the leaky STEM pipeline. Economics of Education Review, 93, 102371. https://doi.org/10.1016/j.econedurev.2023.102371

Stevens, M. L., Armstrong, E. A., & Arum, R. (2008). Sieve, incubator, temple, hub: Empirical and theoretical advances in the sociology of higher education. Annual Review of Sociology, 34(1), 127–151. https://doi.org/10.1146/annurev.soc.34.040507.134737

Wang, X. (2013a). Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. Research in Higher Education, 54(6), 664–692. https://doi.org/10.1007/s11162-013-9291-x

Wang, X. (2015). Pathway to a baccalaureate in STEM fields: Are community colleges a viable route and does early STEM momentum matter? Educational Evaluation and Policy Analysis, 37(3), 376–393. https://doi.org/10.3102/0162373714552561

Weeden, K. A., Gelbgiser, D., & Morgan, S. L. (2020). Pipeline dreams: Occupational plans and gender differences in STEM major persistence and completion. Sociology of Education, 93(4), 297–314. https://doi.org/10.1177/0038040720928484

Xie, Y., Fang, M., & Shauman, K. (2015). STEM education. Annual Review of Sociology, 41(1), 331–357. https://doi.org/10.1146/annurev-soc-071312-145659

Xue, Y. & Larson, R. (2015). STEM crisis or STEM surplus? Yes and yes. Monthly Labor Review, US Bureau of Labor Statistics. https://doi.org/10.21916/mlr.2015.14

Yu, H., McKinney, L., Burridge, A., & Hernández, S. H. (2024). How first-year academic momentum influences transfer outcomes among different racial and ethnic groups. Research in Higher Education, 65(7), 1461–1490. https://doi.org/10.1007/s11162-024-09803-9

Appendices

Appendix A. Technical Specification of Discrete-Time Hazard Models

This appendix provides a concise, non-technical description of the discrete-time hazard modeling framework used to analyze attrition from postsecondary STEM pathways. The purpose is to document key modeling choices and assumptions in accessible language while maintaining transparency and reproducibility. This summary complements the Methods section in the main text.

Definition of STEM Pathway Exit
The analysis models exit from a STEM pathway among individuals who begin their postsecondary careers in STEM fields. Exit is defined as the first academic term in which an individual either (1) leaves postsecondary enrollment following a sufficiently long enrollment gap or (2) completes a non-STEM degree of the same or higher level as the degree of interest within the observation window. Completion of the STEM degree of interest is treated as a terminal outcome, while individuals who neither exit nor complete a degree by the end of the observation window are treated as right-censored.

Time Scale, Risk Set, and Observation Window
Time is measured in discrete academic terms beginning with an individual's first observed postsecondary enrollment. Each individual enters the risk set at initial enrollment and remains at risk until exiting the STEM pathway, completing the STEM degree of interest, or reaching the end of the observation window. Observation windows are defined separately by degree level and are set empirically at 150 percent of the average observed time to degree, ensuring sufficient follow-up while limiting the influence of extreme durations.

Modeling Strategy
Attrition is analyzed using discrete-time hazard models estimated via logistic regression. In this framework, the model estimates the probability that an individual exits a STEM pathway in a given academic term, conditional on having persisted in STEM through all prior terms. A flexible representation of elapsed enrollment time is included to capture how attrition risk changes as students progress through postsecondary education, with particular attention to early enrollment periods.

Covariates capture student background and demographics, academic preparation and enrollment patterns, institutional context, and program characteristics. Both time-invariant and time-varying measures are included, and some specifications incorporate cohort or calendar-period controls to account for broader temporal influences.

Interpretation and Limitations

Model estimates describe associations between observed characteristics and the likelihood of exiting a STEM pathway in a given term and should not be interpreted causally. Predicted probabilities and survival curves derived from the models are used to illustrate how attrition risk evolves over time and varies across subgroups and institutional contexts. As with all analyses based on administrative data, findings are subject to limitations related to measurement, coding consistency, and unobserved factors. As such, the models are intended to characterize patterns of persistence and attrition rather than to estimate causal effects of specific policies or interventions.

Table A1. Results from Discrete-Time Logistic Regression Predicting STEM Attrition

covariate	Estimate	Std. Error	z value	Pr>|z| -	exp_coef	sig
Intercept -	-1.307936036	0.2435475142	-5.370352641	7.86E-08	0.2703775303	***
school_type - 4 (4-year)	-0.05957277797	0.07742187907	-0.7694566276	0.4416222829	0.9421669621	
school_type - C (Liberal Arts)	-0.3208440517	0.08831966657	-3.632758866	2.80E-04	0.7255363883	***
school_type - T (Technical)	-0.03304620843	0.08102238548	-0.4078651626	0.6833726703	0.9674938522	
tuition_status – 0 (In-state)	-1.307244635	0.2067558664	-6.322648339	2.57E-10	0.2705645342	***
tuition_status – 1 (Out-of-state)	-0.8350827328	0.2051512362	-4.070571293	4.69E-05	0.4338385873	***
tuition_status – 2 (In-district)	-1.040164699	0.2042462149	-5.092700001	3.53E-07	0.3533964731	***
tuition_status – 3 (Out-of-district)	-1.032515957	0.2039641677	-5.062241904	4.14E-07	0.3561098756	***
tuition_waiver - Y	-0.08255468891	0.02913654035	-2.833373074	0.004605959544	0.9207610811	***
census_region - NA	-0.126628701	0.0979129707	-1.293278103	0.195914891	0.8810607488	
census_region - Northeast	0.01802176366	0.1343604557	0.1341299683	0.8932998025	1.018185136	
census_region - South	-0.09061235297	0.03272904214	-2.768561102	0.005630442694	0.9133717081	***
census_region - West	-0.01914720133	0.07494595323	-0.2554801227	0.7983522941	0.981034942	
geo_county_agg - 2	0.007957369546	0.06423473688	0.1238795383	0.9014106542	1.007989114	
geo_county_agg - 3	0.1504556599	0.05930498712	2.53698158	0.01118128371	1.162363765	**
geo_county_agg - 4	0.1855189632	0.06839964627	2.71227957	0.00668222083	1.203843028	***
geo_county_agg - 6	0.1363913868	0.04202481637	3.245496318	0.001172460807	1.146130386	***
geo_county_agg - 7	0.05571716845	0.04160738621	1.339117246	0.1805325068	1.057298604	
geo_county_agg - 8	0.09284232445	0.04660479098	1.992119748	0.04635792073	1.097288706	**
geo_county_agg - 9	0.05032147704	0.04139689369	1.215585822	0.2241427378	1.05160911	
geo_county_agg - NA	0.09115678504	0.03668800219	2.484648375	0.01296794542	1.095440741	**
enroll_status – 02 (Other first-year continuing student)	0.2040841132	0.0402481506	5.070645737	3.96E-07	1.226401306	***

enroll_status – 03 (First-time entering undergraduate transfer student)	-0.155278018	0.04360037623	-3.561391699	3.69E-04	0.8561771117	***
enroll_status – 04 (First-time entering graduate student)	1.206670645	1.341020069	0.8998155013	0.3682184436	3.342338279	
enroll_status – 05 (First-time entering doctoral student)	-1.41896089	1.689148745	-0.8400449598	0.4008831786	0.241965315	
enroll_status - 06 (Continuing undergraduate student)	-0.06974249069	0.04175101564	-1.670438182	0.09483270052	0.9326339509	*
enroll_status – 07 (Continuing graduate or doctoral student)	0.2141132842	1.344585732	0.1592410801	0.8734789395	1.238762979	
enroll_status – 08 (Readmitted undergraduate student)	0.1250000838	0.04511986611	2.770400149	0.005598746353	1.133148548	***
enroll_status – 09 (Unclassified pre-baccalaureate)	-0.1527209336	0.0934379924	-1.634462917	0.1021616866	0.8583692304	
enroll_status – 10 (Unclassified post-baccalaureate)	-0.4282044903	0.1338722109	-3.198606249	0.001380936367	0.6516781393	***
enroll_status – 11 (First-time entering graduate transfer)	-14.64095197	1496.174589	-0.009785590587	0.992192353	4.38E-07	
enroll_status – 12 (Readmitted graduate student)	-0.1738723029	1.415322838	-0.1228499239	0.9022259401	0.840404208	
enroll_status – 14 (Undergraduate transfer transient student)	-0.6169564517	0.113157711	-5.452182148	4.98E-08	0.5395841915	***
enroll_status – 15 (Unclassified post-baccalaureate graduate student)	0.7588056342	1.491693664	0.5086873078	0.6109714203	2.135723861	
student_level – 01 (Freshman)	0.07664963684	0.05896538257	1.299909091	0.193632129	1.079663736	
student_level – 02 (Sophomore)	-0.1674282371	0.05927084114	-2.824799411	0.004731022338	0.8458373148	***
student_level – 03 (Junior)	-0.9003966903	0.06321296453	-14.24386116	4.89E-46	0.4064084095	***
student_level – 04 (Senior)	-0.4344199024	0.06205587127	-7.000464155	2.55E-12	0.6476402527	***
student_level – 05 (Master’s level)	-0.9151910173	1.347858506	-0.6789963585	0.4971401692	0.4004401278	
student_level – 06 (Post-Bacc Certificate)	-2.221381776	1.46947607	-1.511682852	0.1306145631	0.108459139	
student_level – 07 (Doctoral)	-3.740082379	1.680440059	-2.225656523	0.02603719703	0.02375214637	**

student_level – 09 (Doctoral, Professional)	-2.120876807	1.367636849	-1.550760209	0.1209591596	0.1199264301	
student_level – 10 (Unclassified Post-Bacc)	0.1574415083	0.1198725379	1.31340765	0.1890456212	1.170512291	
student_level – 11 (Unclassified Graduate)	-1.456122004	1.374598141	-1.059307416	0.2894597988	0.2331386346	
degree_intent – 3 (Non-award or non-degree seeking)	-0.0712145201	0.0645154147	-1.103837283	0.2696637285	0.9312620963	
degree_intent – 4 (Degree-seeking, complete associate level program)	0.4935972883	0.02616572483	18.86426963	2.24E-79	1.638198706	***
degree_intent – 6 (Degree-seeking, complete baccalaureate level (or higher) program at another institution after transfer)	0.3554581671	0.05016683574	7.085520979	1.39E-12	1.426834233	***
degree_intent – 7 (Degree-seeking, complete Certificate of Proficiency)	1.469468158	0.04753634884	30.9125163	8.11E-210	4.346922652	***
degree_intent – 8 (Degree-seeking, complete Technical Certificate)	0.7058946635	0.03157736141	22.3544537	1.09E-110	2.025658157	***
degree_intent – 9 (Degree-seeking, but not an undergraduate student)	-1.136537397	0.2074577259	-5.478404781	4.29E-08	0.3209283475	***
attend_status – NT (Missing)	-16.98436347	23.49005098	-0.7230449814	0.4696522494	4.21E-08	
attend_status – PT	0.5640782283	0.0176171382	32.01872075	5.99E-225	1.757826721	***
totalcrhours	0.01732783832	0.00183539756	9.44091825	3.70E-21	1.017478836	***
diploma_ged - 1 (Student has GED from Arkansas)	0.5095789792	0.05070684006	10.04951163	9.23E-24	1.66459022	***
diploma_ged – 2 (Student does not have diploma or GED (student is still in high school or is high school age))	-0.005373427202	0.341520826	-0.01573382	0.9874467459	0.9946409838	
diploma_ged – 3 (Student does not have diploma or GED)	-0.1211690578	0.1096501571	-1.105051383	0.2691373251	0.8858841813	
diploma_ged – 4 (Student was home-schooled and has the equivalent of a high school diploma/GED)	-0.1899108879	0.1420708261	-1.336733889	0.1813095158	0.8270328293	

diploma_ged – 5 (Student has GED from a non-Arkansas state)	0.144695787	0.0855692778	1.690978242	0.09084095822	1.155687941	*
diploma_ged – 9 (Not Applicable, student is not a first-time entering undergraduate or a high school student)	-0.07500228468	0.03194622257	-2.347766924	0.01888633542	0.9277413667	**
gender - M	0.1792212953	0.01184203829	15.13432831	9.62E-52	1.196285447	***
gender - U	0.5833442466	0.4749982588	1.228097653	0.2194103121	1.792021382	
non_us_resident - Y	-0.2450952644	0.09716364216	-2.522499764	0.01165240283	0.7826299779	**
hispanic - U	0.1145232058	0.07564149954	1.514026116	0.1300192158	1.121338662	
hispanic - Y	-0.01286511028	0.0309219862	-0.4160505797	0.677372992	0.9872172915	
asian - U	0.2032122461	0.2268267363	0.8958919454	0.3703104764	1.225332513	
asian - Y	-0.08467261731	0.03926429655	-2.156478652	0.03104630894	0.9188130387	**
black - U	0.1850574022	0.1422656102	1.300788026	0.1933310218	1.20328751	
black - Y	0.1834380483	0.02847980241	6.440987392	1.19E-10	1.201340538	***
amerind - U	-0.08692047927	0.4040919361	-0.2151007518	0.8296887678	0.9167499934	
amerind - Y	0.1118207378	0.03361882961	3.326134166	8.81E-04	1.118312372	***
white - U	0.08040028571	0.0732754906	1.097232991	0.2725395549	1.083720779	
white - Y	0.06526450722	0.02690470894	2.42576522	0.0152761492	1.067441333	**
hawaiian - NA	-0.3515182254	0.4518341745	-0.7779806072	0.4365804459	0.7036190261	
hawaiian - Y	-0.0519788123	0.09553704825	-0.5440696908	0.5863935165	0.9493489811	
cip_code - 01	0.5072279663	0.09939100573	5.103358826	3.34E-07	1.660681344	***
cip_code - 02	-0.0933853714	0.1274170014	-0.7329113883	0.4636124767	0.91084242	
cip_code - 03	-0.09552560417	0.06801944785	-1.404386645	0.1602037455	0.9088950899	
cip_code - 04	0.1019545514	0.3378832444	0.3017449167	0.7628465247	1.107333144	
cip_code - 09	0.3614365737	0.06582146316	5.491165896	3.99E-08	1.435389978	***
cip_code - 10	0.7579731497	0.4959295507	1.528388757	0.1264160445	2.133946644	
cip_code - 11	-0.1108464382	0.04647765382	-2.384940484	0.01708188462	0.8950761879	**
cip_code - 12	0.6558472109	0.1664568249	3.940043981	8.15E-05	1.92677421	***
cip_code - 13	0.2360676834	0.0535363077	4.409487571	1.04E-05	1.266260012	***
cip_code - 14	-0.600225219	0.05132729219	-11.69407529	1.37E-31	0.5486880472	***
cip_code - 15	-0.09357229566	0.05165929793	-1.811335024	0.07008900722	0.9106721774	*

cip_code - 16	1.117430385	0.1365104423	8.185676978	2.71E-16	3.056988817	***
cip_code - 19	0.801773008	0.09079513816	8.830572036	1.04E-18	2.22949033	***
cip_code - 20	0.07024760763	0.1608861688	0.4366292525	0.6623802445	1.072773775	
cip_code - 22	0.1298555936	0.2664023059	0.4874417027	0.6259453465	1.138663941	
cip_code - 23	0.9134925736	0.08929209117	10.23038616	1.45E-24	2.493014382	***
cip_code - 24	0.2213302737	0.048259501	4.586252843	4.51E-06	1.247735457	***
cip_code - 26	-0.4442421571	0.04731844107	-9.388351499	6.09E-21	0.6413101042	***
cip_code - 27	-0.2328905969	0.06815602202	-3.417021563	6.33E-04	0.7922402423	***
cip_code - 29	0.6734838354	0.2688642117	2.504921838	0.01224784563	1.961057435	**
cip_code - 30	0.3821578795	0.06158508163	6.205364503	5.46E-10	1.46544343	***
cip_code - 31	-0.1471708858	0.05873650815	-2.505611763	0.0122239759	0.8631464654	**
cip_code - 38	1.225142653	0.2113475633	5.796814661	6.76E-09	3.404651733	***
cip_code - 40	-0.4592682911	0.05241177989	-8.76269213	1.91E-18	0.6317457304	***
cip_code - 41	-0.0086688794	0.2081675016	-0.04164376923	0.9667826807	0.991368587	
cip_code - 42	0.7876256069	0.06065263164	12.9858439	1.47E-38	2.198170903	***
cip_code - 43	0.2506101163	0.04590590953	5.459212525	4.78E-08	1.284809061	***
cip_code - 44	0.676364744	0.08856419166	7.637000139	2.22E-14	1.966715208	***
cip_code - 45	0.4864997316	0.05634878282	8.633722101	5.94E-18	1.626612663	***
cip_code - 46	0.4146817645	0.4415930096	0.9390587159	0.347700599	1.513888891	
cip_code - 47	0.2222903401	0.09705878836	2.29026494	0.02200596293	1.248933941	**
cip_code - 48	0.5522853685	0.1333770124	4.140783773	3.46E-05	1.73721867	***
cip_code - 49	-0.02959794532	0.07436488204	-0.3980097125	0.6906230264	0.9708357842	
cip_code - 50	0.4522593202	0.07980567213	5.66700722	1.45E-08	1.571859511	***
cip_code - 51	-0.04258086128	0.04463494168	-0.9539804394	0.3400935438	0.958312972	
cip_code - 52	0.332015785	0.04690426449	7.078584188	1.46E-12	1.393774849	***
cip_code - 54	1.069731782	0.2166118561	4.938472904	7.87E-07	2.914597648	***
ar_res1 - 1	0.2180751708	0.01783582489	12.226806	2.24E-34	1.243680553	***
crdiff	0.02622838285	0.00280035482	9.366092703	7.53E-21	1.026575374	***
fc_start - 001086	-0.1381333834	0.04435421799	-3.114323499	0.001843672257	0.8709825094	***
fc_start - 001089	-0.04448140008	0.03411718961	-1.303782656	0.1923076979	0.9564933906	
fc_start - 001090	-0.1770409986	0.03524048494	-5.023795754	5.07E-07	0.8377454374	***

fc_start - 001091	0.1366321433	0.07920759835	1.724987832	0.08452966543	1.146406358	*
fc_start - 001092	-0.286228737	0.03814092593	-7.504504152	6.17E-14	0.751090794	***
fc_start - 001098	0.02304306423	0.04119214064	0.5594043881	0.5758857673	1.023310607	
fc_start - 001101	-0.1666222836	0.03851039193	-4.326683663	1.51E-05	0.8465192952	***
fc_start - 001104	0.2572518033	0.08178552607	3.145444135	0.001658348631	1.293370761	***
fc_start - 001107	-0.172334082	0.0441828888	-3.900471125	9.60E-05	0.8416979301	***
fc_start - 001108	-0.3597603595	0.03528169771	-10.19679842	2.05E-24	0.6978435376	***
fc_start - 001109	0.2922125555	0.06142295477	4.757383565	1.96E-06	1.339387682	***
fc_start - 001110	0.04798390208	0.03817167301	1.257055253	0.2087336287	1.049153766	
fc_start - 005245	0.5743167989	0.07017767661	8.183753391	2.75E-16	1.775916804	***
fc_start - 005707	0.1668809502	0.05836835466	2.859099785	0.00424845123	1.181613586	***
fc_start - 005732	0.5379383074	0.05462900809	9.847118339	7.05E-23	1.712472628	***
fc_start - 007738	0.9807800669	0.0582722729	16.83099042	1.45E-63	2.666535507	***
fc_start - 009976	0.479026446	0.07078982649	6.766882612	1.32E-11	1.614501832	***
fc_start - 012105	0.3390904569	0.04707402311	7.203345593	5.88E-13	1.403670311	***
fc_start - 012260	0.3991344732	0.05319069072	7.503840763	6.20E-14	1.490534042	***
fc_start - 012261	0.3470220154	0.05015314892	6.919246803	4.54E-12	1.414847873	***
fc_start - 012860	0.6511789187	0.07129783328	9.133221709	6.65E-20	1.917800428	***
fc_start - 020522	0.3535142988	0.05573446947	6.342830606	2.26E-10	1.424063349	***
fc_start - 020735	0.6298855669	0.080827459	7.792965098	6.55E-15	1.877395731	***
fc_start - 020746	0.3412192762	0.05147141615	6.629296446	3.37E-11	1.406661654	***
fc_start - 020753	-0.2557977977	0.05326267674	-4.802571207	1.57E-06	0.7742985179	***
fc_start - 020870	0.2173550648	0.06586367234	3.300075096	9.67E-04	1.242785293	***
fc_start - 021111	0.7484644834	0.07776922608	9.624172969	6.32E-22	2.113751823	***
fc_start - 022209	0.4076607648	0.06013129052	6.779511321	1.21E-11	1.503297103	***
fc_start - 023482	0.524646707	0.05739388259	9.141160754	6.18E-20	1.689861727	***
fc_start - 030633	0.3815618246	0.04789145275	7.967221762	1.62E-15	1.464570206	***
fc_start - 042034	0.5389979205	0.06708720407	8.034288029	9.41E-16	1.714288148	***
fc_start - 042544	0.2741015998	0.05565548919	4.924969734	8.44E-07	1.315348435	***
cohort - 2011 3	0.07165739503	0.04688383361	1.528403066	0.1264124941	1.074287225	
cohort - 2012 0	-0.03988091488	0.09218885168	-0.4326001914	0.6653052504	0.9609038617	

cohort - 2012 1	0.1204090074	0.02188629069	5.501572153	3.76E-08	1.1279581	***
cohort - 2012 2	0.2392616711	0.03383328513	7.071783606	1.53E-12	1.270310897	***
cohort - 2012 3	0.06900675082	0.0585210625	1.179178024	0.2383272951	1.071443442	
cohort - 2013 0	0.0909168465	0.1211851632	0.7502308374	0.4531156892	1.095177934	
cohort - 2013 1	0.07391096653	0.02350110026	3.145000265	0.001660866796	1.076710938	***
cohort - 2013 2	0.204628382	0.03551500321	5.761744712	8.32E-09	1.227068979	***
cohort - 2013 3	3.47E-04	0.06297694383	0.005514290974	0.9956002547	1.000347333	
cohort - 2014 0	-0.06418385235	0.1237239387	-0.5187666432	0.6039234818	0.937832561	
cohort - 2014 1	0.0581628878	0.02405427793	2.417985191	0.01560671198	1.059887624	**
cohort - 2014 2	0.2368785305	0.03946809725	6.001772241	1.95E-09	1.267287172	***
cohort - 2014 3	-0.07601417687	0.07180918995	-1.058557782	0.2898012223	0.9268030673	
cohort - 2015 0	0.04402548551	0.098471761	0.4470874195	0.6548119421	1.045008987	
cohort - 2015 1	6.53E-04	0.02426104223	0.0269142704	0.9785281115	1.000653181	
cohort - 2015 2	0.2148351362	0.04337645801	4.952804956	7.32E-07	1.239657506	***
cohort - 2015 3	-0.04956251245	0.0696826135	-0.7112608147	0.4769226291	0.9516456666	
cohort - 2016 0	0.06326586409	0.1058333786	0.5977874366	0.5499817746	1.065310029	
cohort - 2016 1	0.0305349545	0.02474364481	1.23405241	0.2171833754	1.031005928	
cohort - 2016 2	0.2663805651	0.04559533075	5.842277283	5.15E-09	1.30523169	***
cohort - 2016 3	-0.07035178404	0.07416645575	-0.9485660779	0.3428413507	0.9320658763	
cohort - 2017 0	-0.04415769904	0.09687994394	-0.4557981482	0.6485351359	0.9568030586	
cohort - 2017 1	0.06411483085	0.02436666015	2.631252311	0.008507085241	1.066214826	***
cohort - 2017 2	0.2192638493	0.04609401577	4.756883202	1.97E-06	1.245159768	***
cohort - 2017 3	-0.1064346351	0.08044237175	-1.323116572	0.1857966145	0.8990338115	
cohort - 2018 0	-0.02385654021	0.1012807709	-0.2355485647	0.8137829898	0.9764257775	
cohort - 2018 1	0.06723764763	0.02505036257	2.68409878	0.007272561535	1.069549624	***
cohort - 2018 2	0.2402680612	0.04643303118	5.174507352	2.29E-07	1.271589969	***
cohort - 2018 3	0.07060438083	0.0726027654	0.9724750902	0.3308142513	1.073156581	
age_group - Age 20-30	0.352153331	0.0207321996	16.98581616	1.05E-64	1.422126563	***
age_group - Age 30-40	0.3194729691	0.02533384522	12.61052029	1.85E-36	1.376402167	***
age_group - Age 40+	0.2770758196	0.02707845669	10.23233424	1.42E-24	1.319266393	***
first_rec_type - Non-FTIC undegraduate	0.1934045008	0.01578698988	12.25087887	1.66E-34	1.213373505	***

elapsed - 1	-0.2190860481	0.02159858634	-10.14353646	3.54E-24	0.8032525968	***
elapsed - 2	-1.083487701	0.03502568865	-30.93408702	4.16E-210	0.3384131811	***
elapsed - 3	-0.5246333648	0.02517359174	-20.84062419	1.85E-96	0.5917722892	***
elapsed - 4	-0.4519912986	0.02299720012	-19.65418817	5.32E-86	0.6363597069	***
elapsed - 5	-0.3620704886	0.02717009553	-13.32606609	1.63E-40	0.6962332896	***
elapsed - 6	-0.8048244627	0.04174965307	-19.27739283	8.32E-83	0.447166414	***
elapsed - 7	-0.6074501672	0.0325698099	-18.6507127	1.25E-77	0.5447380908	***
elapsed - 8	-0.3213581205	0.02779609518	-11.56126853	6.47E-31	0.7251635086	***
elapsed - 9	-0.2299429349	0.03236825659	-7.103964163	1.21E-12	0.794578944	***
elapsed - 10	-0.9048869365	0.0535515942	-16.89747896	4.70E-64	0.4045876266	***
elapsed - 11	-0.5724776768	0.04102680389	-13.95374786	2.98E-44	0.5641259838	***
elapsed - 12	-0.1438405184	0.03236948176	-4.443707794	8.84E-06	0.8660258523	***
elapsed - 13	0.5891058754	0.03294390988	17.88208739	1.63E-71	1.802376146	***
elapsed - 14	0.3099686882	0.0496064501	6.248556137	4.14E-10	1.363382423	***
elapsed - 15	-0.2032360847	0.04953365125	-4.102990181	4.08E-05	0.8160855534	***
elapsed - 16	-0.6534268644	0.05022834493	-13.00912593	1.09E-38	0.5202598585	***
elapsed - 17	-0.178024241	0.05512181985	-3.229651007	0.001239414093	0.8369221354	***
elapsed - 18	-0.9027105519	0.1095564995	-8.239680492	1.73E-16	0.4054691238	***
elapsed - 19	-0.2209919769	0.07278858387	-3.036080181	0.002396756189	0.8017231126	***

* = p<0.1, ** = p<0.05, *** = p<0.01

Table A2. Results from Discrete-Time Logistic Regression Predicting STEM Attrition, including K-12 Covariates

covariate	Estimate	Std. Error	z value	Pr>|z| -	exp_coef	sig
Intercept -	0.5841920822	0.9897775919	0.5902256092	5.55E-01	1.793541366	
cte - 1	0.09974989501	0.04078225482	2.445914171	0.01444853914	1.104894544	**
college_stem - 1	0.0674944425	0.0217557865	3.102367385	1.92E-03	1.069824314	***
gpa_grade_point_average	-0.7079607208	0.01523672646	-46.46409599	0	0.4926478202	***
school_type - 4 (4-year)	-0.2523432817	0.06622565835	-3.8103552	1.39E-04	0.77697797	***
school_type - C (Liberal Arts)	0.1042614238	0.2148164566	0.4853511947	6.27E-01	1.109890569	
school_type - T (Technical)	0.01556726557	0.06317935669	0.2463979754	8.05E-01	1.015689067	
tuition_status – 0 (In-state)	-0.5805124981	0.9524314262	-0.6095058207	5.42E-01	0.5596114932	

tuition_status – 1 (Out-of-state)	-0.2332044094	0.945732149	-0.2465861075	0.8052285572	0.7919916665	
tuition_status – 2 (In-district)	-0.5895920882	0.9299321867	-0.6340162182	0.5260702379	0.5545534475	
tuition_status – 3 (Out-of-district)	-0.650653306	0.9301916933	-0.6994830321	0.4842502134	0.5217048325	
tuition_waiver - Y	-0.3145876012	0.1052994189	-2.987553061	0.002812204739	0.7300899004	***
census_region - NA	-0.5169136126	0.3923947714	-1.317330531	0.1877278555	0.5963583032	
census_region - Northeast	-1.798890832	1.061896194	-1.694036425	0.09025836172	0.1654823341	*
census_region - South	-0.5494917835	0.2087209232	-2.632662673	0.008471844234	0.5772431003	***
census_region - West	-0.6712882506	0.4731672235	-1.418712491	0.1559828539	0.5110497933	
geo_county_agg - 2	0.2208255729	0.08871884035	2.489049361	0.01280851811	1.247105883	**
geo_county_agg - 3	0.4128055366	0.08079843102	5.109078622	3.24E-07	1.511051153	***
geo_county_agg - 4	0.2907806544	0.09478356575	3.067838312	0.002156132453	1.337471183	***
geo_county_agg - 6	0.3087629554	0.05367862641	5.752065134	8.82E-09	1.361739539	***
geo_county_agg - 7	0.2182365074	0.05391240512	4.047983148	5.17E-05	1.24388122	***
geo_county_agg - 8	0.328939567	0.06036041226	5.449591126	5.05E-08	1.389493882	***
geo_county_agg - 9	0.3227563094	0.05348652734	6.034347815	1.60E-09	1.380928791	***
geo_county_agg - NA	0.236388405	0.04647165295	5.08672255	3.64E-07	1.266666194	***
enroll_status – 02 (Other first-year continuing student)	0.24847272	0.05965607688	4.165086492	3.11E-05	1.282065847	***
enroll_status – 03 (First-time entering undergraduate transfer student)	-0.01816879533	0.07183487408	-0.2529244404	0.8003266027	0.9819952622	
enroll_status – 04 (First-time entering graduate student)	-0.14832008	2.103666134	-0.07050552252	0.9437913052	0.8621551122	
enroll_status – 05 (First-time entering doctoral student)	-16.91663042	695.9029226	-0.02430889406	0.9806062188	4.50E-08	
enroll_status - 06 (Continuing undergraduate student)	0.02577711439	0.06353752032	0.4056990935	0.6849637084	1.026112217	
enroll_status – 07 (Continuing graduate or doctoral student)	-1.561534651	2.109265411	-0.7403215559	0.4591049043	0.209813833	

enroll_status – 08 (Readmitted undergraduate student)	0.2453086582	0.07477365337	3.280683063	0.001035560382	1.28E+00	***
enroll_status – 09 (Unclassified pre-baccalaureate)	-0.1554055422	0.1758318557	-0.883830416	0.3767877645	0.8560679354	
enroll_status – 10 (Unclassified post-baccalaureate)	0.2893208398	0.4005795029	0.7222557264	4.70E-01	1.335520148	
enroll_status – 11 (First-time entering graduate transfer)	-17.64056603	4734.664858	-0.00372583204	0.997027223	2.18E-08	
enroll_status – 12 (Readmitted graduate student)	-18.54885257	2568.056893	-0.007222913411	0.994236999	8.80E-09	
enroll_status – 14 (Undergraduate transfer transient student)	-0.6597243727	0.1793262027	-3.678906722	0.0002342359117	0.5169938125	***
enroll_status – 15 (Unclassified post-baccalaureate graduate student)	-14.24710424	2047.70034	-0.006957611894	9.94E-01	6.49E-07	
student_level – 01 (Freshman)	0.142048024	0.1294901418	1.096979446	2.73E-01	1.152632001	
student_level – 02 (Sophomore)	-0.3633579315	0.1303742892	-2.787036721	0.005319245103	0.6953375058	***
student_level – 03 (Junior)	-1.20416801	0.1356087803	-8.879720084	6.70E-19	0.299941444	***
student_level – 04 (Senior)	-0.5121929435	0.134390479	-3.811229391	0.0001382773742	0.5991801688	***
student_level – 05 (Master's level)	0.6222797657	2.11239029	0.2945856022	0.7683104621	1.863170797	
student_level – 06 (Post-Bacc Certificate)	-16.31789562	1774.191438	-0.009197370293	0.9926616637	8.19E-08	
student_level – 07 (Doctoral)	-15.87208405	1368.813361	-0.01159550637	0.9907483318	1.28E-07	
student_level – 09 (Doctoral, Professional)	-0.5211883954	2.138516447	-0.2437149343	0.8074516062	0.5938144421	
student_level – 10 (Unclassified Post-Bacc)	-0.268821185	0.396479859	-0.6780197755	4.98E-01	0.7642799081	
student_level – 11 (Unclassified Graduate)	0.9705529175	2.20042134	0.441075943	6.59E-01	2.639403428	

degree_intent – 3 (Non-award or non-degree seeking)	-0.5484445168	0.09936146261	-5.519690455	3.40E-08	0.5778479444	***
degree_intent – 4 (Degree-seeking, complete associate level program)	0.5415198603	0.03883395504	13.94449419	3.40E-44	1.718616936	***
degree_intent – 6 (Degree-seeking, complete baccalaureate level (or higher) program at another institution after transfer)	0.1456931181	0.0670732324	2.172149945	2.98E-02	1.15684112	**
degree_intent – 7 (Degree-seeking, complete Certificate of Proficiency)	1.13465221	0.08381829679	13.53704685	9.45E-42	3.11E+00	***
degree_intent – 8 (Degree-seeking, complete Technical Certificate)	0.6212246123	0.04711405957	13.18554627	1.06E-39	1.861205902	***
degree_intent – 9 (Degree-seeking, but not an undergraduate student)	-0.5033837213	0.3424042129	-1.470144649	1.42E-01	0.6044817974	
attend_status – NT (Missing)	-18.02962124	57.14695054	-0.3154957714	7.52E-01	1.48E-08	
attend_status – PT	0.422397291	0.0328101737	12.87397302	6.31E-38	1.525614517	***
totalcrhours	2.35E-05	0.003339932563	0.00704603565	0.9943781235	1.000023534	
diploma_ged - 1 (Student has GED from Arkansas)	0.2668968587	0.1715596191	1.555709089	0.1197772843	1.305905747	
diploma_ged – 2 (Student does not have diploma or GED (student is still in high school or is high school age))	-0.08259862834	0.4024396133	-0.2052447761	0.8373808877	0.9207206242	
diploma_ged – 3 (Student does not have diploma or GED)	-0.3540265661	0.1680986959	-2.106063728	0.03519881708	0.7018563215	**
diploma_ged – 4 (Student was home-schooled and has the equivalent of a high school diploma/GED)	-0.5478837041	0.4172080504	-1.313214602	1.89E-01	0.5781720998	
diploma_ged – 5 (Student has GED from a non-Arkansas state)	0.2160455997	0.3774965085	0.5723115178	0.5671109481	1.241158974	
diploma_ged – 9 (Not Applicable, student is not a	-0.06636490133	0.04108680362	-1.615236414	0.1062595077	0.9357893312	

first-time entering undergraduate or a high school student)						
gender - M	0.03780173417	0.01781426542	2.121992307	0.03383838407	1.038525308	**
gender - U	0.4616615593	0.5265671982	0.8767381654	0.3806288718	1.586708206	
non_us_resident - Y	-0.9623778177	0.4382858756	-2.195776481	0.02810794582	0.3819835181	**
hispanic - U	0.202565476	0.1348535105	1.502114963	0.1330674211	1.224540261	
hispanic - Y	-0.04979045516	0.04785720515	-1.040396216	0.2981558589	0.9514287706	
asian - U	0.1557149609	0.5061653678	0.307636537	7.58E-01	1.168493089	
asian - Y	-0.2994490971	0.06626752804	-4.518790816	6.22E-06	0.741226452	***
black - U	0.6112572013	0.3008599372	2.031700223	4.22E-02	1.842746647	**
black - Y	0.06209978984	0.04633296276	1.340293954	0.1801497965	1.064068523	
amerind - U	-0.8623657422	0.6230839959	-1.384028073	0.166349859	0.4221621732	
amerind - Y	0.1455170853	0.05468098122	2.661201061	0.007786244676	1.156637496	***
white - U	0.1115431074	0.1557034431	0.7163817655	0.4737556511	1.118001937	
white - Y	0.1584597368	0.04437340099	3.571052326	3.56E-04	1.171704747	***
hawaiian - NA	-0.07597108177	0.6571203003	-0.1156121364	0.907959944	0.9268430088	
hawaiian - Y	0.08077139792	0.1546704164	0.5222162053	0.6015198041	1.084123035	
ar_res1 - 1	0.1514225888	0.0237970837	6.363073335	1.98E-10	1.163488231	***
crdiff	0.0352221627	0.004543593928	7.752048985	9.04E-15	1.03584981	***
cohort - 2011 3	-0.4463812507	0.1277289313	-3.494754448	0.0004744980807	0.6399397481	***
cohort - 2012 0	-0.1641995791	0.2032102422	-0.8080280666	0.4190744253	0.8485726476	
cohort - 2012 1	-0.09593237964	0.04477948822	-2.14232863	3.22E-02	0.9085254488	**
cohort - 2012 2	0.6383545429	0.09942534905	6.420440552	1.36E-10	1.893362867	***
cohort - 2012 3	0.02193007456	0.1371811605	0.1598621449	8.73E-01	1.022172306	
cohort - 2013 0	-0.2612414865	0.2349307168	-1.111993741	0.2661408417	0.7700949296	
cohort - 2013 1	-0.05270520463	0.04511071781	-1.168352161	2.43E-01	0.9486596317	
cohort - 2013 2	0.3894549752	0.1015110744	3.836576231	1.25E-04	1.476176022	***
cohort - 2013 3	0.01566437582	0.1453635319	0.1077600112	0.9141860646	1.015787705	
cohort - 2014 0	-0.3005686321	0.2605612345	-1.153543169	0.2486875095	0.7403970874	
cohort - 2014 1	-0.07604256946	0.04466071085	-1.702672618	8.86E-02	0.9267767533	*
cohort - 2014 2	0.3573168367	0.0964722403	3.703830611	2.12E-04	1.429488713	***

cohort - 2014 3	-0.1273422826	0.1480919824	-0.8598864065	3.90E-01	0.8804322644	
cohort - 2015 0	-0.4162335957	0.1887945432	-2.204690817	2.75E-02	0.65952619	**
cohort - 2015 1	-0.06122382481	0.04360558982	-1.404036158	0.1603080826	0.9406126838	
cohort - 2015 2	0.4287751198	0.1047089908	4.094921711	4.22E-05	1.53537572	***
cohort - 2015 3	-0.1580527877	0.1295997083	-1.21954586	0.2226370809	0.8538047104	
cohort - 2016 0	-0.2736572229	0.1672162144	-1.636547173	1.02E-01	0.7605927444	
cohort - 2016 1	-0.007979607806	0.0439902515	-0.1813949122	8.56E-01	0.9920521448	
cohort - 2016 2	0.3504237127	0.1112471269	3.149957417	0.001632942615	1.419668953	***
cohort - 2016 3	-0.3565321783	0.1518003454	-2.348691483	1.88E-02	0.7000999431	**
cohort - 2017 0	0.1423848463	0.1576676734	0.9030693683	3.66E-01	1.153020299	
cohort - 2017 1	0.007367930651	0.04456748232	0.1653207735	8.69E-01	1.007395141	
cohort - 2017 2	0.1088271585	0.0985697215	1.104062758	2.70E-01	1.114969621	
cohort - 2017 3	-0.1720486589	0.1274986341	-1.349415703	0.1772034799	0.8419382044	
cohort - 2018 0	-0.007881876021	0.1557249808	-0.05061407605	0.9596330461	0.9921491045	
cohort - 2018 1	0.07796161697	0.04619645755	1.687610287	9.15E-02	1.081081163	*
cohort - 2018 2	0.3485555809	0.09149311678	3.809637196	0.0001391708628	1.4170193	***
cohort - 2018 3	0.2041643338	0.1122044308	1.819574614	6.88E-02	1.226499692	*
age_group - Age 20-30	0.3068467387	0.02879328425	10.65688568	1.62E-26	1.35913265	***
age_group - Age 30-40	-0.203545672	0.3625049211	-0.5614976795	5.74E-01	0.8158329427	
age_group - Age 40+	0.207286007	0.4507366946	0.4598826976	6.46E-01	1.230334406	
first_rec_type - Non-FTIC undegraduate	0.01032161974	0.03168699272	0.3257368041	7.45E-01	1.010375071	
elapsed - 1	0.0566257999	0.04355609272	1.300066107	1.94E-01	1.058259735	
elapsed - 2	-1.285577705	0.07734304353	-16.62176255	4.85E-62	0.2764908075	***
elapsed - 3	-0.4596531184	0.05703004863	-8.059840899	7.64E-16	0.6315026642	***
elapsed - 4	-0.1788561044	0.04777626385	-3.743618484	1.81E-04	0.83622622	***
elapsed - 5	0.08463708966	0.05293098989	1.599008253	0.1098187689	1.088322032	
elapsed - 6	-0.5976626281	0.08118528781	-7.361711022	1.82E-13	0.5500959133	***
elapsed - 7	-0.4125896494	0.07233579136	-5.703810542	1.17E-08	0.6619338521	***
elapsed - 8	0.06991761733	0.05627456515	1.242437274	2.14E-01	1.072419829	
elapsed - 9	0.2841564022	0.06017290295	4.722331619	2.33E-06	1.328640717	***

elapsed - 10	-0.5877692003	0.09258184744	-6.348644108	2.17E-10	0.5555652582	***
elapsed - 11	-0.3303345228	0.08358741413	-3.951964853	7.75E-05	0.7186832773	***
elapsed - 12	0.3609392387	0.06177891739	5.842433858	5.14E-09	1.434676285	***
elapsed - 13	1.273498337	0.06112589397	20.83402392	2.13E-96	3.57333144	***
elapsed - 14	0.954502263	0.07914227418	12.06058675	1.71E-33	2.59737745	***
elapsed - 15	0.3582253465	0.08703788553	4.115740454	3.86E-05	1.430788007	***
elapsed - 16	0.1171129723	0.07778555415	1.505587684	1.32E-01	1.124246431	
elapsed - 17	0.515484775	0.08251435146	6.247213556	4.18E-10	1.674450037	***
elapsed - 18	-0.1603172171	0.141979748	-1.129155526	2.59E-01	0.8518735173	
elapsed - 19	0.6888658505	0.1111622748	6.196939128	5.76E-10	1.991455644	***

Appendix B. Identification of STEM-Related Industries Using NAICS Codes

To characterize early labor market outcomes associated with postsecondary STEM pathways, employer industry classifications from Unemployment Insurance wage records were used to identify employment in industries commonly associated with STEM production and innovation. Industries were classified using the North American Industry Classification System (NAICS) at the six-digit level.

The NAICS system categorizes industries into broad sectors. STEM-related industries are found across several sectors, and there is no single, systematic way to isolate STEM industries. Therefore, we selected industries that were representative of the activities and industries examined in past research on the STEM workforce.

Consistent with prior research on STEM-intensive industries (for example Roberts and Wolf, 2018; Chow and Goldschlag, 2023), we selected industries engaged in software development, engineering services, scientific research, telecommunications, and advanced technology manufacturing.

Because NAICS codes classify industries rather than occupations, this approach identifies employment in industries that commonly employ STEM workers rather than identifying specific STEM occupations.

Note that, while we did include medical and healthcare majors in our postsecondary analysis, we specifically excluded any medical industries from the industry analysis. We did this because there is some debate as to whether medical occupations should be included when studying the

STEM workforce (Xue and Larson, 2015; Bhatti, 2025). This will likely lead to some underrepresentation in workforce outcomes for students with healthcare-related degrees, but we wanted to be sure to examine industries that have a consensus in the literature as being STEM industries.

Table B.1

NAICS 2 digit	NAICS 6 digit	Sector_title	Industry_Description	STEM_Domain
32	325412	Manufacturing	Pharmaceutical Preparation Manufacturing	Advanced Manufacturing
32	325414	Manufacturing	Biological Product Manufacturing	Advanced Manufacturing
33	334111	Manufacturing	Electronic Computer Manufacturing	Advanced Manufacturing
33	334220	Manufacturing	Communications Equipment Manufacturing	Advanced Manufacturing
33	334413	Manufacturing	Semiconductor and Related Device Manufacturing	Advanced Manufacturing
33	334511	Manufacturing	Search, Detection, Navigation, and Guidance Instruments Manufacturing	Advanced Manufacturing
33	334512	Manufacturing	Automatic Environmental Control Manufacturing	Advanced Manufacturing
33	336411	Manufacturing	Aircraft Manufacturing	Aerospace Manufacturing
33	339112	Manufacturing	Surgical and Medical Instrument Manufacturing	Surgical and Medical Supplies Manufacturing
33	339113	Manufacturing	Surgical and Medical Instrument Manufacturing	Surgical and Medical Supplies Manufacturing
33	339114	Manufacturing	Dental Equipment and Supplies Manufacturing	Surgical and Medical Supplies Manufacturing
33	339115	Manufacturing	Ophthalmic Goods Manufacturing	Surgical and Medical Supplies Manufacturing
33	339116	Manufacturing	Dental Laboratories	Surgical and Medical Supplies Manufacturing
51	511210	Information	Software Publishers	Information Technology
51	517311	Information	Wired Telecommunications Carriers	Telecommunications
51	517312	Information	Wireless Telecommunications Carriers	Telecommunications
51	517919	Information	Other Telecommunications	Telecommunications
51	518210	Information	Data Processing, Hosting, and Related Services	Information Technology
51	519130	Information	Internet Publishing and Broadcasting	Information Technology
54	541330	Professional, Scientific, and Technical Services	Engineering Services	Engineering
54	541380	Professional, Scientific, and Technical Services	Testing Laboratories	Engineering / Scientific Services
54	541511	Professional,	Custom Computer Programming Services	Information Technology

		Scientific, and Technical Services		
54	541512	Professional, Scientific, and Technical Services	Computer Systems Design Services	Information Technology
54	541513	Professional, Scientific, and Technical Services	Computer Facilities Management Services	Information Technology
54	541519	Professional, Scientific, and Technical Services	Other Computer Related Services	Information Technology
54	541620	Professional, Scientific, and Technical Services	Environmental Consulting Services	Environmental Consulting
54	541690	Professional, Scientific, and Technical Services	Other Scientific and Technical Consulting Services	Scientific Services
54	541714	Professional, Scientific, and Technical Services	Research and Development in Biotechnology	Research & Development
54	541715	Professional, Scientific, and Technical Services	Research and Development in Physical, Engineering, and Life Sciences	Research & Development
61	611310	Educational Services	Colleges, Universities, and Professional Schools	Colleges, Universities, and Professional Schools
61	611519	Educational Services	Other Technical and Trade Schools	Other Technical and Trade Schools

Appendix C. Data Dictionary

Variable	Definition
academic_year	Academic year beginning July 1 of prior calendar year. e.g., AY2016 begins July 1, 2015.
age	Student's age at the time of the record.
age_group	Categorical grouping of student age used for descriptive analysis. <20, 20-29, 30-39, 40+
amerind	Y/N/U; ethnicity and race identification
asian	Y/N/U; ethnicity and race identification
attend_status	Indicates the student's enrollment intensity for the term. FT = full-time enrollment; PT = part-time enrollment.

black	Y/N/U; ethnicity and race identification
broad_enroll_group	Aggregates detailed enrollment status codes into broader categories describing the student’s stage in postsecondary education. FTIC: 01; Non-FTIC undergraduate: 02,03,06,08,14; Pre-baccalaureate: 09,13,16; Post-baccalaureate: 04,05,07,10,11,12,15
calendar_year	Calendar year corresponding to the academic term record. Four-digit year (e.g., 2015, 2016).
census_region	Indicates the U.S. Census geographic region corresponding to the student’s state of origin. Derived from state using U.S. Census regional classification (Northeast, Midwest, South, West).
cip_code	2-digit
cip_code_name	cip code name
cip_detail	4-digit
cohort	Entry cohort defined by starting academic year and term.
college_stem	Whether the student took a STEM course with an institution of higher learning while in high school.
continuous_attend	Indicator for ‘effective’ continuous enrollment across Spring–Fall transitions (i.e., ignores missing Summer when not substantively meaningful). 1 if attendance is continuous under Spring→Fall continuity rule; 0 otherwise.
country	Student's country of origin; country abbreviations
crdiff	Difference between first and last observed credit hours within a student–institution–academic-year. Used to detect intra-year enrollment changes (e.g., withdrawal within year).
ct_person_term	The number of records for a given student–institution–term combination after applying data cleaning and disambiguation procedures. Counts the number of observations for each combination of: student_id × institution_id × term_val
cte	Whether or not a student took Career and Technical Education Courses in high school

degree_intent	Indicates the type of credential a student intends to pursue at the reporting institution during the given term. 2: Degree-seeking, bachelor's degree at the reporting institution; 3: Non-degree seeking enrollment; 4: Degree-seeking, associate degree; 5: Degree-seeking, certificate of proficiency (< 9 credit hours); 6: Degree-seeking, bachelor's degree (or higher) intended to be completed at another institution after transfer; 7: Degree-seeking, certificate of proficiency (≥ 9 credit hours); 8: Degree-seeking, technical certificate; 9: Degree-seeking, non-undergraduate program; D: Diploma-seeking (specific institutions only); C: High school student seeking a credential (dual enrollment).
degree_level	Indicates the level of a degree awarded to the student. 00, 01: Certificate of Proficiency; 02: Technical Certificate; 03, 04: Associate Degree; 05: Bachelor's Degree; 06: Post-Baccalaureate Certificate; 07: Master's Degree; 08: Master's Certificate; 17: Research Doctoral Degree; 18: Professional Doctoral Degree
degree_name	details of degree
degreestem	Indicator that the awarded credential (when graduation=1) is classified as STEM. 1 if degree_type_stem indicates STEM; 0 otherwise.
diploma_ged	This field identifies high school completion type: 0 = Student has high school diploma (not home-school credential) 1 = Student has GED from Arkansas 2 = Student does not have diploma or GED (student is still in high school or is high school age) 3 = Student does not have diploma or GED (student is not in high school or is not high school age) 4 = Student was home-schooled and has the equivalent of a high school diploma/GED 5 = Student has GED from a non-Arkansas state 9 = Not Applicable, student is not a first-time entering undergraduate or a high school student
early_stem_flag	Indicator equal to 1 if the student is enrolled in a STEM major during one of the first four observed terms (time_seq) following entry into the dataset (approximately the first academic year). Otherwise 0.
elapsed	Number of terms observed since the student's first analytical record (duration in discrete time).
enroll_status	IPEDS enrollment classification for the term. 01 = First-time entering undergraduate 02 = Other first-year continuing student 03 = First-time entering undergraduate transfer student

	04 = First-time entering graduate student 05 = First-time entering doctoral student 06 = Continuing undergraduate student 07 = Continuing graduate or doctoral student 08 = Readmitted undergraduate student 09 = Unclassified pre-baccalaureate 10 = Unclassified post-baccalaureate 11 = First-time entering graduate transfer 12 = Readmitted graduate student 13 = High school student 14 = Undergraduate transfer TRANSIENT student 15 = Unclassified post-baccalaureate graduate student
ever_grad_per_student	Indicates whether a student has received a degree from any institution associated with the student in the dataset. 1: The student has at least one recorded degree completion at any institution in the dataset; 0: No degree completion is recorded for the student.
ever_stem	Indicates whether student ever declared STEM major.
ever_stem_first_year	Student-level indicator equal to 1 if the student was enrolled in a STEM major during any of the first four observed terms in the dataset; Otherwise 0.
exit_full	Discrete-time event indicator: equals 1 in the term a student is classified as exiting STEM without a STEM credential. GAP_CUTOFF=4. exit=1 if (graduation=1 and degreestem=0) OR ((continuous_attend!=1) and (degreestem=0) and (fwd_gap is not NA) and (fwd_gap>=4)). End-of-file without exit is treated as right-censoring (not an exit).
exited	Cumulative indicator of whether an exit has occurred up through the current record (within student sequence).
fc_start	Identifies the institution where the student first enrolled in postsecondary education within the dataset. Note that high school or pre-college enrollment records are excluded when determining the student's starting institution.
fice_code	Federal institution code for reporting institution. IPEDS FICE
fwd_gap	Number of term units between the current record and the next observed record within a student-institution sequence. For record t: next_time_seq - time_seq (mapped to term units); NA on final record.

gap	Difference in discrete seasonal term index between the current observation and the previous observed term for the same student (within institution if applicable); This is not an enrollment indicator. It is a calendar continuity indicator. Computed from ordered term sequence; used to characterize discontinuities (e.g., missing terms).
gender	Indicates the student's reported gender. M: Mael; F: Female; O: Other / non-binary / not exclusively male or female; U: Unknown or not reported.
geo_county	Indicates the student's county of origin when the student's home state is Arkansas. Three-digit Federal Information Processing Standard (FIPS) county codes ranging from 001–075, corresponding to counties in Arkansas. This variable is populated only when geo_state = 'AR'. For students from other states, the field is typically missing. Only Arkansas residents.
geo_county_agg	This is an aggregate of geo_county where counties were grouped based on population and location. Below is a list of counties found in each group. 1=035 2=007, 045, 053, 3=031, 033, 051 4=055, 069 6=015, 067, 023, 057, 059, 001, 075, 037, 029 7=005, 043, 011, 003, 009, 019, 063, 071, 027 8=061, 039, 073, 047, 025 9=017, 021, 013, 065, 049, 041, 101 All other counties grouped into other or NA category
geo_state	Indicates the student's state of origin, two-letter U.S. state abbreviations (e.g., AR, TX, CA).
gpa_grade_point_average	The student's cumulative high school GPA
hawaiian	Y/N/U; ethnicity and race identification
hispanic	Y/N/U; ethnicity and race identification
hs_grad_year	Indicates the year the student graduated from high school or completed an equivalent credential (e.g., GED or HiSET). For students who are still enrolled in high school at the time of the record (such as dual-enrollment students), the variable reflects the expected high school graduation year. Four-digit year (e.g., 2018, 2019, 2020).

ipeds_race	Indicates the student's race and ethnicity category as defined by the Integrated Postsecondary Education Data System (IPEDS) reporting standards. NR: Non-resident alien (international student); UN: Race/ethnicity unknown or not reported; HS: Hispanic or Latino; IN: American Indian or Alaska Native; AS: Asian; BL: Black or African American; NH: Native Hawaiian or Other Pacific Islander; WH: White; TW: Two or more races.
major_1	Indicates the student's primary academic major for the given term. Program code corresponding to the institution's approved degree program list maintained by the Arkansas Higher Education Coordinating Board (AHECB).
non_us_resident	Indicates whether the student is classified as a non–U.S. resident. Y: Student is a non–U.S. resident (international student); N: Student is a U.S. citizen or permanent resident.
nperterm	Number of enrollments within same term across institutions. Count of student–institution records within a given seasonal term.
resident_state	Indicates the student's current legal state of residence. Two-letter U.S. state abbreviations (e.g., AR, TX, CA).
school_type	The type of institution the student attended for the given term. T = Technical C = College (Liberal Arts) 2 = 2-year (Community College) 4 = 4-year (College or University)
ssn_id	Hashed 64-character identifier for student.
start_time_seq	time_seq value of first analytical record. time_seq where rn_final=1
stem	Indicator equal to 1 if CIP code matches AR STEM list. 1=STEM, 0=Non-STEM
stem_degree_ever_student	Indicates whether a student has earned a degree in a STEM field at any institution associated with the student in the dataset. 1: The student has at least one recorded graduation in a STEM field; 0: No STEM degree completion is recorded for the student. Note that this variable aggregates STEM degree completion across all institutions associated with the student in the dataset.
stem_starter_student	Equals 1 if first analytical record is STEM. 1 if rn_final=1 and stem=1
student_level	Academic level of student in the term. 00 = Unclassified undergraduate 01 = Freshman

	02 = Sophomore 03 = Junior 04 = Senior 05 = Master's level 06 = Post-Baccalaureate Certificate or Specialist or First-Professional Certificate/Degree 07 = Doctoral - Research/Specialist 08 = Doctoral - Other 09 = Doctoral - Professional Practice 10 = Unclassified post-baccalaureate (see note 3) 11 = Unclassified graduate (see note 5) 13 = High school student, but NOT a high school senior 14 = High school senior
t_start	Indicates the term value (term_val) corresponding to the student's first observed enrollment term in the dataset. t_start identifies when a student first appears in the analytical dataset. It essentially marks the entry point of the student's longitudinal record. For each student: t_start = minimum(term_val).
term_val_imp	Standardized academic term indicator that collapses original "on" (enrolled) and "off" (not enrolled) term codes into a single seasonal term value. 0=Summer II, 1=Fall, 2=Spring, 3=Summer I
time_seq	Sequential index enumerating all academic terms.
timediff	Number of terms elapsed between the first analytical record and the current record.
totalcrhours	Number of credit hours enrolled in during the term.
tuition_status	Indicates the student's tuition residency classification for the term, based on the tuition amount charged by the institution. 0: In-state tuition; 1: Out-of-state tuition; 2: In-district tuition; 3: Out-of-district tuition; 4: In-state tuition classification under Arkansas Act 844.
tuition_waiver	Indicates whether the student received a waiver that reduced or eliminated out-of-state tuition charges. Y: Student received a partial or full out-of-state tuition waiver; N: Student did not receive a tuition waiver.
white	Y/N/U; ethnicity and race identification

[i] Note that the temporal scope overlaps with the Covid-19 pandemic. The analytic framework presented in this report did not formally investigate any effects due to Covid. However, we did see some interesting results that lined up with the pandemic. We saw that STEM PhDs from the 2014 graduating class had reduced wages 5 years out, which could be an artifact of Covid. We also see that belonging to the 2018 cohort increased chances of exit. However, that effect was similar to the effect of belonging to a few other

cohorts, so we hesitate to say the increased rate of exit for the 2018 cohort was due to Covid without further investigation. While the pandemic likely had an influence on STEM attrition rates, the current framework did not formally investigate the influence of Covid-19, and therefore we refrained from making any claims about the effect of the pandemic.

[ii] Any student who declared a STEM major within their first 4 enrolled terms is designated as a STEM starter. Any student who does not meet this definition is by default classified as a non-STEM starter. The non-STEM starter category includes students who declare a non-STEM major or defer declaration beyond their first 4 terms. Some non-STEM starters may have eventually gone on to declare a STEM major after their 4 first terms, but these students never make it into the STEM starter group. The majority of the analyses in this report focus on the STEM starter group. We wanted to study why students enter a STEM major early in their postsecondary career and either stay or exit, rather than students who enter the STEM pathway late in their postsecondary career. Thus, we focus only on students who declared a STEM major within their first 4 enrolled terms.