

SUMMARY OF FIRST2 DATA

John Stewart

WVU

11-15-2025



FIRST2 DATA SUMMARY

Institution	Year	Began Survey	Completed First Gen Question	Completed Survey	Analysis Records
Fairmont	Spr 2025	92	50	47	47
Glenville	Spr 2025	103	78	76	72
Blue Ridge	Fall 2024	106	77	73	74
EKU	Spr25				55
WVU	Spr22-Fall23				1009
Marshall	Spring 2025	97	83	77	69
Marshall	Spring 2023	33	21	22	
Charleson	Spring 2025	25	17	16	
Blue Ridge	Spring 2022	56			
Fairmont	Spring 2022	84	70	70	70
Fairmont	Fall 19 to Spring 21	199			184

FIRST2 ANALYSIS NOTES

- Marshall Spring 2023 too small for meaningful analysis
- Charleston Spring 2025 too small for meaningful analysis
- Blue Ridge Spring 2022 given as two surveys: most variables in first survey with rural/first gen in second survey. Survey 1 – 56 started, 45 finished. Survey 2: 38 started, 34 finished. Matched -29. Again, too small for meaningful analysis. A report on this same was given at a First2 conference.
- For Spring 2021 and earlier, only FSU and WVU data exist and we were not giving the self-reported FirstGen or Rural Question.
- For Spring 2021 and earlier, FSU we use similar but not identical questions. Asked about graduating with major not graduating in general, only student with major retained.
- In Fall 21, two surveys were given at FSU. These will have the same matching issues and have not been analyzed.

RURAL AND FIRST GENERATION



The surveys asked the student to self-report rural and first-generation status. Using the prompts:

Do you identify as a first-generation college student?

Would you describe yourself as having a rural upbringing?



Prior First2 analysis used institutional reporting of First Gen and Rural markers derived from the location of the high school.

VARIABLES

Self-Efficacy (SE, 1-5) – The student's belief they can succeed in their current science class.

Belonging (BL, 1-5) – The student's feeling of belonging in their current science class.

Science Identity (ID, 1-5) – The student's identification as a science person.

Grad.STEM (%) - The percentage chance the student graduates with a STEM degree.

Grad.NoSTEM (%) - The percentage chance the student graduates with a non-STEM degree

NoGrad (%) - The percentage chance the student does not graduate.

VARIABLES

First 2 Activities – Activities like those provided by the First2 program, but not necessarily provided by the program. Collected on a 4 point scale, so 2.5 represents a balance of yes and no. Some captured dichotomously but scaled to 4 points.

- o Did you have a support network when you came to campus?
- o If yes, was that support network developed during high school?
- o Did you have a STEM research experience before you started college?
- o Did you have a STEM research experience as a college freshman?
- o Did you participate in professional development in science process skills?
- o Did you a college student come to talk at your high school about college?
- o Did you a college student come to talk at your high school about STEM careers?
- o Have you provided mentorship to high school students or students earlier in their college studies about STEM careers?

VARIABLES

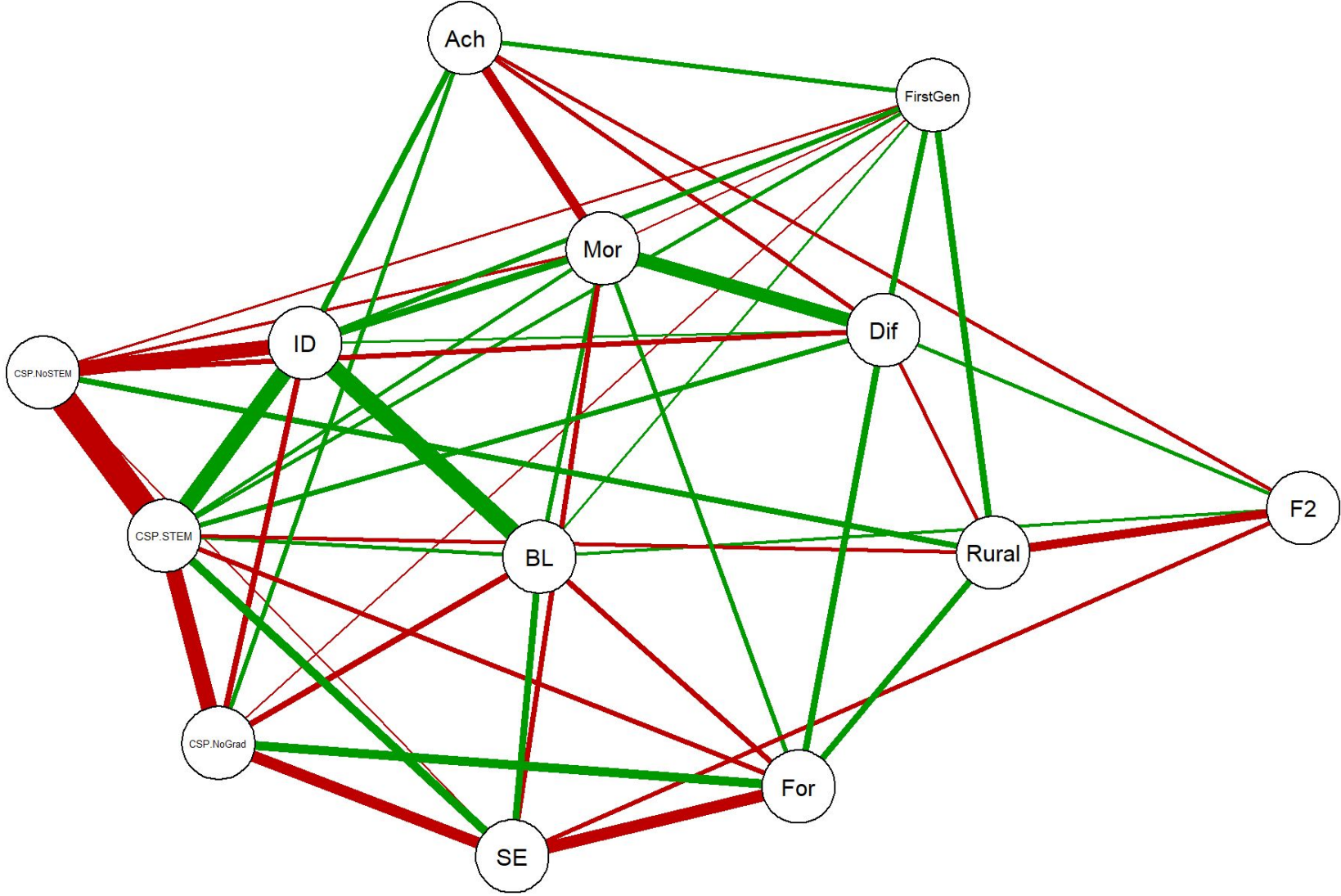
- Career Identity Status – Has the student developed a well-researched professional identity.
 - Achievement (Ach, 1-5) – The student has selected a career path after serious investigation.
 - Foreclosure (For, 1-5) – The student has had a career path imposed on them generally by a parent.
 - Diffusion (Dif, 1-5) – Is working toward a career, but is not committed and is open to a change.
 - Moratorium (Mor, 1-5) – Cannot decide on a career.

Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.49	0.57	4.02	3.96	3.72	2.69	82.19	13.59	4.21	3.77	2.57	2.74	1.54
Rural	0.59	1.00	4.04	3.90	3.67	2.49	77.59	19.22	3.19	3.83	2.44	2.69	1.70
FirstGen	1.00	0.70	4.05	4.04	3.90	2.67	87.39	10.22	2.39	3.93	2.78	2.63	1.59
Majority	0.00	0.00	3.91	3.89	3.58	2.89	83.69	8.46	7.85	3.62	2.46	2.73	1.31

RESULTS FAIRMONT STATE SPRING 2025 (N=47)

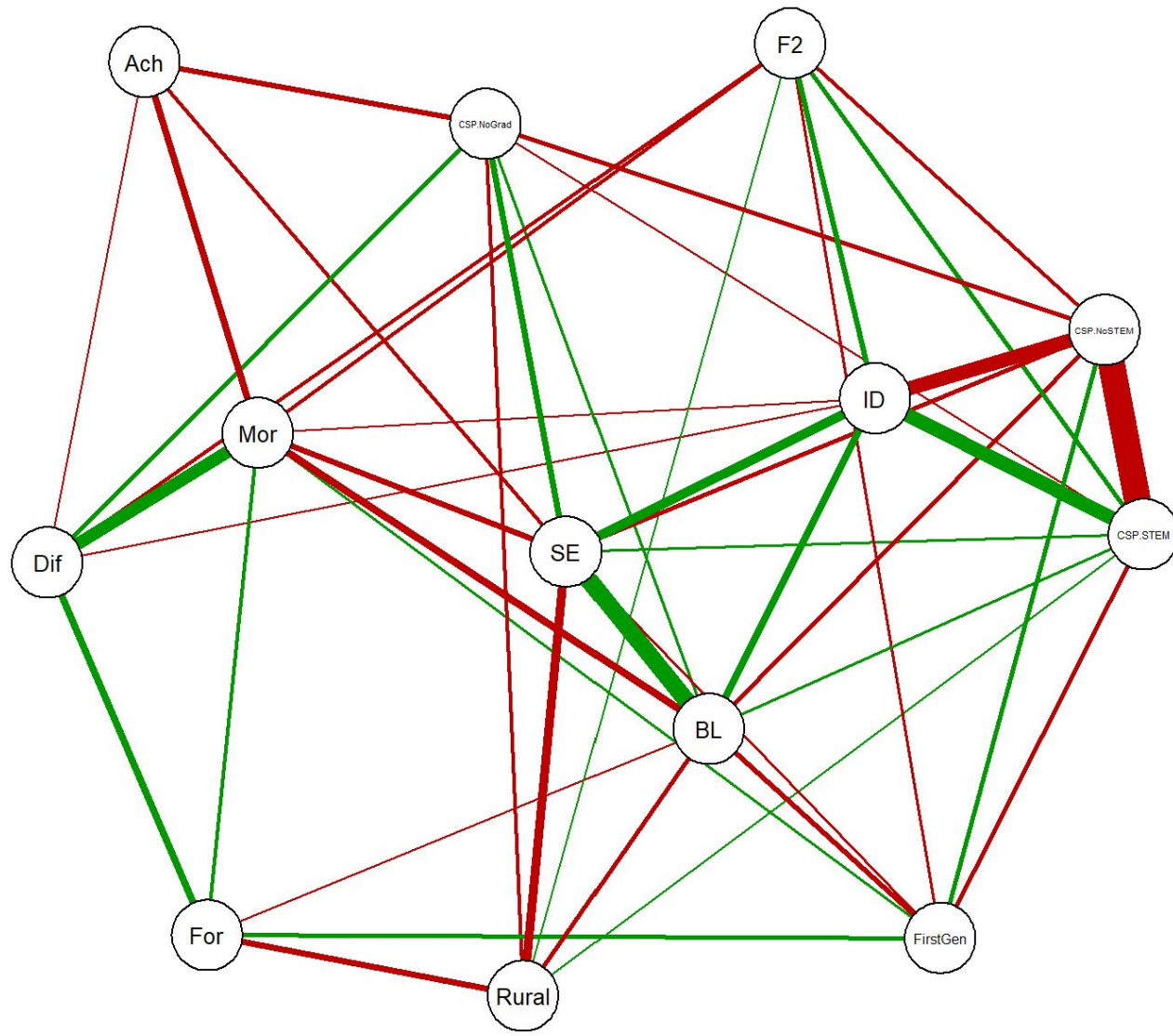
CORRELATION

- Correlation measures how two variables change together.
- The following shows positive correlations as solid green lines, negative as dashed red lines.
- Thicker lines are stronger correlations.
- The figure groups strongly correlated variables together



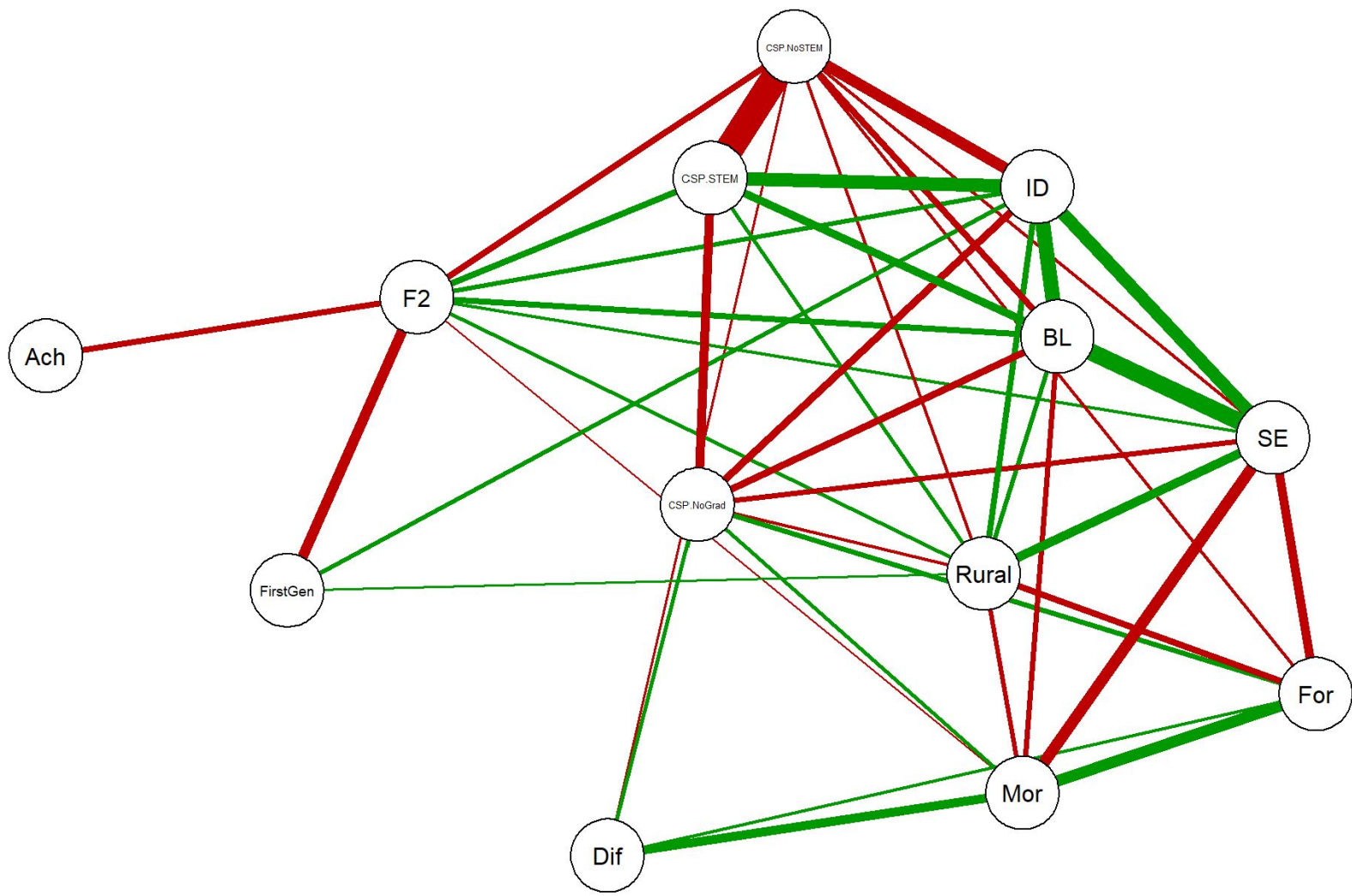
Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.33	0.61	3.49	3.72	2.51	2.68	38.01	59.65	2.33	3.67	2.64	2.62	1.62
Rural	0.32	1.00	3.28	3.58	2.46	2.74	41.18	57.95	0.86	3.68	2.59	2.67	1.44
FirstGen	1.00	0.58	3.33	3.43	2.37	2.55	27.79	70.54	1.67	3.65	2.71	2.81	1.85

RESULTS GLENVILLE SPRING 2025 (N=72)



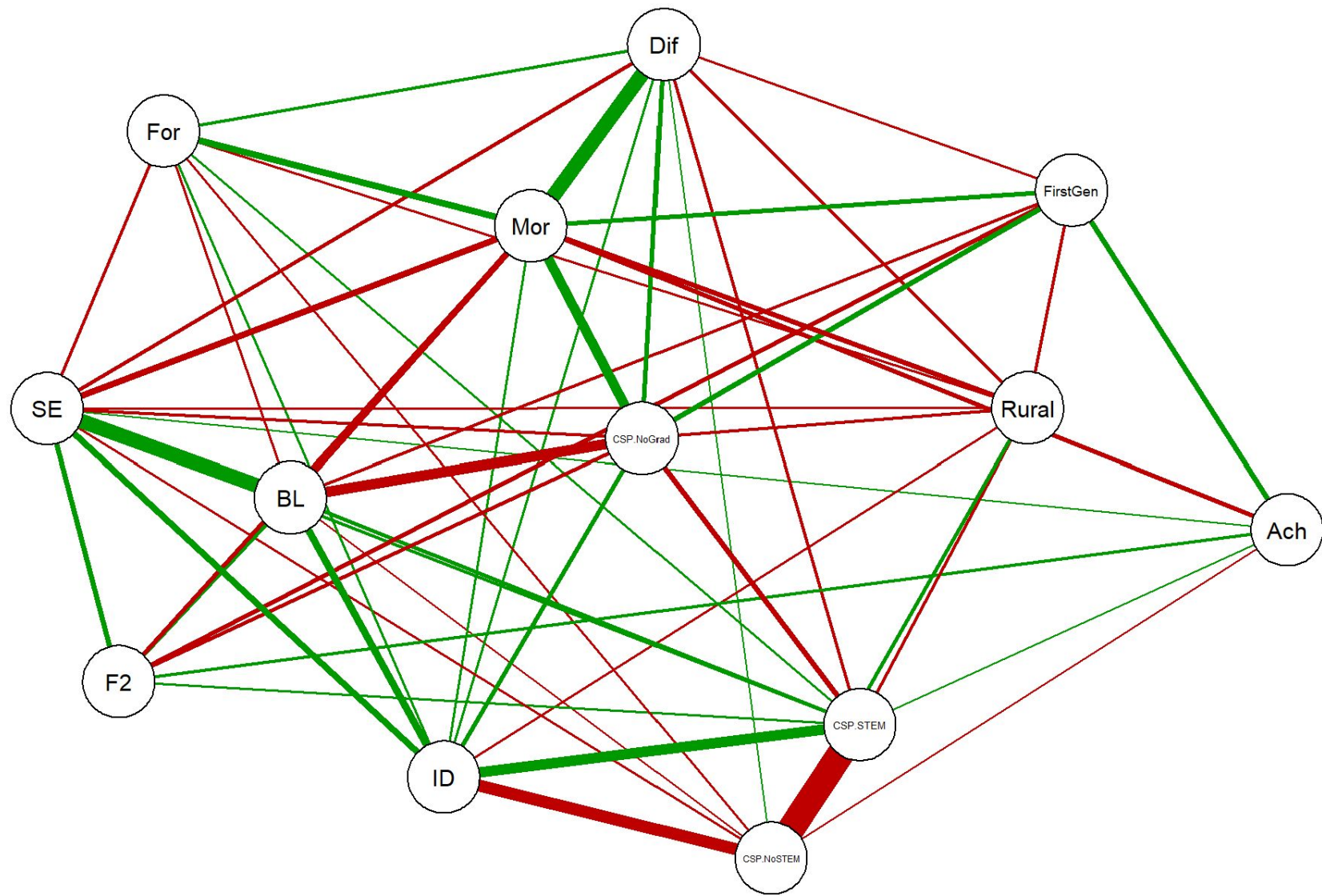
Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.42	0.47	3.93	3.68	2.94	2.50	66.43	29.86	3.70	3.55	2.78	2.70	1.55
Rural	0.49	1.00	4.21	3.85	3.24	2.64	73.66	24.46	1.89	3.57	2.70	2.47	1.33
FirstGeneration	1.00	0.55	3.98	3.61	3.22	2.17	68.87	26.94	4.19	3.58	2.73	2.79	1.60
Majority	0.00	0.00	3.72	3.66	2.65	2.56	53.52	42.96	3.52	3.60	2.74	2.86	1.58

BLUE RIDGE FALL 24
(N=74)



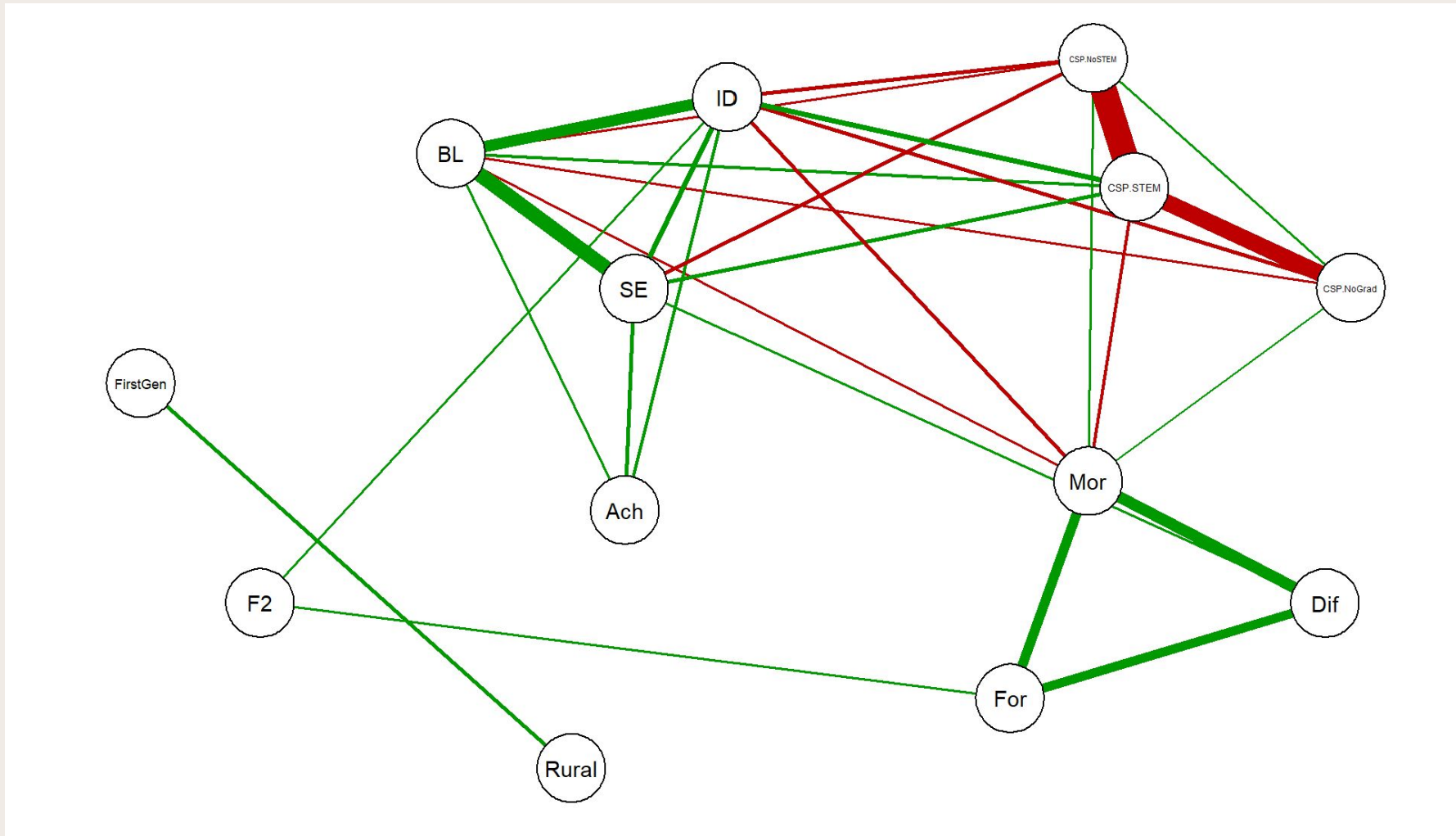
Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.16	0.38	4.09	4.20	3.62	2.91	81.87	16.44	1.69	3.56	2.45	2.47	1.25
Rural	0.10	1.00	3.98	4.30	3.49	2.98	76.33	22.86	0.81	3.62	2.26	2.17	1.17
FirstGen	1.00	0.22	4.07	3.96	3.76	2.60	86.00	10.11	3.89	4.06	2.17	2.94	1.28
Majority	0.00	0.00	4.20	4.22	3.61	2.97	84.30	14.19	1.52	3.44	2.63	2.54	1.28

EKU SPRING 2025
(N=55)



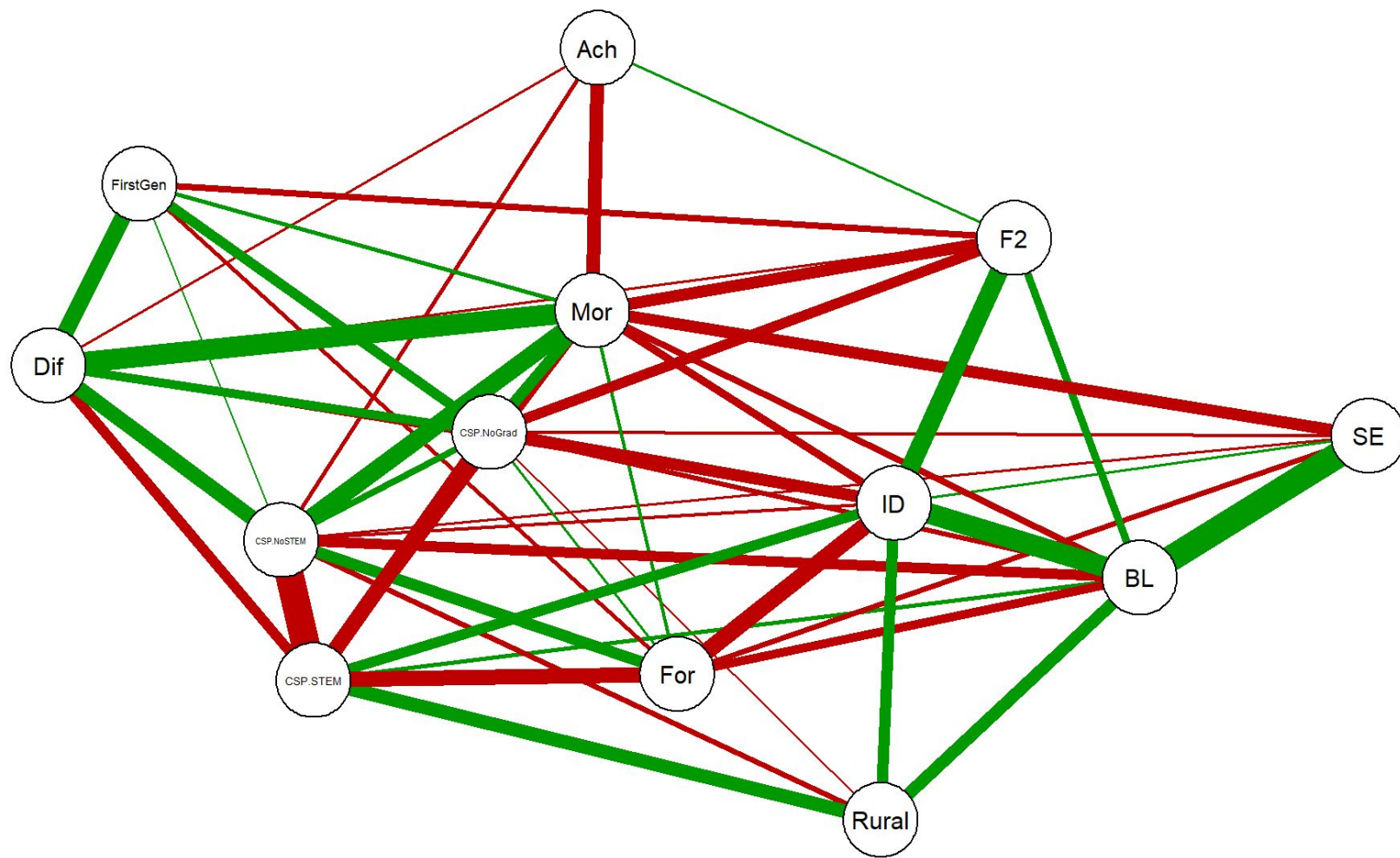
Type	First Gen	Rural	SE	BL	ID	F2	Grad. STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.15	0.39	3.77	3.98	3.94	1.91	93.69	4.63	1.68	3.36	3.03	2.74	1.96
Rural	0.23	1.00	3.70	3.94	3.94	1.90	93.34	4.73	1.93	3.39	2.96	2.63	1.91
First Gen	1.00	0.59	3.56	3.83	3.96	1.79	92.83	4.71	2.46	3.39	3.00	2.66	1.78
Majority	0.00	0.00	3.84	4.03	3.94	1.92	93.94	4.59	1.47	3.35	3.08	2.83	2.01

WVU SPRING 2022 TO FALL 2023
(N=1009)



Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.33	0.58	4.09	4.18	4.06	2.73	89.90	4.90	2.30	3.52	2.75	2.84	1.53
Rural	0.35	1.00	4.14	4.33	4.23	2.73	94.75	3.45	1.80	3.46	2.75	2.79	1.54
FirstGen	1.00	0.61	4.06	4.09	4.04	2.53	89.35	6.48	4.17	3.63	3.24	3.07	1.35
Majority	0.00	0.00	4.11	4.08	3.90	3.01	82.65	6.05	1.30	3.68	2.58	2.63	1.50

MARSHALL SPRING 25
(N=69)

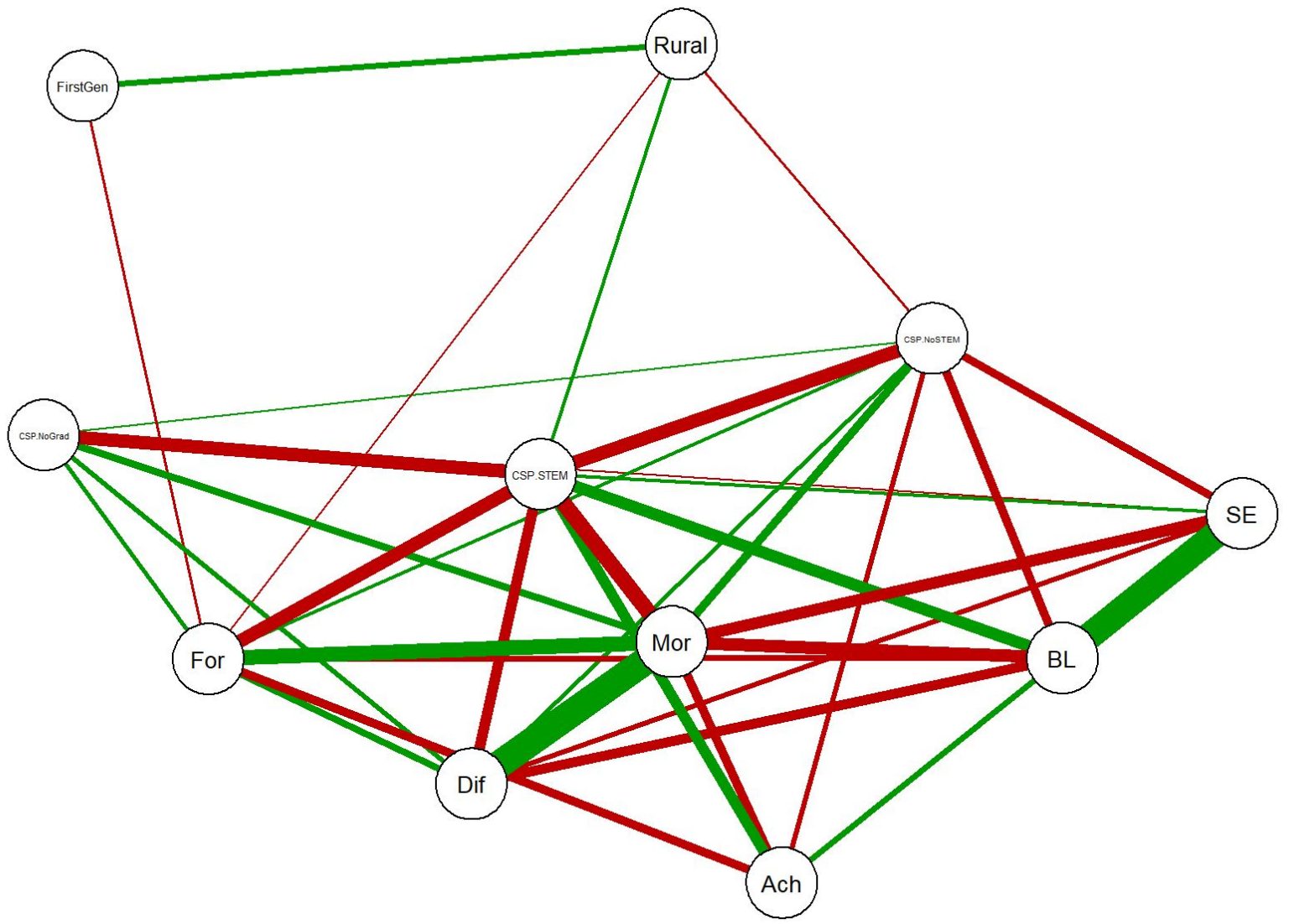


Type	FirstGen	Rural	SE	BL	ID	F2	Grad.STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.31	0.61	3.98	4.09	3.89	2.83	84.03	11.76	4.21	3.48	2.86	2.89	1.63
Rural	0.33	1.00	4.02	4.25	4.03	2.92	83.30	11.34	5.36	3.62	2.80	2.71	1.51
FirstGen	1.00	0.64	3.89	3.98	3.95	2.68	79.23	15.43	5.34	3.39	2.84	3.05	1.34
Majority	0.00	0.00	3.89	3.92	3.66	2.82	86.58	11.42	2.00	3.32	2.97	3.26	1.89

FAIRMONT STATE SPRING 2022
(N=70)

Type	First Gen	Rural	SE	BL	ID	F2	Grad. STEM	Grad.NoSTEM	NoGrad	ACH	DIF	MOR	FOR
Overall	0.38	0.29	4.04	3.97	NA	NA	0.86	0.15	0.10	3.48	2.57	2.80	1.76
Rural	0.52	1.00	4.12	4.05	NA	NA	0.89	0.11	0.11	3.63	2.50	2.67	1.59
First Gen	1.00	0.40	4.04	4.00	NA	NA	0.85	0.16	0.10	3.52	2.47	2.84	1.60
Majority	0.00	0.00	4.00	3.94	NA	NA	0.86	0.14	0.08	3.47	2.64	2.84	1.84

FAIRMONT STATE FALL 2019 TO SPRING 2021
(N=184)



SUMMARY

- The samples were quite similar despite very different academic contexts.
- SE and BL were high ~4 at all institutions except Glenville. These were fairly consistent for rural and first generation students.
- Science Identity was also generally high, but more variable with lower levels of ID at ECU, Glenville and Blue Ridge. Again, fairly consistent for FirstGen and Rural students.
- Projected graduation percentage was generally very high, above 80% with the large four year colleges (WVU, Marshall) with the highest percentage. The percentage reported by Glenville was quite low (did this sample only collect STEM students?).
- FirstGen and Rural students reported generally lower graduation probabilities at some institutions, but this was not consistent.
- All students reported high levels of successfully selecting a career identity.
- The patterns of correlation between the variables were extremely inconsistent between institutions and no general pattern was identifiable.

PART 2 – WVU DATA

What do these words mean?

PRELIMINARIES



FIRST GENERATION STUDENTS

First generation students are defined as those students where neither parent earned a 4-year college degree.



URBANICITY/RURALITY (NCES)

- THE NCES locale codes were used to code each student's high school. The codes are as follows:
- City – Large (11): Territory inside an Urbanized Area and inside a Principal City with population of 250,000 or more.
- City – Midsize (12): Territory inside an Urbanized Area and inside a Principal City with population less than 250,000 and greater than or equal to 100,000.
- City – Small (13): Territory inside an Urbanized Area and inside a Principal City with population less than 100,000.
- Suburban – Large (21): Territory outside a Principal City and inside an Urbanized Area with population of 250,000 or more.
- Suburban – Midsize (22): Territory outside a Principal City and inside an Urbanized Area with population less than 250,000 and greater than or equal to 100,000.
- Suburban – Small (23): Territory outside a Principal City and inside an Urbanized Area with population less than 100,000.
- Town – Fringe (31): Territory inside an Urban Cluster that is less than or equal to 10 miles from an Urbanized Area.
- Town – Distant (32): Territory inside an Urban Cluster that is more than 10 miles and less than or equal to 35 miles from an Urbanized Area.
- Town – Remote (33): Territory inside an Urban Cluster that is more than 35 miles from an Urbanized Area.
- Rural – Fringe (41): Census-defined rural territory that is less than or equal to 5 miles from an Urbanized Area, as well as rural territory that is less than or equal to 2.5 miles from an Urban Cluster.
- Rural – Distant (42): Census-defined rural territory that is more than 5 miles but less than or equal to 25 miles from an Urbanized Area, as well as rural territory that is more than 2.5 miles but less than or equal to 10 miles from an Urban Cluster.
- Rural – Remote (43): Census-defined rural territory that is more than 25 miles from an Urbanized Area and also more than 10 miles from an Urban Cluster.

- The National Center for Education Statistics (NCES) classifies the rurality of each US high school.
- These codes are called Locale codes.



RURAL STUDENTS

We define rural students as students who attended a high school with a NCES Locale code of 42 or 43.



WHAT THE DATA CAN'T TELL YOU



DECISIONS MADE IN HIGH SCHOOL ARE CRITICAL

- All students in the dataset were at one time enrolled at West Virginia University (WVU); therefore, all students had made the decision to attend college at a major university.
- The data cannot explain the decision to attend college and elect a STEM major, or more importantly, the decision not to attend college.



STEM MATRICULATION



THE DATASET

- We accessed the course records of all students who were enrolled as undergraduates at West Virginia University's (WVU) Morgantown campus from 2000 to present; 146,000 students.
- We restricted the dataset to US citizens who were admitted as First Time Freshman beginning after the Fall 2005 semester when WVU began collecting reliable first-generation data.



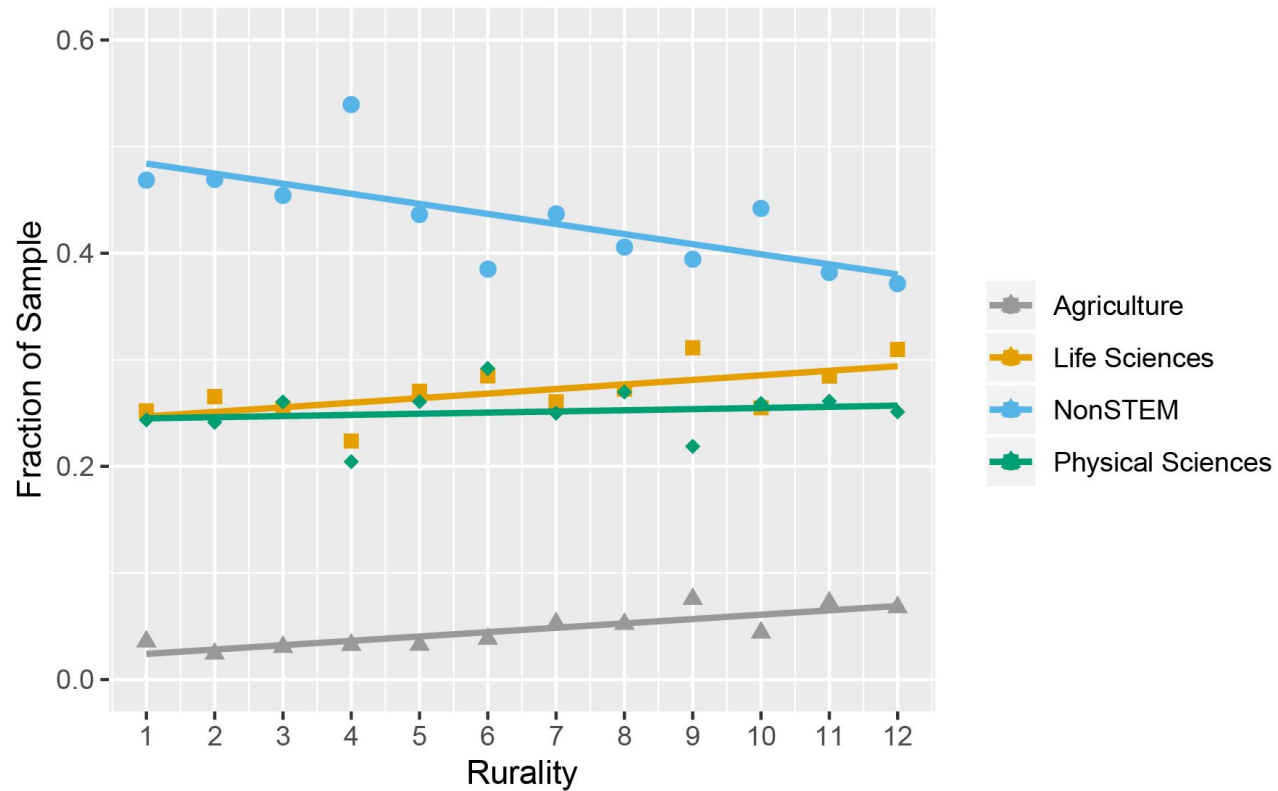
STEM

- Defining Science, Technology, Engineering, and Mathematics (STEM) students is not straightforward.
- For this talk, we divide students into four classes:
 - Non-STEM
 - Agriculture
 - Life sciences including biology and health professions. This also includes general science majors.
 - Physical science including engineering, math, chemistry, and physics.



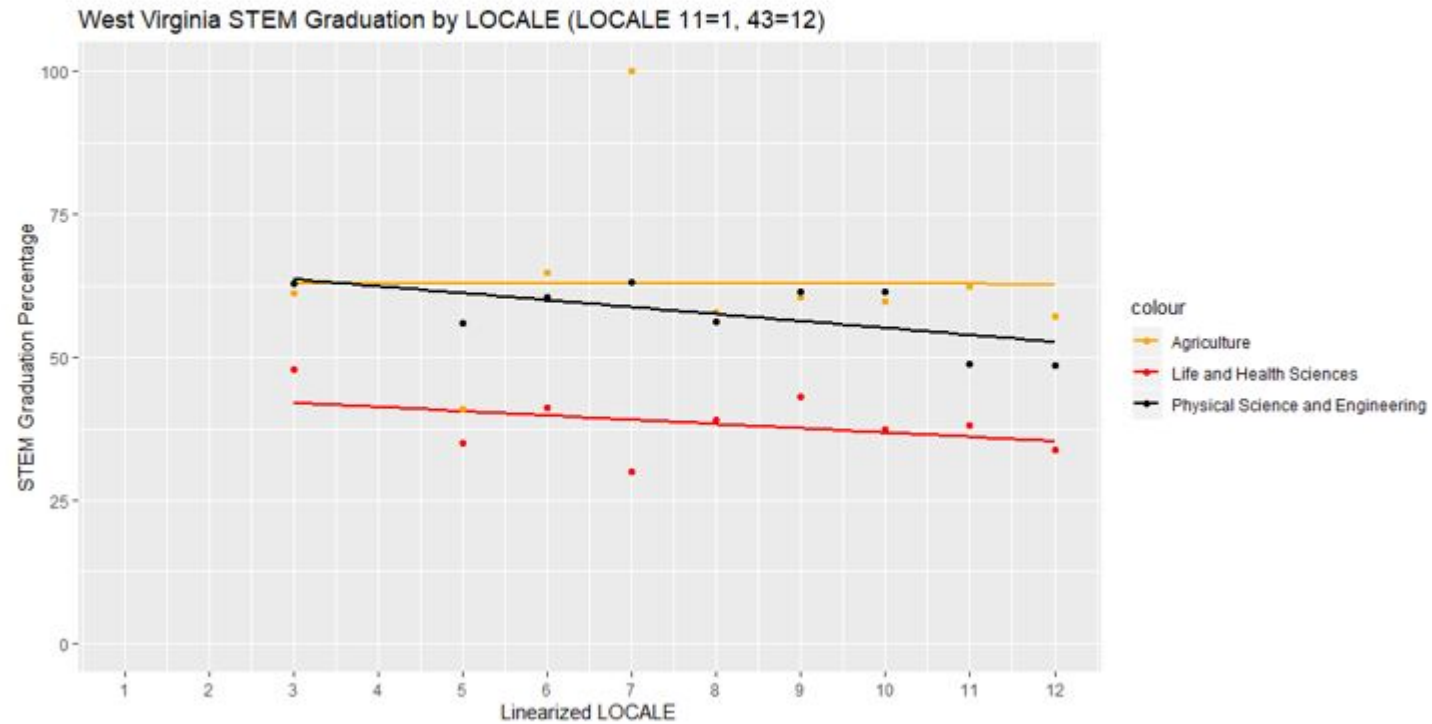
STEM MAJOR ELECTION RATES

Major Selection by Rural/Urban Classification



COLLEGE SUCCESS – GRADUATION IN STEM BY LOCALE

Graduation rates in STEM are independent of Locale for Agriculture majors, decline with increasing rurality for Life Sciences majors, and decline more sharply for Physical Science majors. For Physical Science majors, Locale 41 graduate at a 10% higher rate than Locale 42 and 43.



COLLEGE SUCCESS

- First-generation students graduate at a lower rate ($p < .001$).

	Graduation Rate Non-STEM Majors	Graduation Rate Physical Science Majors
First Generation	50%	56%
Not First Generation	62%	69%



When are rural or first-generation students most at risk of leaving STEM?

HAZARD

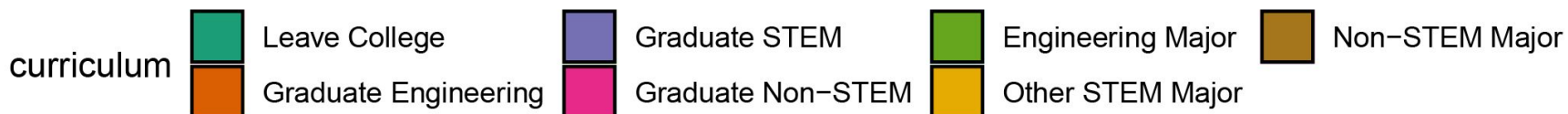
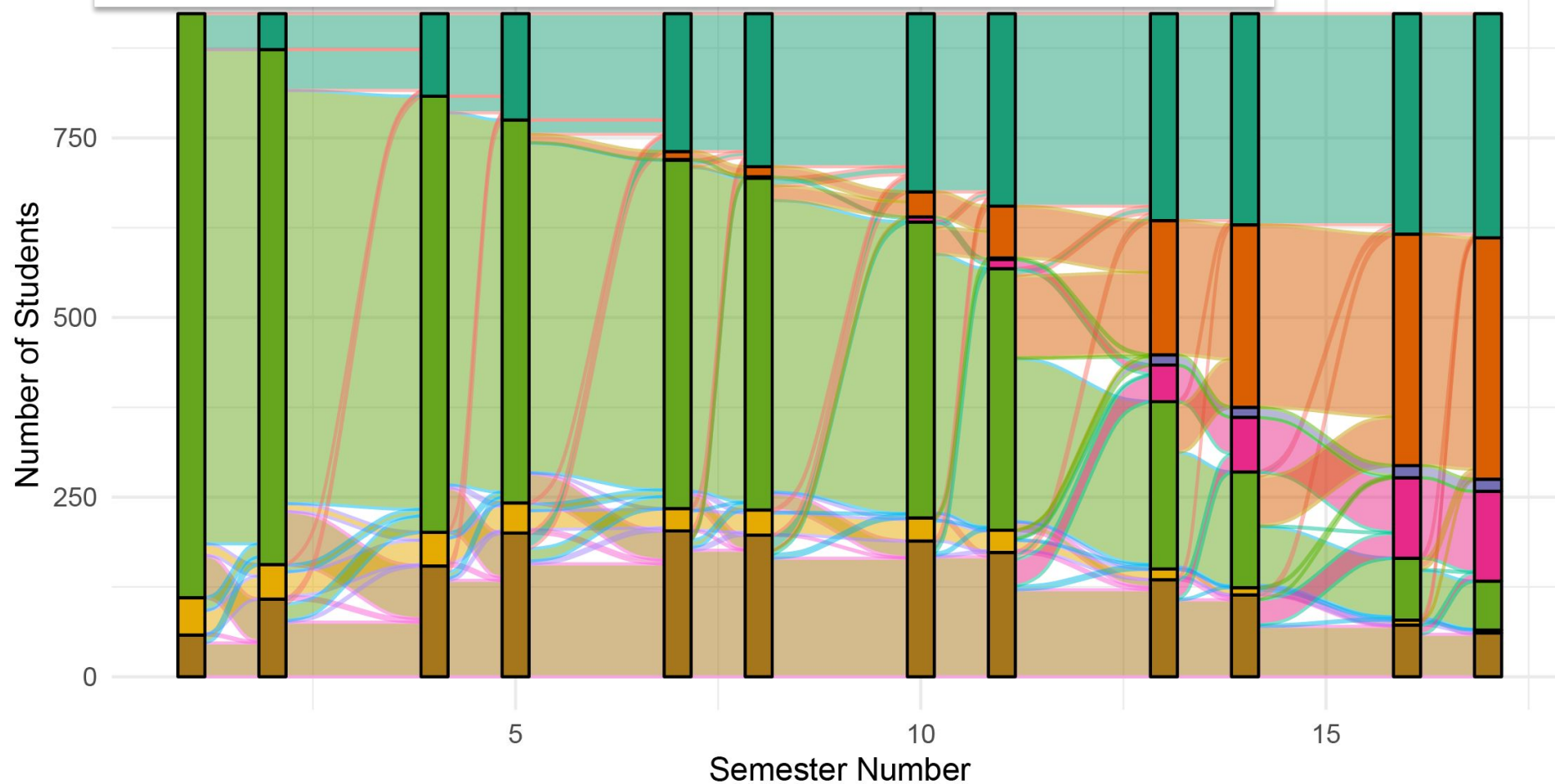


DATASET

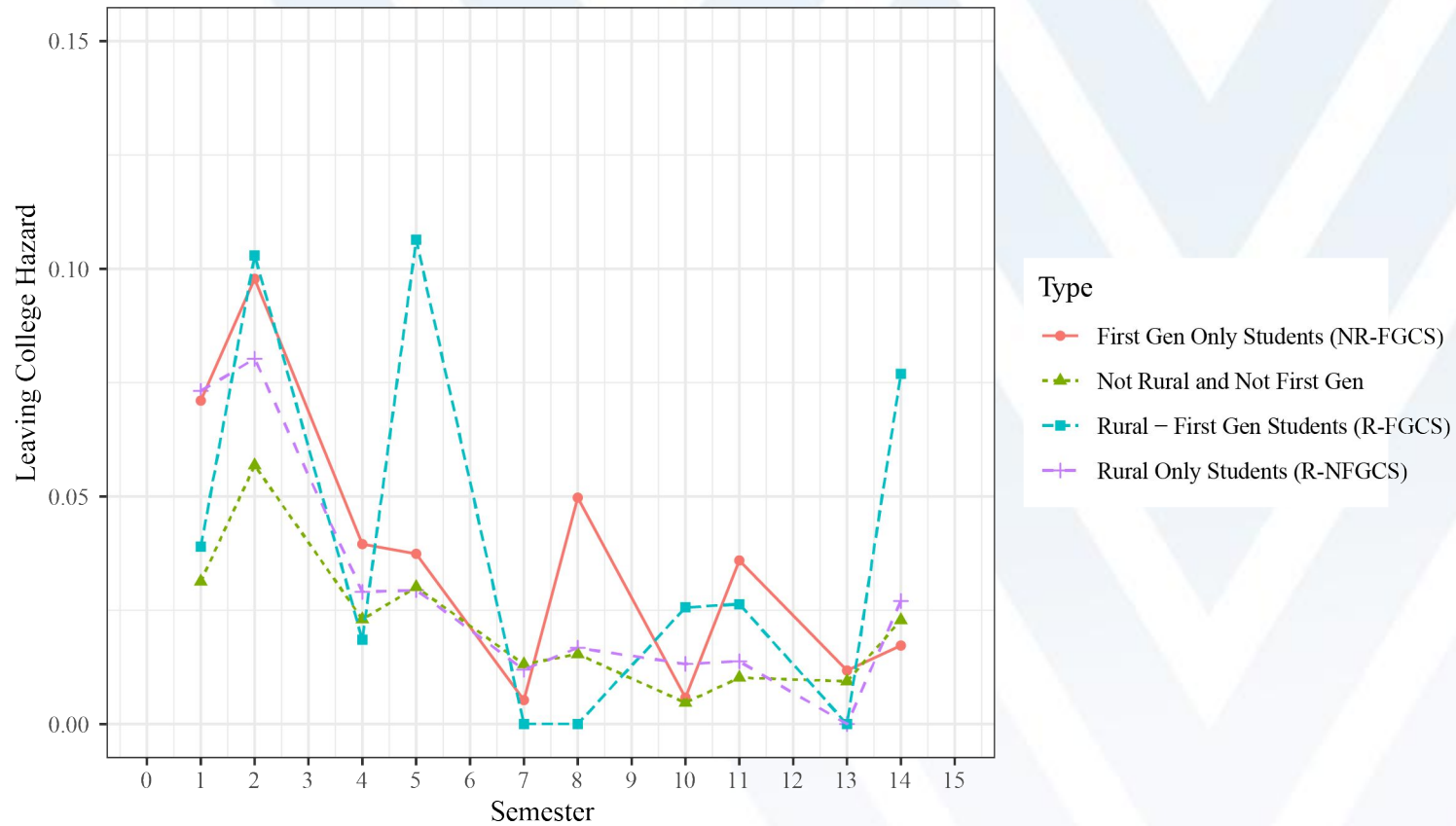
- Data-Products/DataProduct-Institutional-Rural/Ver4/DemoRpt
MajSeq-Eng-rural-v4.csv
- WV engineering students with LOCALE code.
- Admit code filtered FTF, readmit, transfer.
- Survival plots filtered engineering first semester, fall first semester, FTF.
- All plots with 6 year graduation window.



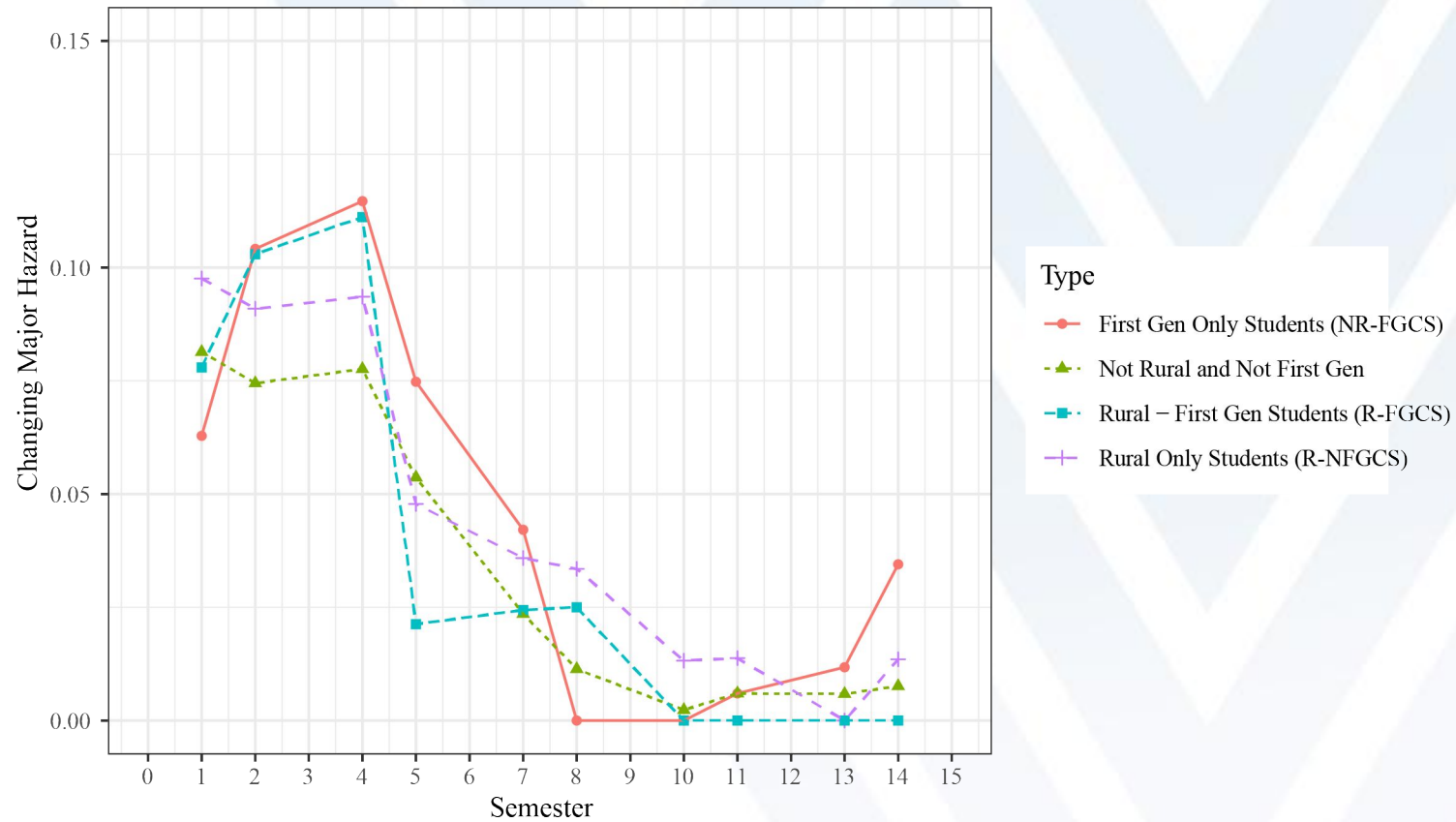
Major Changing Patterns of Rural WVU Engineering Students



Rural First-generation students are at elevated risk of leaving college in both the spring freshmen and spring sophomore semesters.



Rural First-generation students are at lower risk of changing to a non-STEM major in both the spring freshmen and spring sophomore semesters.



When are rural or first-generation students most at risk of leaving STEM

RISK FACTORS



MATH READINESS

- For engineering students in this dataset, the student is “Math Ready” if they are ready to enroll in Calculus I (MATH 155) in their first semester.
- For current WVU students, the math department has recently changed its placement criteria, placing the majority of students in MATH 153. We will have to wait a bit to see what this ultimately does to graduation rates.



RURAL/FIRST GENERATION DIFFERENCES

Table 1 - Variable Definitions and Descriptive Statistics. Only students enrolled as first-time freshman are included.

Variable	Description				
Rural Status	Not Rural	Not Rural	Rural	Rural	
First Gen Status	Not First Gen	First Gen	Not First Gen	First Gen	
N	3437	422	550	93	
ACTM%	76±18	73±20	74±18	74±21	ACT or SAT mathematics percentile score.
ACTV%	71±19	68±20	70±19	69±21	ACT English or SAT verbal percentile score.
HSGPA	3.67±.5	3.56±.5	3.65±.4	3.64±.4	High school grade point average.
AP Count	1.25±2.6	0.81±1.9	0.32±1.0	0.45±1.2	Total number of AP credit hours applied to college credit.
Transfer Count	1.21±2.7	1.26±2.8	1.62±2.4	1.66±2.3	Total number of dual enrollment credit hours applied to college credit.
Math Ready %	64	51	52	39	Percentage of students enrolled in Calculus I or more advanced mathematics course in the first semester of college.



RISK FACTORS – LEAVING COLLEGE

Table 2 – Logistic regression predicting the probability of leaving college before the second year of college where B is the normalized regression coefficient, SE the standard error, and OR the odds ratio.

Variable	B	(SE)	OR (95% CI)	<i>p</i>
Intercept	-2.32	(0.13)	-	<0.001
ACTM%	0.13	(0.08)	1.14 (0.98, 1.33)	0.085
ACTV%	0.03	(0.07)	1.03 (0.90, 1.18)	0.623
HSGPA	-0.76	(0.07)	0.47 (0.41, 0.53)	<0.001
AP Count	-0.64	(0.15)	0.53 (0.38, 0.70)	<0.001
Transfer Count	-0.04	(0.06)	0.97 (0.85, 1.07)	0.532
Math Ready				
Not Math Ready (ref)	-	-	-	-
	-0.41	(0.15)	0.66 (0.49, 0.89)	0.006
Rural				
Not Rural (ref)	-	-	-	-
	0.34	(0.16)	1.41 (1.03, 1.92)	0.031
FGCS				
Not FGCS (ref)	-	-	-	-
	0.22	(0.13)	1.25 (0.96, 1.61)	0.092

High School is still important

Rural Students are 40% more likely to leave college



RISK FACTORS – LEAVING STEM WHILE STAYING IN COLLEGE

Table 3 – Logistic regression predicting the probability of change to a major other than engineering by second year of college where B is the normalized regression coefficient, SE the standard error, and OR the odds ratio.

Variable	B	(SE)	OR (95% CI)	<i>p</i>
Intercept	-1.31	(0.10)	-	< 0.001
ACTM%	-0.12	(0.07)	0.89 (0.77, 1.02)	0.086
ACTV%	0.08	(0.07)	1.09 (0.96, 1.24)	0.208
HSGPA	-0.04	(0.06)	0.96 (0.86, 1.09)	0.560
AP Count	-0.02	(0.07)	0.98 (0.84, 1.13)	0.777
Transfer Count	-0.00	(0.05)	1.00 (0.89, 1.10)	0.942
Math Ready				
Not Math Ready (ref)	-	-	-	-
	-1.13	(0.14)	0.32 (0.24, 0.42)	< 0.001
Rural				
Not Rural (ref)	-	-	-	-
	-0.05	(0.16)	0.95 (0.69, 1.28)	0.733
FGCS				
Not FGCS (ref)	-	-	-	-
	0.19	(0.12)	1.21 (0.95, 1.53)	0.124

Only Math Readiness matters for staying in STEM.

Math Ready students have 200% higher odds of staying in STEM

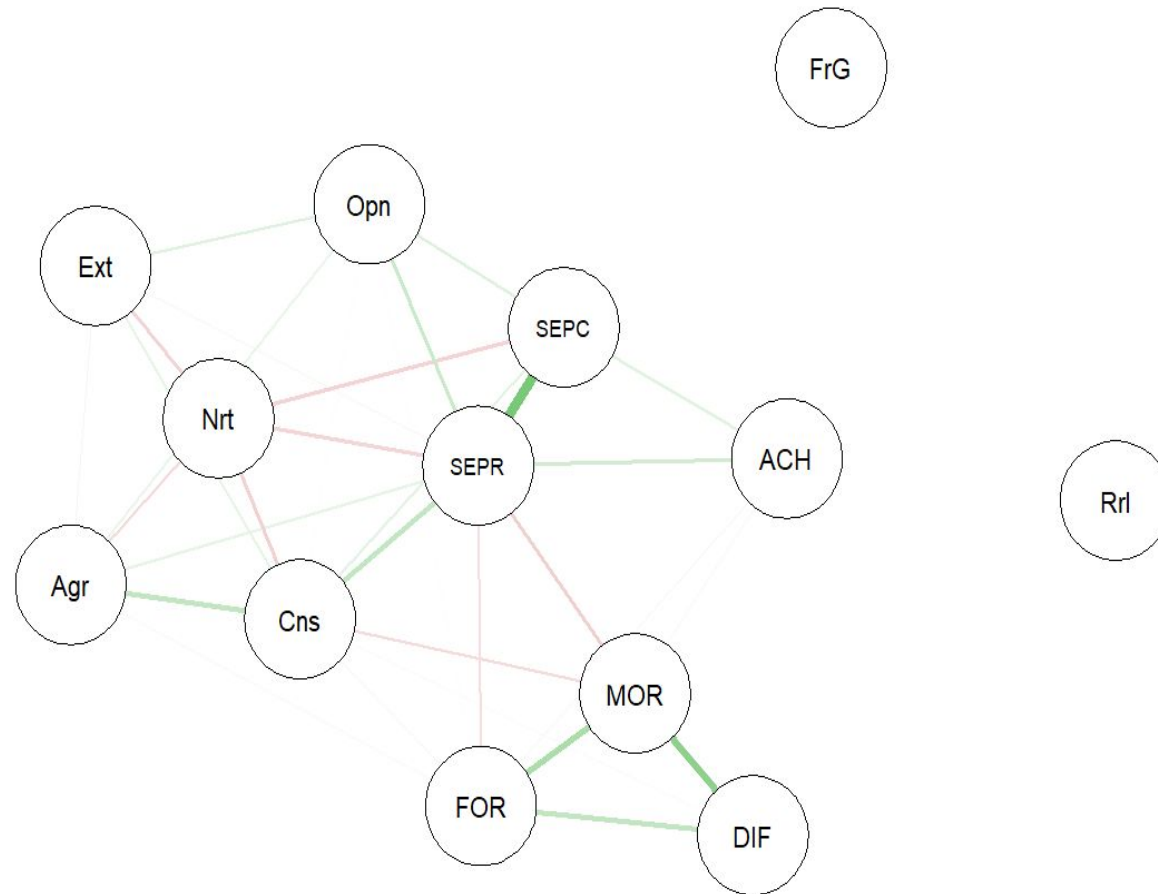


But feelings and beliefs are important too.

NON-COGNITIVE FACTORS



NON-COGNITIVE FACTORS



We also measure a broad collection of non-cognitive variables possibly related to rurality, first generation status, and retention. Thus far, these factors appear much less important than academic factors.



Preliminary results from the First2 survey instruments.

DIVING DEEPER



FIRST2 NETWORK DATA

- Preliminary results from the First2 survey instruments.
- Data collected at Fairmont State University and Blue Ridge Community and Technical College this spring.



FIRST2 NETWORK DATA

- Which of these choices best expresses your feelings about the statement, "**My parents/guardians often discussed preparing for college**"? (Options range from 1: strongly disagree to 5:strongly agree.)
- First-generation students' mean was 2.92; continuing generation students' mean was 4.01.
- This was a significant difference ($F=18.19$, $p<.001$) between the answers of students.



FIRST2 NETWORK DATA

Which choice best expresses your feelings about the statement, "I knew many college-educated adults when I was growing up"?
(Options range from 1: strongly disagree to 5:strongly agree.)

First-generation students' mean was 2.50.
Continuing generation students' mean was 3.84.

This was a significant difference ($F=28.53$, $p<.001$) between the answers of first-generation and continuing generation students.

There was not a significant difference between answers of rural and non-rural students.



FIRST2 NETWORK DATA

Which choice best expresses your feelings about the statement, "I found the transition from high school to college to be stressful"?(
Options range from 1: strongly disagree to 5:strongly agree.)

First-generation students' mean was 3.81.
Continuing generation students' mean was 3.19.

This was a significant difference ($F=6.409$, $p<.05$) between the answers of first-generation and continuing generation students.

There was not a significant difference between answers of rural and non-rural students.

