Aug 17, 2012   
LSAY LCA Project Report

Kyusang Park

1. Reverse coding of all the indicator variables to represent higher means as pro-math and pro-science disposition:

Math as a Subject Matter Items (From Cohort 2, Year 1, Fall 1987)

AB39A = I enjoy math

AB39H = Math is useful in everyday problems

AB39I = Math helps a person think logically

AB39K = It is important to know math to get a good job

AB39L = I will use math in many ways as an adult

Science as a Subject Matter Items (From Cohort 2, Year 1, Fall 1987)

AB39M = I enjoy science

AB39T = Science is useful in everyday problems

AB39U = Science helps a person think logically

AB39W = It is important to know science to get a good job

AB39X = I will use science in many ways as an adult

1 (strongly agree) 🡪 5

2 (agree) 🡪 4

3 (not sure) 🡪 3

4 (disagree) 🡪 2

5 (strongly disagree) 🡪 1

1. Recoding of distal outcome variables by dichotomizing the original distal outcome variable of RSTEMMA:

STEMJOB, 0 (Not in Workforce) 🡪 0

1 (Non-STEMM) 🡪 0

2 (STEMM Support) 🡪 0

3 (STEMM) 🡪 1

Proportions, 0 = 92.51% (n=3411)

1 = 7.49% (n=276)

STEMJOB1, 0 (Not in Workforce ) 🡪 0

1 (Non-STEMM) 🡪 0

2 (STEMM Support) 🡪 1

3 (STEMM) 🡪 1

Proportions, 0 = 85.82% (n=3164)

1 = 14.18% (n=523)

1. Recoding of distal outcome variables by dichotomizing the original distal outcome variable of RSTEMMX:

ENGJOB, 11 (scientists) 🡪 0

12 (mathematicians) 🡪 0

14 (social scientists) 🡪 0

15 (health professionals I) 🡪 0

16 (other scientific/technical professionals) 🡪 0

17 (law and related) 🡪 0

18 (other professionals) 🡪 0

19 (social scientists level II) 🡪 0

20 (secondary math/sci teachers) 🡪 0

21 (health professionals II) 🡪 0

22 (business) 🡪 0

23 (public administration) 🡪 0

24 (technicians/technologists) 🡪 0

25 (teachers) 🡪 0

31 (sales) 🡪 0

32 (clerical) 🡪 0

33 (crafts) 🡪 0

41 (operatives) 🡪 0

42 (transportation) 🡪 0

43 (protective services) 🡪 0

51 (labors) 🡪 0

52 (service) 🡪 0

61 (homemaker) 🡪 0

62 (student) 🡪 0

63 (retired, unemployed, out of workforce) 🡪 0

97 (other occ, non-stemm) 🡪 0

13 (engineers) 🡪 1

Proportions, 0 = 97.67% (n=3601)

1 = 2.33% (n=86)

1. Recoding of a covariate variable by dichotomizing RACETH:

RACECOV, 3 (White) 🡪 0

4 (Asian) 🡪 0

1 (Hispanic) 🡪 1

2 (Black) 🡪 1

5 (Native American) 🡪 1

6 (Other) 🡪 1

Proportions, 0 = 77.66% (n=4313)

1 = 22.34% (n=1241)

1. Final LCA model:

Sample size = 2861

Number of indicator variables = 10 (AB39A, AB39H, AB39I, AB39K, AB39L, AB39M, AB39T, AB39U, AB39W, AB39X)

Number of covariate variables = 3 (AMTHIRT, ASCIIRT, RACECOV)

Number of distal outcome variables = 5 (STEMJOB, STEMJOB1, ENGJOB, CMTHIRT, CSCIIRT)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AB39A | AB39H | AB39I | AB39K | AB39L | AB39M | AB39T | AB39U | AB39W | AB39X |

AMTHIRT

ASCIIRT

RACECOV

STEMJOB

STEMJOB1

ENGJOB

CMTHIRT

CSCIIRT

c

1. Results of LCA with covariates

* Categorical latent variable
* Continuous covariates (AMTHIRT, ASCIIRT) and categorical covariate (RACECOV) with indirect effects on Ys through C which can be thought of as predictors of class membership
* This modeling allows us to understand how different classes relate to students’ background information such as academic achievement and demographics

1. Class enumeration

Investigating the unconditional model, 5-class solution was chosen for the final model.

1. Overall picture

* Sample size = 2861
* Log-likelihood = -37680.13
* Number of parameters = 76
* BIC = 75965.14
* ABIC = 75723.66
* Entropy = 0.804
* VLMR = 0.0479
* BLRT = 0

1. Understanding latent class regression

Estimate S.E. Est./S.E. P-Value

Categorical Latent Variables

C#1 ON

AMTHIRT -0.008 0.019 -0.443 0.658

ASCIIRT -0.056 0.018 -3.033 0.002

RACECOV -0.648 0.381 -1.703 0.089

C#2 ON

AMTHIRT -0.027 0.010 -2.817 0.005

ASCIIRT -0.019 0.009 -2.214 0.027

RACECOV -0.127 0.151 -0.847 0.397

C#3 ON

AMTHIRT 0.006 0.011 0.573 0.566

ASCIIRT -0.034 0.010 -3.279 0.001

RACECOV -0.344 0.203 -1.691 0.091

C#4 ON

AMTHIRT -0.002 0.009 -0.265 0.791

ASCIIRT 0.010 0.009 1.072 0.284

RACECOV 0.470 0.151 3.117 0.002

Intercepts

C#1 0.689 0.748 0.922 0.357

C#2 1.886 0.367 5.140 0.000

C#3 0.104 0.411 0.253 0.800

C#4 -1.262 0.430 -2.936 0.003

Interpretation of regression coefficients:

* Given membership in either class 1 or 5, students with higher science achievement scores are more likely to be in class 1 than students with lower scores. (p < .01)
* Given membership in either class 2 or 5, students with higher math achievement scores are more likely to be in class 2 than students with lower scores. (p < .01)
* Given membership in either class 2 or 5, students with higher science achievement scores are more likely to be in class 2 than students with lower scores. (p < .05)
* Given membership in either class 3 or 5, students with higher science achievement scores are more likely to be in class 2 than students with lower scores. (p < .01)
* Given membership in either class 4 or 5, students who are neither White nor Asian are more likely to be in class 4 than students who are either White or Asian. (p < .01)

1. Understanding LCA with distal outcome variables

STEMJOB

Mean S.E. Mean S.E.

Class 1 0.000 0.000 Class 2 0.042 0.011

Class 3 0.066 0.019 Class 4 0.134 0.019

Class 5 0.091 0.011

Chi-Square P-Value Chi-Square P-Value

Overall test 45.011 0.000 Class 1 vs. 2 14.372 0.000

Class 1 vs. 3 11.890 0.001 Class 1 vs. 4 49.803 0.000

Class 1 vs. 5 64.304 0.000 Class 2 vs. 3 1.113 0.291

Class 2 vs. 4 17.703 0.000 Class 2 vs. 5 8.490 0.004

Class 3 vs. 4 6.348 0.012 Class 3 vs. 5 1.291 0.256

Class 4 vs. 5 3.433 0.064

STEMJOB1

Mean S.E. Mean S.E.

Class 1 0.121 0.048 Class 2 0.109 0.016

Class 3 0.112 0.025 Class 4 0.193 0.022

Class 5 0.176 0.015

Chi-Square P-Value Chi-Square P-Value

Overall test 6.530 0.163 Class 1 vs. 2 0.051 0.821

Class 1 vs. 3 0.027 0.870 Class 1 vs. 4 1.851 0.174

Class 1 vs. 5 1.158 0.282 Class 2 vs. 3 0.008 0.930

Class 2 vs. 4 9.591 0.002 Class 2 vs. 5 7.924 0.005

Class 3 vs. 4 6.055 0.014 Class 3 vs. 5 4.499 0.034

Class 4 vs. 5 0.402 0.526

ENGJOB

Mean S.E. Mean S.E.

Class 1 0.000 0.000 Class 2 0.008 0.005

Class 3 0.022 0.011 Class 4 0.058 0.013

Class 5 0.021 0.006

Chi-Square P-Value Chi-Square P-Value

Overall test 17.860 0.001 Class 1 vs. 2 2.419 0.120

Class 1 vs. 3 3.684 0.055 Class 1 vs. 4 20.120 0.000

Class 1 vs. 5 13.309 0.000 Class 2 vs. 3 1.209 0.271

Class 2 vs. 4 13.097 0.000 Class 2 vs. 5 2.668 0.102

Class 3 vs. 4 4.335 0.037 Class 3 vs. 5 0.007 0.935

Class 4 vs. 5 6.392 0.011

CMTHIRT

Mean S.E. Mean S.E.

Class 1 51.222 1.305 Class 2 50.912 0.471

Class 3 53.529 0.690 Class 4 55.599 0.563

Class 5 55.267 0.379

Chi-Square P-Value Chi-Square P-Value

Overall test 19.161 0.001 Class 1 vs. 2 0.049 0.825

Class 1 vs. 3 2.423 0.120 Class 1 vs. 4 9.508 0.002

Class 1 vs. 5 8.876 0.003 Class 2 vs. 3 9.307 0.002

Class 2 vs. 4 40.651 0.000 Class 2 vs. 5 47.696 0.000

Class 3 vs. 4 5.483 0.019 Class 3 vs. 5 4.703 0.030

Class 4 vs. 5 0.222 0.638

CSCIIRT

Mean S.E. Mean S.E.

Class 1 50.858 1.322 Class 2 51.536 0.484

Class 3 54.199 0.696 Class 4 55.578 0.580

Class 5 55.668 0.372

Chi-Square P-Value Chi-Square P-Value

Overall test 13.026 0.011 Class 1 vs. 2 0.228 0.633

Class 1 vs. 3 4.930 0.026 Class 1 vs. 4 10.839 0.001

Class 1 vs. 5 12.183 0.000 Class 2 vs. 3 9.528 0.002

Class 2 vs. 4 28.474 0.000 Class 2 vs. 5 41.803 0.000

Class 3 vs. 4 2.262 0.133 Class 3 vs. 5 3.402 0.065

Class 4 vs. 5 0.016 0.900

Class 4 showed the highest mean values of STEMJOB, STEMJOB1, ENGJOB, CMTHIRT and CSCIIRT. Overall tests of mean differences were all significant ( p < .01) for STEMJOB, ENGJOB, CMTHIRT, and CSCIIRT, except for STEMJOB1 ( p > .05)