Battle of the conjunctions: Disjunctive vs. compensatory course placement

Terrence Willett, Cabrillo College
Mallory Newell, De Anza College
Loris Fagioli, Irvine Valley College

November 17, 2016

http://www.rpgroup.org/projects/multiple-measures-assessment-project
Agenda

• Project Overview
• Research Basis
  • Impact Analysis and Relationship to Equity
• Comparing different ways to combine test and non-test data and information in placement Systems
• Integration with Common Assessment Initiative (CAI)
Project Overview

**Collaboration**
- CAI
- CCCCCO
- Cal-PASS+
- RP Group
- 60 CCCs

**Model Development**
- English
- Math
- ESL
- Reading
- Non-cognitive Variables
- Self-reported transcript data

**Engagement**
- Local replication
- Webinars
- Professional development
- Support
- Pilot results inform statewide implementation
Growing body of evidence

• **Weak relationship** between assessment tests and college course outcomes: bit.ly/CCRCAssessment

• **Incredible variability** in cut scores; CCCs often use HIGHER cut scores than 4-year institutions: bit.ly/NAGB2012

• **Underestimates** students of color, women, first generation college students, low SES: bit.ly/DefiningPromise

• Long thread of research in the CCCs
  – Hetts, Fuenmayor, & Rothstein, 2012 http://www.lbcc.edu/PromisePathways
Why Multiple Measures?

- Tests used in isolation have been under-placing students
- Multiple measures
  - provides a more complete picture of student ability
  - provides a way to increase the accuracy of placement, particularly reducing underplacement
  
  - are required by law (Title V)
  - supported by statewide senate
Methods

• Matched data from high schools and community colleges in CalPASS Plus
• Recursive decision trees with Poisson model
• Rules and R code:
  
  http://rpgroup.org/projects/multiple-measures-assessment-project/decision-rules

  #load a data file into R
  setwd("C:/Users/MyName/Documents/ProjectFolder")

  MyData <- read.delim("DataFile.txt", quote = ",", row.names = NULL, stringsAsFactors = FALSE)
Variables Explored in the Models

• High School Cumulative GPA (primary predictor)
• Grades in high school courses
• CST scores
• Advanced Placement course taking
• Taking higher level courses (math)
• Delay between HS and CCC (math)
• Type of English or Math course
Transfer Level English

Diagram:

- If HS_12_GPA_CUM >= 2.6:
  - If HS_12_GPA_CUM >= 2.1:
    - 0.42 (12%)
    - 0.59 (26%)
  - If HS_12_GPA_CUM >= 3:
    - 0.73 (27%)
    - 0.87 (36%)

- If HS_12_GPA_CUM < 2.6:
  - No path
## Transfer Level MMAP Rule Sets

<table>
<thead>
<tr>
<th>Transfer Level Course</th>
<th>Direct Matriculant</th>
<th>Non-Direct Matriculant</th>
</tr>
</thead>
<tbody>
<tr>
<td>College Algebra (STEM)</td>
<td>HS 11 GPA &gt;=3.2 OR HS 11 GPA &gt;=2.9 AND Pre-Calculus C (or better)</td>
<td>HS 12 GPA &gt;=3.2 OR HS 12 GPA &gt;=3.0 AND Pre-Calculus or Statistics (C or better)</td>
</tr>
<tr>
<td><em>(and high school Alg II recommended)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics (Non STEM)</td>
<td>HS 11 GPA &gt;=3.0 OR HS 11 GPA &gt;=2.3 AND Pre-Calculus C (or better)</td>
<td>HS 12 GPA &gt;=3.0 OR HS 12 GPA &gt;=2.6 AND Pre-Calculus C (or better)</td>
</tr>
<tr>
<td><em>(and high school Alg I recommended)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>HS 11 GPA &gt;=2.6</td>
<td>HS 12 GPA &gt;=2.6</td>
</tr>
<tr>
<td>Reading</td>
<td>HS 11 GPA &gt;=2.7</td>
<td>HS 12 GPA &gt;=2.8</td>
</tr>
<tr>
<td>ESL</td>
<td>HS 11 GPA &gt;=2.7</td>
<td>HS 12 GPA &gt;=2.6</td>
</tr>
</tbody>
</table>
Various Placement Systems and Their Impact on Student Equity
What are some possible placement systems?

**Disjunctive placement:**
- Take the highest placement of multiple measures
  - i.e. Test or High School (HS) Transcripts or AP score or EAP or...
- Recommended by MMAP

**Compensatory placement:**
- Combination of all multiple measures with equal or varying weights
  - i.e. Placement = Test + HS GPA + HS Course + AP score + ...

**Conjunctive placement:**
- Lowest placement where all measures agree
  - i.e. exceed both Test score threshold and HS GPA criteria
- Highly restrictive
  - Not recommended by the CCCCCO

Special thanks to Dr. Barry Gribbons of College of the Canyons for first highlighting these systems to the MMAP Team
How can we compare these systems?

• **Accuracy**: The proportion of students who are correctly predicted to be successful or to be unsuccessful.

• **Other Classification Metrics**: Positive predictive value, Sensitivity, Specificity, etc.

• **1 year throughput rate**: The number of students successfully completing the gatekeeper course at the end of a course sequence divided by the number of students in the initial cohort within 1 year.

• **Underrepresented Minority Placement Rate**: Equity and disproportionate impact are major considerations when evaluating the performance of placement systems.
Classification metrics

- **Accuracy**: proportion of students correctly predicted to be successful or to be unsuccessful = \((TP+TN)/(TP+FP+TN+FN)\)
- **PPV**: Positive predictive value, the number of passing students (i.e., true positives) divided by the number of students predicted to succeed = \(TP/(TP+FP)\)
- **NPV**: Negative predictive value = \(TN/(TN+FN)\)
- **Specificity**: \(TN/(TN+FP) = 1\) - Type I error = True Positive Rate
- **Sensitivity**: \(TP/(TP+FN) = 1\) - Type II error = Power = \(1\) - False Positive Rate

TP=True Positive, FP=False Positive
TN=True Negative, FN=False Negative
Information that can be used to evaluate placement systems. Two way contingency table or “confusion matrix”.

<table>
<thead>
<tr>
<th></th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actually Failed</td>
<td><strong>True Negative</strong></td>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Actually Passed</td>
<td><strong>False Negative</strong></td>
<td><strong>True Positive</strong></td>
</tr>
</tbody>
</table>
Information that can be used to evaluate placement systems

<table>
<thead>
<tr>
<th>Actual Outcome</th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actually Failed</td>
<td><strong>True Negative</strong></td>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Actually Passed</td>
<td><strong>False Negative</strong></td>
<td><strong>True Positive</strong></td>
</tr>
</tbody>
</table>

Accuracy = \( \frac{TP + TN}{TP + FP + TN + FN} \)
Information that can be used to evaluate placement systems

\[ PPV = \frac{TP}{TP+FP} \]

<table>
<thead>
<tr>
<th>Actually Failed</th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Negative</td>
<td>False Positive</td>
<td>True Positive</td>
</tr>
<tr>
<td>False Negative</td>
<td>False Positive</td>
<td>True Positive</td>
</tr>
</tbody>
</table>

PPV
Information that can be used to evaluate placement systems

NPV = TN/(TN+FN)

<table>
<thead>
<tr>
<th></th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actually Failed</td>
<td><strong>True Negative</strong></td>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Actually Passed</td>
<td><strong>False Negative</strong></td>
<td><strong>True Positive</strong></td>
</tr>
</tbody>
</table>
Information that can be used to evaluate placement systems

**Specificity** = \( \frac{TN}{(TN + FP)} \)

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actually Failed</td>
<td><strong>True Negative</strong></td>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Actually Passed</td>
<td><strong>False Negative</strong></td>
<td><strong>True Positive</strong></td>
</tr>
</tbody>
</table>
Information that can be used to evaluate placement systems

Sensitivity = $\frac{TP}{TP+FN}$

<table>
<thead>
<tr>
<th></th>
<th>Predicted to Fail</th>
<th>Predicted to Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actually Failed</td>
<td><strong>True Negative</strong></td>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Actually Passed</td>
<td><strong>False Negative</strong></td>
<td><strong>True Positive</strong></td>
</tr>
</tbody>
</table>

Sensitivity
Response Operating Curve

ROC
- random
- good
- better
- best

True positive rate
False positive rate
R Resources for Model Metrics

• “caret” package for testing many models
• "e1071" corrects some errors in caret
• "pROC" draws ROC curves
• Max Kuhn caret site and webinar slides
Managing Errors

- Typically have trade off in specificity v. sensitivity and must consider consequences of false positives v. false negatives.
- High Sensitivity
  - Airport security
  - Allow students a chance to pass a course
- High Specificity
  - Convicting someone of a serious crime
  - Protect students from failing a course
Types of Placement Error

- **Overplacement**: Student is placed above their ability to succeed. Highly visible.
- **Underplacement**: Student could have been successful at a higher level than where placed. Tends to be invisible.
- Current placement systems tend to result in much greater underplacement error.
- Total placement error is minimized when over- and underplacement are balanced.
- Consequences to students of each error not equal
Evaluating Placement Systems

**Disjunctive placement:**
- Take the highest placement (Test or MMAP)
- Recommended by MMAP

**Compensatory placement:**
- Logistic regression (combines Test, MMAP simultaneously)
- Run with two cut-values: 0.70, 0.50

**Conjunctive placement:**
- Only if Test and MMAP in agreement
- Highly restrictive
- Not recommended by the CCCCCO
Accuracy: Statistics Course

Accurate Placement in College Statistics

- Conjunctive: 61%
- Compensatory (0.70): 57.40%
- Disjunctive (0.70)*: 67%
- Compensatory (0.50): 68.10%
Positive Predicted Value (PPV): Statistics Course

PPV for College Statistics by Placement System

73%  77.90%  65%  69.70%

Proportion Correctly Predicted to Pass

- Conjunctive
- Compensatory (0.70)
- Disjunctive (0.70)
- Compensatory (0.50)
One Year Throughput Rate: College Statistics Course

Statistics Class Throughput rate by Placement System

Percent placed accurately:
- Conjunctive
- Compensatory (0.70)
- Disjunctive (0.70)
- Compensatory (0.50)
Percentage of Transfer-placed Students who are URM

Percentage Transfer-placed Students who are URM by Placement System

- Conjunctive
- Compensatory (0.70)
- Disjunctive (0.70)
- Compensatory (0.50)
Placement of Under-represented Minorities into Transfer English by Assessment Model

- Original: 40%
- Conjunctive: 27%
- Disjunctive: 49%
- Compensatory: 34%
Summary of Placement Models

• No single metric is sufficient but several well-chosen metrics can allow for a more informed decision.
• Throughput is an important metric to consider.
• PPV can be calculated for all placement systems; metrics that require a True Negative cannot be calculated for disjunctive placement systems.
• When requiring >70% probability of passing transfer-level course, disjunctive models have higher access and throughput than compensatory models.
• If compensatory model is set to a 0.50 criterion or cut-value, it can outperform a disjunctive model (with a .70 criterion) in terms of accuracy, access, PPV and throughput.
• The conjunctive model was very restrictive and had the lowest throughput rates and URM placement rates.
Interactive rCharts

http://ramnathv.github.io/rCharts/
Integration of MMAP with CAI
Integration of MMAP with CAI

• Note: Common Assessment updates currently on pause
• Common Assessment platform will house a transcript data repository
  – repository will be source-agnostic & store transcript data from variety of sources, including CalPASS & self-report via CCC Apply
  – statewide decision trees programmed into platform, for internally generated Multiple Measures placement recommendation
  – expect data points used in MM placement recommendation
• Students will receive a single placement recommendation created from a disjunctive placement model
Contacts:

John Hetts
Educational Results Partnership
jhetts@edresults.org

Ken Sorey
Educational Results Partnership
ksorey@edresults.org

Dan Lamoree
Educational Results Partnership
dl amoree@edresults.org

Rachel Baker
UC Irvine
rachelbb@uci.edu

Mallory Newell
The RP Group
newellmallory@fhda.edu

Loris Fagioli
The RP Group
lfagioli@rpgroup.org

Terrence Willett
The RP Group
twillett@rpgroup.org

Craig Hayward
The RP Group
chayward@rpgroup.org
Impact Analysis and Relationship to Equity
Retrospective Impact Analysis by College

- **4 graphics available** for each college and overall based on disjunctive model for Math and English
  - Overall
  - Overall with range
  - By Ethnicity
  - By Ethnicity with range
    - Range = Expected estimates vary depending on data availability

- **Limitations:**
  - Does not include self-reported data
  - Based on students with Cal-PASSPlus data
  - Retrospective projections
  - Placement vs Enrollment
Math and English Transfer Placement by Ethnicity - IVC
### MMAP Prospective Projections South Orange County District (example)

<table>
<thead>
<tr>
<th>Site</th>
<th>Subject</th>
<th>Level</th>
<th>Enrollment Fall 2016</th>
<th>Enrollment Fall 2015</th>
<th>Placement Fall 2016</th>
<th>Placement Fall 2015</th>
<th>Historic Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irvine</td>
<td>Math</td>
<td>-3</td>
<td>127</td>
<td>181</td>
<td>319</td>
<td>421</td>
<td>40.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>406</td>
<td>417</td>
<td>713</td>
<td>788</td>
<td>56.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>353</td>
<td>428</td>
<td>473</td>
<td>573</td>
<td>74.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1090</td>
<td>855</td>
<td>1331</td>
<td>1054</td>
<td>81.9%</td>
</tr>
<tr>
<td>Irvine</td>
<td>English</td>
<td>-3</td>
<td>184</td>
<td>305</td>
<td>351</td>
<td>607</td>
<td>52.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>279</td>
<td>420</td>
<td>442</td>
<td>673</td>
<td>63.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>245</td>
<td>323</td>
<td>374</td>
<td>518</td>
<td>65.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1137</td>
<td>664</td>
<td>1516</td>
<td>877</td>
<td>75.0%</td>
</tr>
<tr>
<td>Saddleback</td>
<td>Math</td>
<td>-3</td>
<td>193</td>
<td>290</td>
<td>653</td>
<td>920</td>
<td>29.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>477</td>
<td>459</td>
<td>869</td>
<td>846</td>
<td>54.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>375</td>
<td>454</td>
<td>506</td>
<td>611</td>
<td>74.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>861</td>
<td>583</td>
<td>1098</td>
<td>749</td>
<td>78.4%</td>
</tr>
<tr>
<td>Saddleback</td>
<td>English</td>
<td>-3</td>
<td>12</td>
<td>12</td>
<td>50</td>
<td>50</td>
<td>23.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>444</td>
<td>746</td>
<td>904</td>
<td>1439</td>
<td>49.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>279</td>
<td>365</td>
<td>471</td>
<td>582</td>
<td>59.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1250</td>
<td>773</td>
<td>1648</td>
<td>998</td>
<td>75.8%</td>
</tr>
</tbody>
</table>

Notes: Levels: -3 = three levels below transfer, -2: two levels below transfer ... 0 = transfer level.
1 Enrollment projections of new students for a disjunctive MMAP model (Fall 2018). Projections are weighted by historic enrollment rates (see below). MMAP projections assume information available for all students (students with missing data were projected at equal proportions to students with complete information). Thus, these numbers estimate the maximum expected impact/enrollment.
2 Actual enrollments for past fall term (Fall 2015)
3 Historic Weight = Percent of students who enroll in course out of all students who were placed into course (5 year average)
Pilot College Examples
Sierra College
College-Level English

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2011</td>
<td>72%</td>
</tr>
<tr>
<td>Fall 2012</td>
<td>73%</td>
</tr>
<tr>
<td>Fall 2013</td>
<td>70%</td>
</tr>
<tr>
<td>Fall 14 - Accuplacer</td>
<td>73%</td>
</tr>
<tr>
<td>F14 HS Data</td>
<td>79%</td>
</tr>
</tbody>
</table>
Cañada College

- English Transfer: N = 66, 75%
- Math Transfer: N = 70, 68%
Bakersfield College

- Engl Transfer: 69% (Test placed) vs. 77% (MIH placed)
- Math Transfer: 60% (Test placed) vs. 68% (MIH placed)
San Diego District

Pass rates in transfer level Math

<table>
<thead>
<tr>
<th>Course</th>
<th>District pass rates (Spring 2015)</th>
<th>MMAP pass rates (entire sample) - Fall 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>College Algebra</td>
<td>58</td>
<td>67</td>
</tr>
<tr>
<td>Statistics</td>
<td>64</td>
<td>61</td>
</tr>
<tr>
<td>Trigonometry (Pre-calc)</td>
<td>51</td>
<td>63</td>
</tr>
<tr>
<td>Liberal Arts Math</td>
<td>67</td>
<td>67</td>
</tr>
</tbody>
</table>
Self-Reported Transcript Data
Potential use of self-reported high school info

- UC admissions uses self-report but verifies after admission
  - 2008: 9 campuses, 60,000 students. No campus had >5 discrepancies between reported grades and student transcripts: http://bit.ly/UCSelfReportGPA


- ACT research often uses self-reported GPA, generally find it to be a highly powerful predictor and highly correlated with students actual GPA: ACT, 2013: r(1978) = .84 http://bit.ly/ACTSRGPA
# GPA vs. Self-reported HSGPA

**ACT, 2013:**  

<table>
<thead>
<tr>
<th>HSGPA Level</th>
<th>N</th>
<th>Mean HSGPA</th>
<th>Mean diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Self-reported</td>
<td></td>
</tr>
<tr>
<td>3.50–4.00</td>
<td>599</td>
<td>3.79</td>
<td>3.75</td>
</tr>
<tr>
<td>3.00–3.49</td>
<td>451</td>
<td>3.24</td>
<td>3.23</td>
</tr>
<tr>
<td>2.50–2.99</td>
<td>408</td>
<td>2.81</td>
<td>2.76</td>
</tr>
<tr>
<td>2.00–2.49</td>
<td>265</td>
<td>2.24</td>
<td>2.35</td>
</tr>
<tr>
<td>1.50–1.99</td>
<td>172</td>
<td>1.77</td>
<td>2.04</td>
</tr>
<tr>
<td>0.00–1.49</td>
<td>85</td>
<td>1.03</td>
<td>1.85</td>
</tr>
<tr>
<td>Total</td>
<td>1,980</td>
<td>2.95</td>
<td>3.02</td>
</tr>
</tbody>
</table>

**College Board, 2009:**  

Under-reporting was 2-4X as common as over-reporting.
Using Self-Reported GPA from CCC Apply

• New **optional** items included in Open CCCApply
  – Grade Point Average
  – Highest English Course Taken
  – Highest English Course Taken Grade
  – Highest Math Course Taken
  – Highest Math Course Taken Grade
  – Highest Math Course Passed
  – Highest Math Course Passed Grade

• Need to opt-in!
  – Contact CCCAssess Product Manager, John Hadad, [jhadad@ccctechcenter.org](mailto:jhadad@ccctechcenter.org) to opt-in
  – Requesting all MMAP pilot colleges to opt-in and share data with the MMAP team for validation
Self-reported data - Plan for use

• Fully test reliability of self-reported high school info, particularly using the CCCApply questions
• Use as backup/supplement to actual high school transcript data for:
  – Students from high schools not participating in CalPASS Plus or with gaps in CalPASS Plus participation for that student
  – Students from out-of-state
  – Possible use of information from senior year for direct matriculants