Academic Support Services and Student Success
Do They Make a Difference?

*****

http://www.unr.edu/ia/research

Serge Herzog, PhD
Director, Institutional Analysis
Consultant, CRDA StatLab
University of Nevada
Reno, NV 89557, serge@unr.edu

September 2014
Purpose of Study

• Estimate the effect of campus tutoring* services on student
  – First-semester academic success (academic momentum)
  – Enrollment persistence into second semester
  – Enrollment persistence into second year (fall-to-fall retention)
• Measure estimation bias associated with student self-selection

*Use of any of three campus-based centers
Prior Research

• Student engagement has been defined as “participation in educationally effective practices, both inside and outside the classroom…” (Kuh et al., 2007)

• Student engagement research is based almost entirely on student self-reported survey responses (e.g., NSSE, CIRP, UCUES), the validity of which has come under mounting criticism (Porter, 2011; Bowman, 2011; Herzog & Bowman, 2011; Pascarella, Seifert, & Blaich, 2010)

• Call for analytically rigorous research by AERA (Schneider et al., 2007) and USDE’s Institute of Ed Sci. and its What Works Clearinghouse (http://ies.ed.gov/ncee/wwc/ReviewedStudies.aspx)
Approach in this Study

- Rely on empirically verifiable direct measures of student engagement (i.e., use of campus tutoring services) via actuarial records and office logs
- Employ a non(quasi)-experimental design to reduce selection bias associated with student use of tutoring services
- Test the counterfactual hypothesis, “What would have been the effect on academic outcomes had the tutored students not received any campus tutoring service?”
- Gauge the selection bias by comparing naïve estimators with those from data that attempts to mimic randomization
Data Source and Sample

- Fall new freshman cohorts 2011 and 2012 at U. of Nevada, excluding
  - Students with less than 12 attempted credits (PT)
  - Students with complete credit withdrawal after registration
  - Students without entry survey data
  - Statistical outliers (using Cook’s, Mahalanobis’)

- Effective sample: 4,727 freshmen (~85% of cohorts)

- Computed variables
  - Precollege preparation index (GPA-test score composite)
  - Delayed college entry: months from HS graduation to college
  - Academic momentum: 100-pt index (GPA, credits earned)
  - Imputation (mean, predicted value) of missing values
Causal Inference Estimation

- Randomized Control Trials (RCT)
  - Costly, ethical issues, operational difficulties, disruption of natural setting
  - Replication not possible with RCT

- Nonexperimental designs in observational studies
  - Regression discontinuity
  - Instrumental variable (IV) techniques
  - Econometric models to adjust for selection bias
  - Propensity score methods
    - Inverse probability of treatment weighting (IPTW)
    - Subclass stratification
    - Propensity score matching (PSM)
    - Regression covariate adjustment (linear/nonlinear)
Causal Inference Estimation

- **Propensity Scoring: Rubin Causal Model**
  - Two potential outcomes: 
    \[ Y_i(1) \text{ and } Y_i(0) \text{ and let } Z \text{ denote tutoring (1) or no tutoring (0)} \]
  - But only one is observed: 
    \[ Y_i(Y_i = Z_i(1) + (1 - Z_i)Y_i(0) \]
  - Thus, \( Y_i(1) - Y_i(0) = \text{effect of tutoring and } E[Y_i(1) - Y_i(0) = \text{average tutoring effect (ATE)} \]
  - Average tutoring effect (ATT) for \( Y_i = (1) \) is defined as \( E[Y(1) - Y(0)|Z = 1] \) where \( \Pr(Z_i = 1|X_i) \) where \( X_i = \text{pre-Z covariates} \)

- Propensity score (PS) accounts for covariates (characteristics) that predict treatment selection (received tutoring)

- PS captures the conditional probability of treatment selection given observed (measurable) covariates (selection on ‘observables’)

- Find non-treated who are similar to treated in all relevant pre-treatment characteristics (unconfounded, conditionally indepen.)

- Matching treated (tutored) with non-treated (not tutored) students on their PS enables ‘counterfactual’ analysis
Causal Inference Estimation

- The PS is estimated via logit (or probit, log-linear link, Mahalanobis, neural net, classification trees etc.)
- Variable selection for PS estimation should be guided by conceptual theory, prior research, and understanding of the treatment selection process
- Within-study comparison of randomized and quasi-experiment by Steiner et al. (2010):
  - Variables that best predict both treatment selection and outcome(s)
  - Measures of subject preference, motivation, and proxy-'pretests'
- Reliable PS analysis requires
  - Strongly Ignorable Treatment Assignment
  - Covariates are not affected by treatment, must be pre-treatment
  - Treated/non-treated simultaneity (balance, common support)
  - Covariate selection governed by variance-bias balance
- Analysis done in R with MatchIt, Matching, GenMatch packages
Causal Inference Estimation

- Selected variables for PS estimation include
  - Preference and motivation cluster: math-intensive major, admission test date, preference for institution, educational goal, father education level, mother education level
  - Proxy ‘pretest’ cluster: high school GPA, admission test scores (math, verbal)
  - Student demographics: age, gender, race, residency
  - Academic indicators: HS rank, undeclared, advanced standing
  - Financial aid profile: unmet need, loans, merit aid, Pell
  - Other variables: on-campus dorm res., plans to work full-time/not at all, delayed college entry
Causal Inference Estimation

- IPTW with focus on \((ATT)\) reweights PS of non-treated students
  
  \[ w(Y_i(0)) = \frac{P_i(X)}{1-P_i(X)} \]

- PS matching uses 3 algorithms within support area of treated
  - NNR: 1 nearest neighbor matching, with replacement
  - FullSubCl seeks optimal subclassification to minimize weighted distance between treated and control cases within each subclass
  - Genetic seeks set of weights for each covariate to optimize balance, with replacement to result in 1:1 match (with replacement)

- Outcome estimation
  - Regression covariate adjustment in linear model with treatment status \((y/n)\), PS distance, and post-treatment selection covariates (worked on campus, used diversity center, took no English, took no math, took a fully online course)

- Lechner’s (1999) variance approximation for SE estimates
  
  \[ \frac{1}{N_1} Var(Y(1)|D = 1) + \frac{\sum_{i(D=0)}(W_i)^2}{N_1} \ast Var(Y(0)|D = 0) \]
Prop Score Balance Check: Genetic r:1

Distribution of Propensity Scores

Unmatched Treatment Units

Matched Treatment Units

Matched Control Units

Unmatched Control Units

Propensity Score
Prop Score Balance Check: Near r:1 short

Distribution of Propensity Scores

Unmatched Treatment Units

Matched Treatment Units

Matched Control Units

Unmatched Control Units

Propensity Score
Prop Score Balance Check: Nearest r:1

Raw Treated

Matched Treated

Raw Control

Matched Control
Prop Score Balance Check: SubCI low

- Raw Treated
- Matched Treated
- Raw Control
- Matched Control
## Balance Check of PS and Covariates

**Table 1: Covariate and Propensity Score Balance for Matched Data**

<table>
<thead>
<tr>
<th></th>
<th>NNR:1</th>
<th>Full SubCl</th>
<th>Genetic</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Std. Difference in Mean</em> for:</em>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All covariates</td>
<td>0.009</td>
<td>-0.007</td>
<td>0.004</td>
</tr>
<tr>
<td><strong>Balance Improvement %</strong></td>
<td>22.6</td>
<td>-0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Motivation, educ. Goal</td>
<td>0.007</td>
<td>-0.003</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>Balance Improvement %</strong></td>
<td>68.5</td>
<td>54.4</td>
<td>12.3</td>
</tr>
<tr>
<td>Subject preference</td>
<td>0.009</td>
<td>-0.029</td>
<td>-0.010</td>
</tr>
<tr>
<td><strong>Balance Improvement %</strong></td>
<td>-35.2</td>
<td>47.9</td>
<td>-72.9</td>
</tr>
<tr>
<td>Propensity score</td>
<td>0.001</td>
<td>0.001</td>
<td>0.030</td>
</tr>
<tr>
<td><strong>Balance Improvement %</strong></td>
<td>99.8</td>
<td>99.8</td>
<td>91.6</td>
</tr>
<tr>
<td><strong>Covariates with SD &gt; 0.2</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Covariates with SD &gt; 0.1</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Covariates with SD &gt; 0.05</strong></td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

*Formula denominator uses standard deviation of treated, not pooled, unlike Cohen's d*
### Outcome Estimation

#### Average Outcomes and Naïve Estimator

<table>
<thead>
<tr>
<th></th>
<th>Non-Tutored (N=2667)</th>
<th>Tutored (N=2060)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall momentum</td>
<td>77</td>
<td>85</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>(21.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring retention</td>
<td>91%</td>
<td>96%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>(24.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall-to-fall retention</td>
<td>79%</td>
<td>87%</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>(38.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard deviation in parentheses
# Outcome Estimation

Table 2. Estimation of Academic Tutoring Effect Using Standard Regression and PS Reweighting and Matching

<table>
<thead>
<tr>
<th></th>
<th>OLS Regr</th>
<th>IPTW</th>
<th>NNR:1</th>
<th>Full SubCl*</th>
<th>Genetic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treated</strong></td>
<td>(N=2060)</td>
<td>(N=2060)</td>
<td>(N=2060)</td>
<td>(N=2060)</td>
<td>(N=2060)</td>
</tr>
<tr>
<td><strong>Untreated</strong></td>
<td>(N=2667)</td>
<td>(N=2667)</td>
<td>(N=1218)</td>
<td>(N=2667)</td>
<td>(N=1288)</td>
</tr>
<tr>
<td><strong>Fall momentum</strong></td>
<td>5.6</td>
<td>5.4</td>
<td>7.1</td>
<td>2.4</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.51)</td>
<td>(0.19)</td>
<td>(0.19)</td>
<td>(0.19)</td>
</tr>
<tr>
<td><strong>Spring retention (SR)</strong></td>
<td>0.007</td>
<td>0.008</td>
<td><strong>0.018</strong></td>
<td>0.010</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(.008)</td>
<td>(0.006)</td>
<td>(.008)</td>
</tr>
<tr>
<td><strong>SR w/o fall momentum control</strong></td>
<td><strong>0.036</strong></td>
<td>0.034</td>
<td><strong>0.050</strong></td>
<td>0.020</td>
<td><strong>0.034</strong></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(.008)</td>
</tr>
<tr>
<td><strong>Fall-to-fall retention (FF)</strong></td>
<td><strong>0.031</strong></td>
<td>0.030</td>
<td><strong>0.033</strong></td>
<td><strong>0.029</strong></td>
<td><strong>0.042</strong></td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.010)</td>
<td>(.013)</td>
<td>(.010)</td>
<td>(.012)</td>
</tr>
<tr>
<td><strong>FF w/o fall momentum control</strong></td>
<td><strong>0.069</strong></td>
<td><strong>0.066</strong></td>
<td><strong>0.079</strong></td>
<td><strong>0.045</strong></td>
<td><strong>0.073</strong></td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.011)</td>
<td>(.013)</td>
<td>(.011)</td>
<td>(.013)</td>
</tr>
</tbody>
</table>

Bold=sig at 0.05 alpha. Standard errors in parentheses are based on Lechner's formula for NNR:1 and Genetic.

*1,899 subclasses with a range of 2 to 24 control cases per subclass
# Outcome Estimation

## Table 3. Estimation of Academic Tutoring Effect on First Semester Academic Momentum

<table>
<thead>
<tr>
<th></th>
<th>OLS Regr High</th>
<th>OLS Regr Low</th>
<th>NNR:1 High</th>
<th>Full SubCl Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N=688)</td>
<td>(N=694)</td>
<td>(N=688)</td>
<td>(N=694)</td>
</tr>
<tr>
<td>Treated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated</td>
<td>(N=886)</td>
<td>(N=881)</td>
<td>(N=399)</td>
<td>(N=881)</td>
</tr>
<tr>
<td><strong>Fall momentum</strong></td>
<td><strong>2.6</strong></td>
<td><strong>8.5</strong></td>
<td><strong>1.8</strong></td>
<td><strong>7.6</strong></td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(1.1)</td>
<td>(0.28)</td>
<td>(0.69)</td>
</tr>
<tr>
<td>All covariates</td>
<td>-0.005</td>
<td>-0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Balance Improvement %</em></td>
<td>-7.1</td>
<td>-31.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation, educ. Goal</td>
<td>0.003</td>
<td>-0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Balance Improvement %</em></td>
<td>19.0</td>
<td>65.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math preference</td>
<td>0.038</td>
<td>-0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Balance Improvement %</em></td>
<td>49.8</td>
<td>58.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propensity score</td>
<td>0.002</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Balance Improvement %</em></td>
<td>99.5</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bold=sig at 0.05 alpha. Standard errors in parentheses are based on Lechner’s formula for matched data.
Findings

• Naïve estimator overstates positive effect for all outcomes compared to parametric models

• Academic momentum estimates for weighted/matched data differ from standard OLS, which overestimates the effect compared to the most balanced matched data (full subclassification)

• Standard OLS regression estimates similar effect size for spring and fall-to-fall retention compared to average PS weighted data

• Retention is largely influenced by academic momentum

• The most balanced matched data (full subcl) yield consistently the smallest effect sizes, suggesting selection bias across all tested outcomes

• Less prepared students accrue much higher academic momentum benefit from tutoring than well prepared students
Referenced Studies