



College Sports Project

John Emerson, Principal Investigator
Rachelle Brooks, Director



NORTHWESTERN
UNIVERSITY

Overview: Purpose of College Sports Project, Center for Data Collection and Analysis:

The primary purpose of the CSP database is to enable participating colleges and universities to quantify student athletes' academic outcomes in terms of their "representativeness," and to track institutional changes over time.

These data will not be used to "police" institutional behavior. They will be provided to the president who may find them helpful in meeting institutional goals. Responsible tracking of outcomes that colleges and universities care about is increasingly recognized as a "best practice" in higher education.

The Mellon Foundation's experience suggests that research is an important tool, not only for uncovering shortcomings and monitoring trends, but also for documenting and communicating successes.



Research questions and goals:

How “representative” are student athletes of their own student bodies, with regard to academic outcomes?

How do subgroups (gender, recruiting status, types of colleges) of students compare in their representativeness?

Provide useful information to college Presidents about their own institutions, some in a broader context

Provide similar data to 4 participating athletic conferences

How is the picture changing through time?

Can we “explain” the differences in academic outcomes that we observe for various institutions and student subgroups?

Do athletes achieve at a level expected, given their academic qualifications and demographic characteristics?



What data are collected?

- **Student incoming characteristics such as high school class rank, standardized test scores, sport recruitment, high school**
- **Student demographics such as gender, date of birth, citizenship, race**
- **College experiences such as athletic participation, college GPA**
- **Student identifying information: name and permanent home address**
- **These data are augmented by information from other national databases—in 2008 College Board Data about high school quality was received.**



How are data assembled and submitted?

Institutions establish a primary CPS contact, usually an IR officer, who works with admissions and athletics staff on campus to assemble the data file.

Data are submitted to Northwestern through a secure electronic site and are protected behind several firewalls and encrypted.

Once received, data are extensively cleaned by CSP staff prior to analyses to identify and correct inconsistencies between years at the student and institution level.



Data for this talk:

- **37097 students at 71 NCAA Division-III institutions**
- **6 other participating institutions lacked data for the underperformance analysis**
- **All students entering these colleges between July 1, 2005 and June 30, 2006**
- **Data collected in winter 2008 after 2 years of tracking**
- **Admissions data on this cohort collected in winter 2007**

Technical Note: Many of our analyses are in the percentile class rank scale (0 to 100), but this uses only 4pt. GPAs.



Summary of Some Key Findings

Descriptive information about this data set:

Counts and **Mean GPA** for Gender x Recruiting Status

	Non-athlete	Recruited	Non-recruited	Totals
Male	10,279 3.02	3,946 2.83	1,752 3.00	15,977 2.97
Female	16,379 3.22	2,808 3.17	1,933 3.23	21,120 3.22

Combined SAT and HS Standing for Gender x Recruiting Status

	Non-athlete	Recruited	Non-recruited	Totals
Male	1262 1.59	1179 1.34	1248 1.61	1240 1.53
Female	1231 1.86	1199 1.81	1249 1.98	1229 1.86

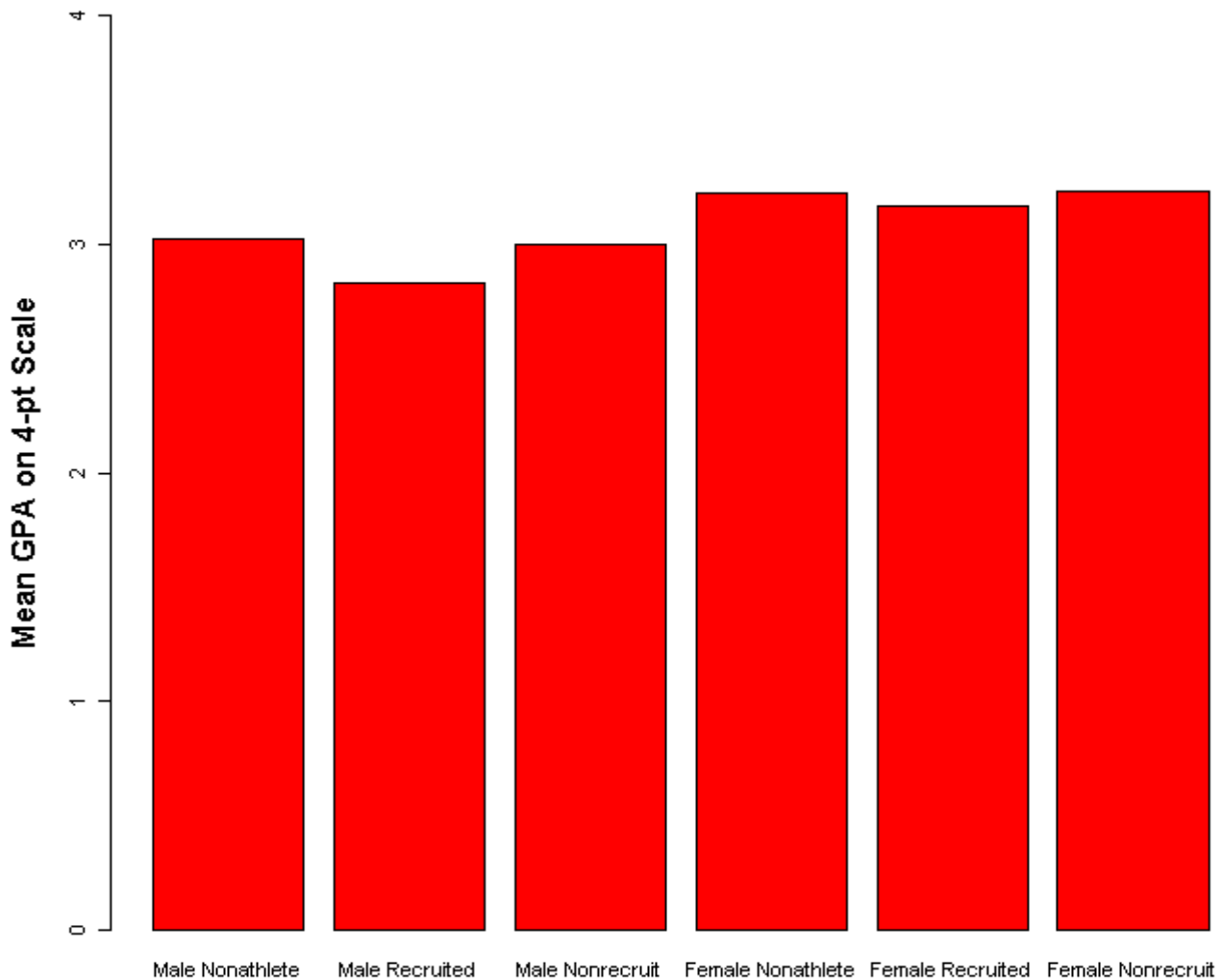
For technical reasons, HS standing is recorded as decile class rank in a logistic scale. Its only use today is as a predictor in the regression model that leads to estimates of underperformance.

Let's take a look at GPAs...



Mean GPA for Sophomore by Gender and Athlete Status

Sophomore GPA by Gender & Athlete Status, All 71 CSP Inst.



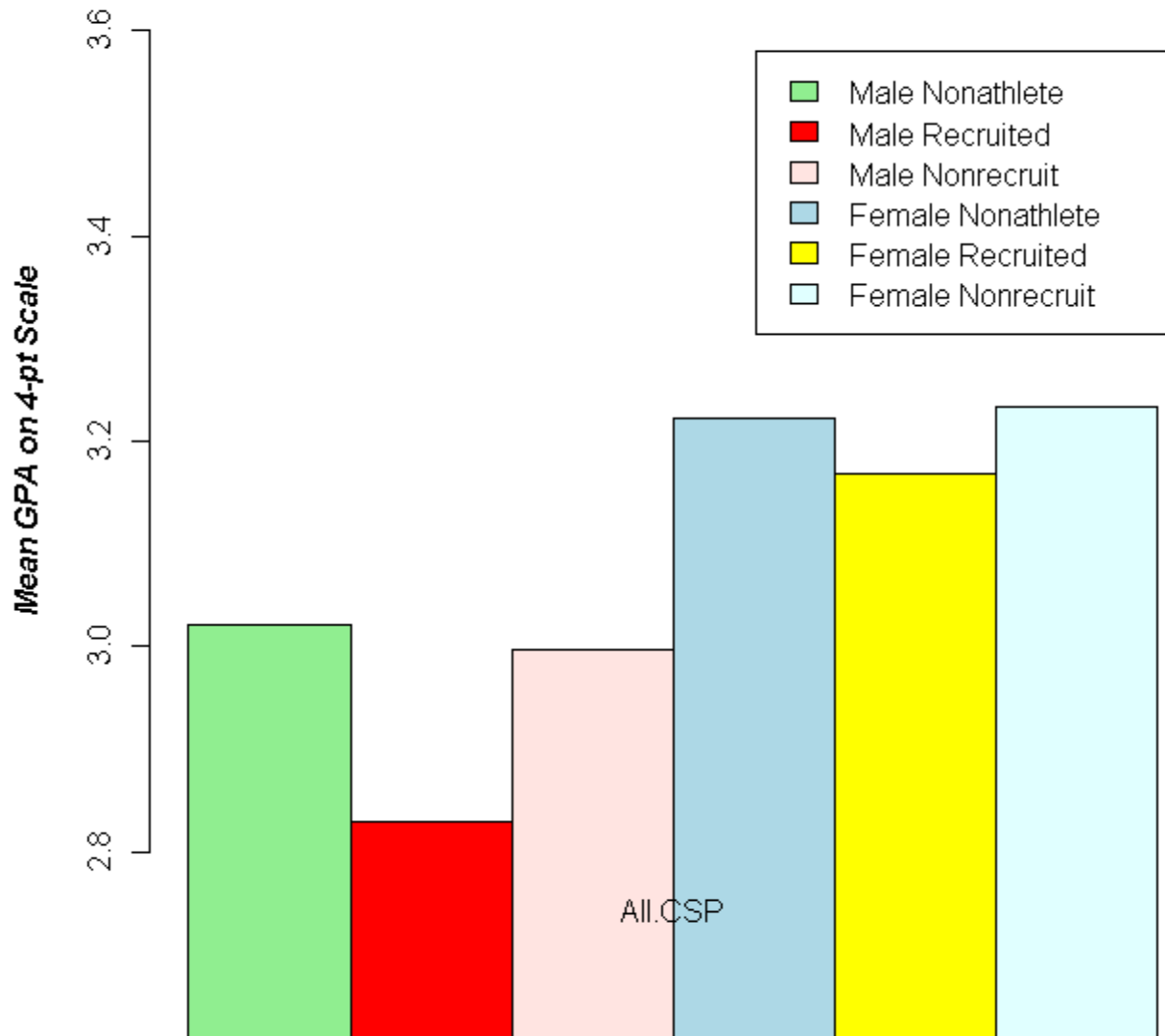
Broad Findings About Students from NCAA D-III

1. Intercollegiate athletes generally have lower college grades than non-athletes at the same institution.
2. Male students generally have lower college grades than female students, and athletes differ from non-athletes by a wider margin for males than for females.
3. The majority of athletes are recruited. The non-recruited athletes perform better academically than do the recruited athletes. Women non-recruited athletes often do as well as and sometimes better than non-athletes.
4. When students entered college, males had higher SATs but lower high school class rank than females. Recruited athletes have lower SATs, lower grades than non-athletes.

**Let's take a closer look using
graphical displays, primarily.**



Sophomore GPA by Gender & Athlete Status, 71 CSP Inst.



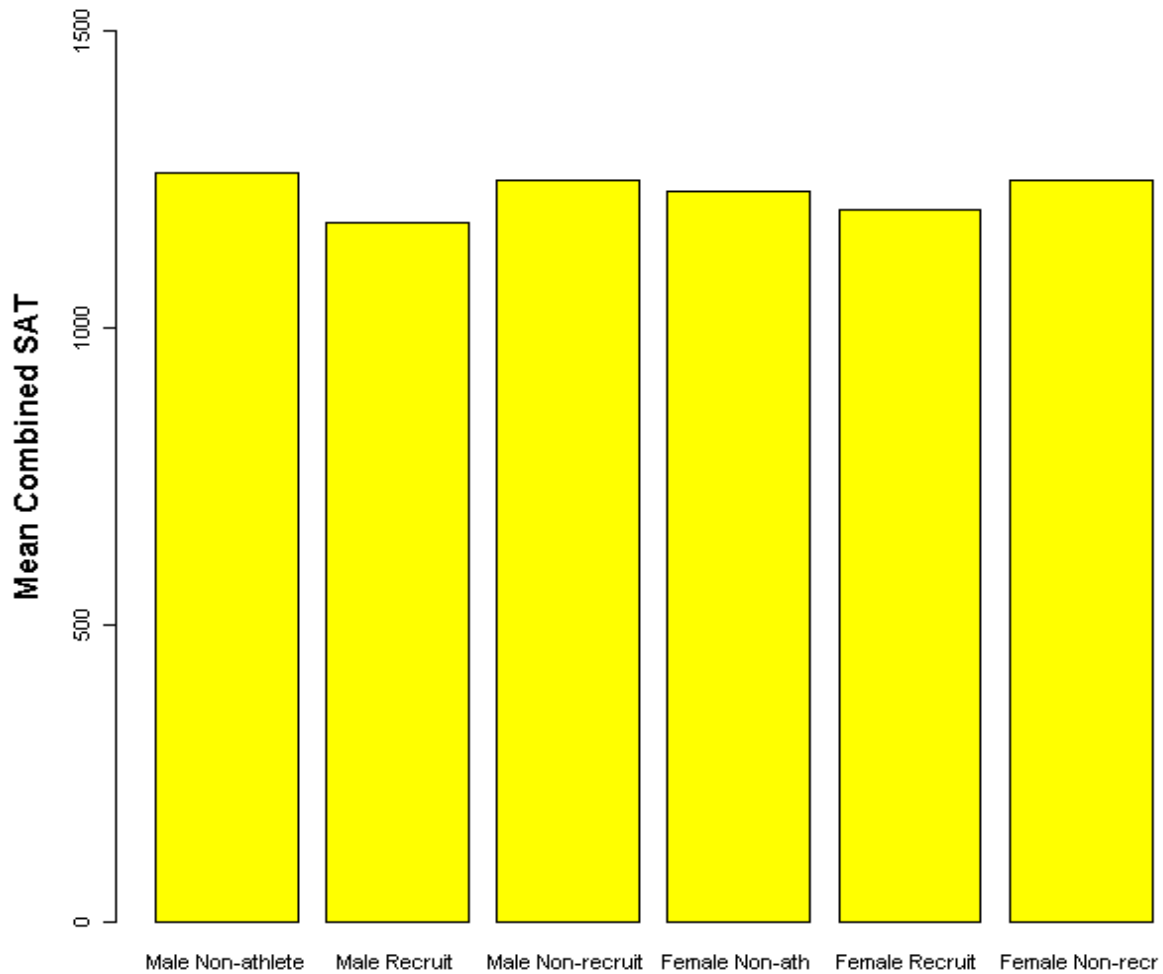
Assessing Academic Underperformance

Can we “explain” the GPA differences using what we know about students when they entered college? One predictor variable is a student’s combined SAT from high school



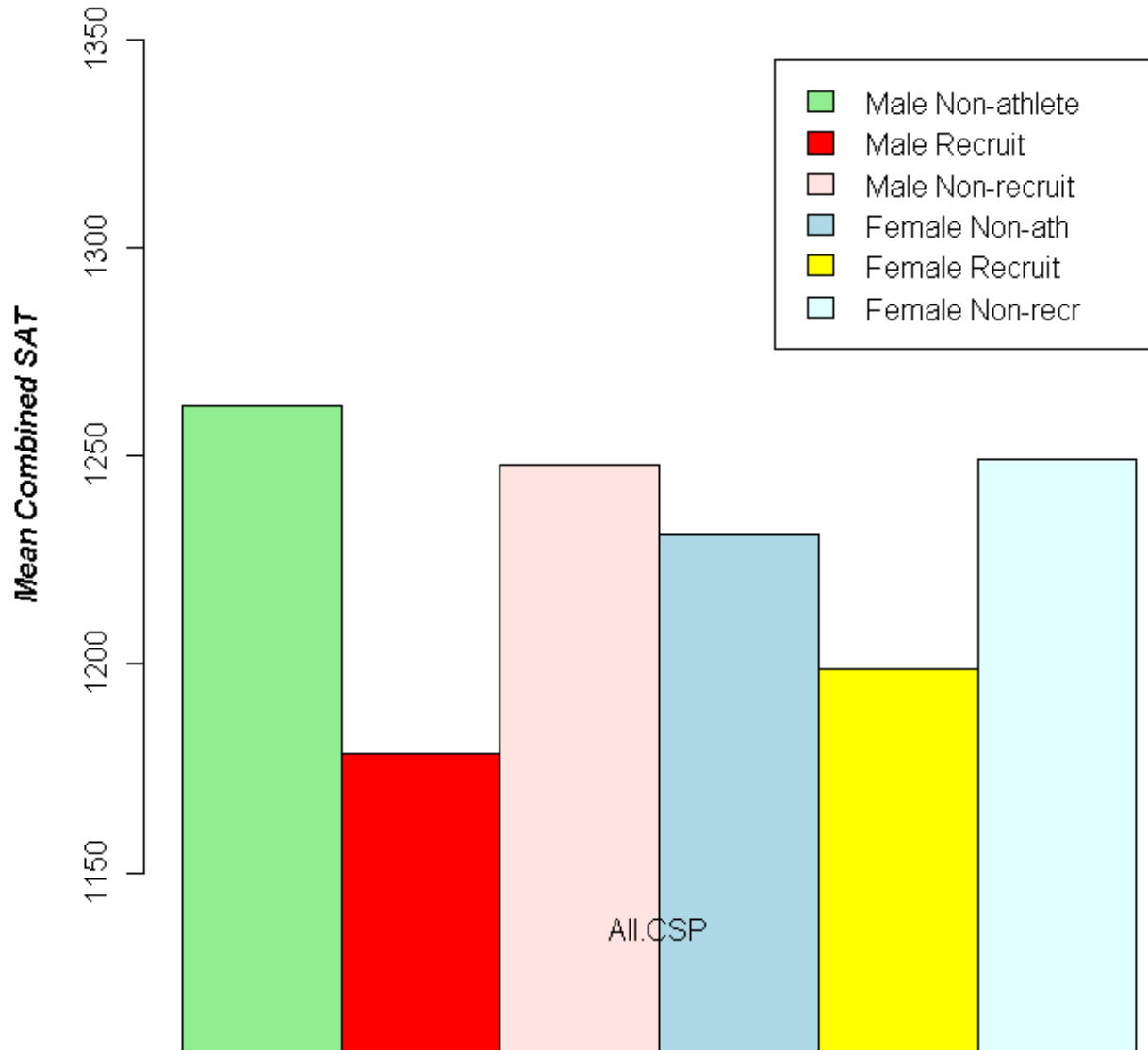
Average Combined SAT

Combined SAT by Gender & Athlete Status, 71 CSP Inst.



Average Combined SAT

Combined SAT by Gender & Athlete Status, 71 CSP Inst.



Multiple Regression Analysis of College GPAs

Use of the model:

- Predict GPAs for groups of athletes when they are recoded as non-athletes. Record differences.
- Athlete's GPA differences in means (observed – predicted) are called “underperformance”

Response variable is 2-Year GPA on 4-point scale



Explanatory variables include:

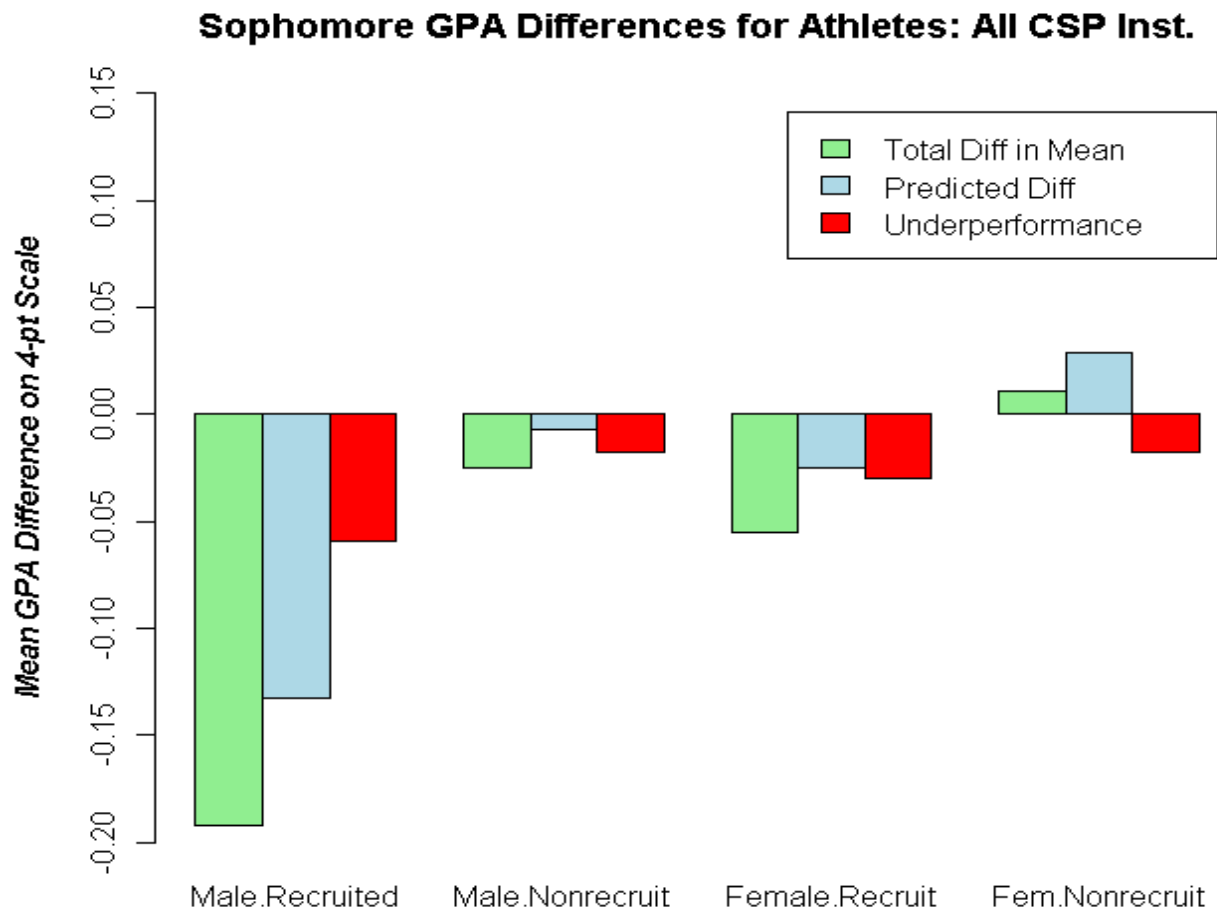
- **Gender (M/F)**
- **Athletic status (Non-ath., Recruited, Non-recruit)**
- **Race/Ethnicity**
- **Citizenship (US/International)**
- **College class standing (Fr., So., Upper)**
- **College Board data on SATs for the high school**
- **College Board data on percent attending college**
- **ID for the 71 colleges– differences among colleges**
- **Many interactions involving above**

A few results:

- **R-squared = 44%**
- **All main effects statistically significant (big data set!)**
- **A few interactions were not statistically significant**



Comparing Athletes to Non-Athletes



Mean Athlete GPA MINUS Mean Nonathlete GPA

Underperformance seems modest. Note that positive differences show groups of athletes with higher college grades than we expect them to have based on their characteristics.



Exploring subgroups of the 71 CSP institutions

- The 71 institutions are heterogeneous: 63 of them are Bachelors Liberal Arts Colleges (Carnegie classification) but these vary considerably by the academic credentials of students and the level of selectivity in admissions.
- We defined subgroups using each college's average combined SAT score as a proxy for selectivity.
- We labeled 8 institutions which are not classified as Bachelor's Liberal Arts Colleges in the Carnegie classification as "Other".

Groups of Institutions Determined by Average Combined SAT

College subgroup	Comb. SAT range	No. colleges	No. students
High	>1250	25	13,694
Middle	1150 to 1250	23	11,089
Lower	<1150	15	5,532
Other (non L.A.)	any SATs	8	6,782

N=71

N=37097



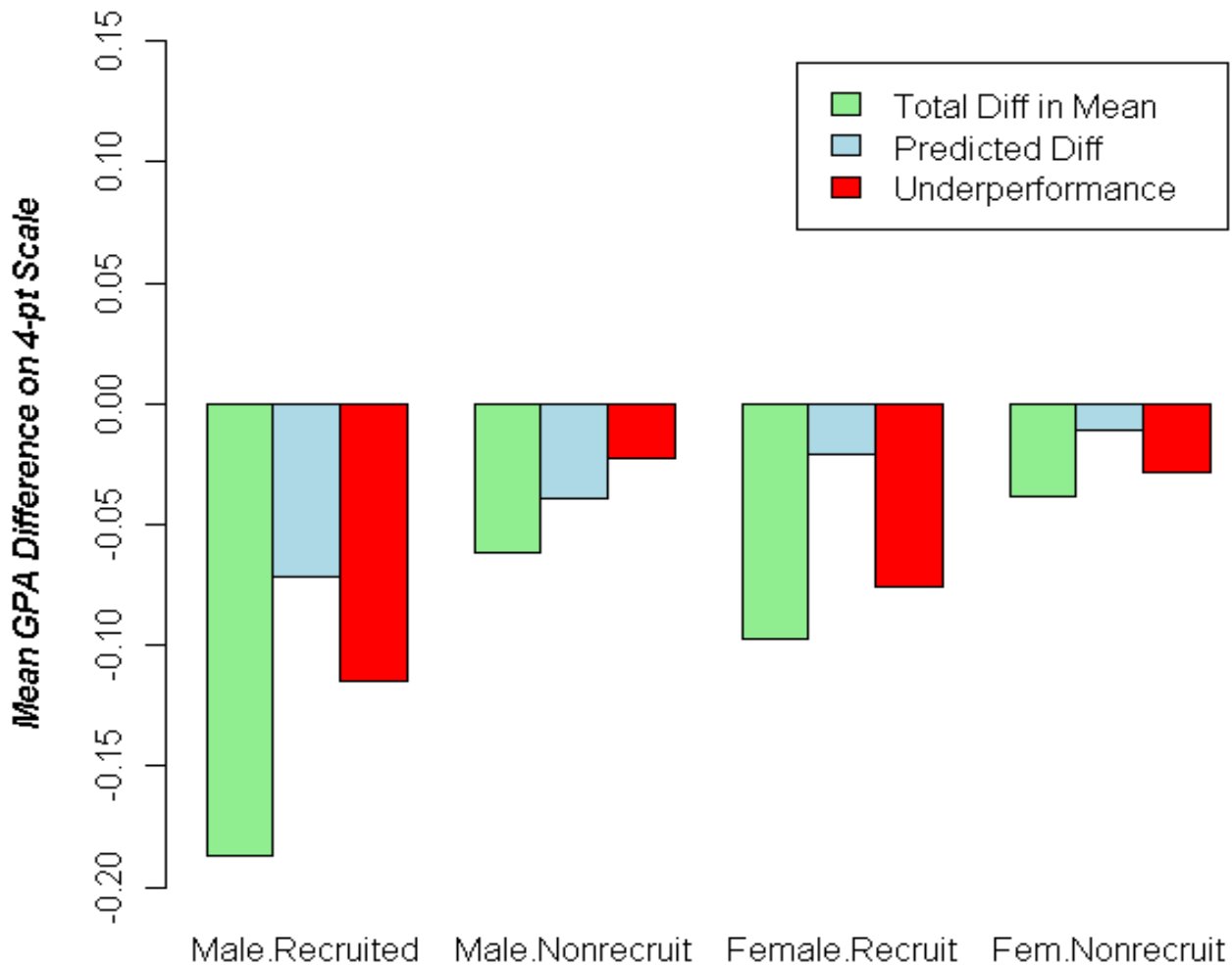
NORTHWESTERN
UNIVERSITY



College Sports Project

Athletes vs. Non-Athletes by Selectivity of Colleges

Sophomore GPA Differences for Athletes: High Selectivity



Mean Athlete GPA MINUS Mean Nonathlete GPA

Note greater “underperformance” among recruited athletes



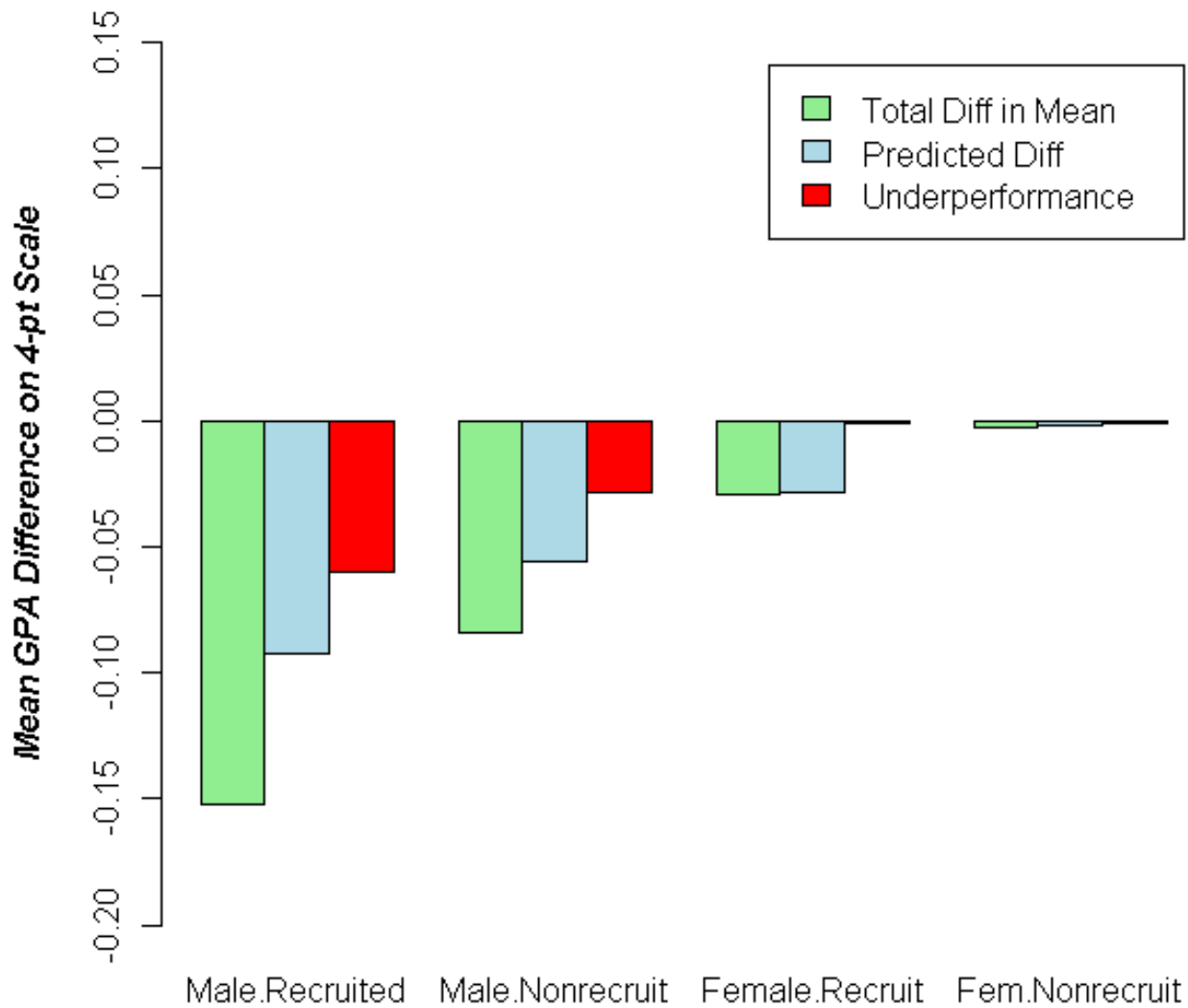
NORTHWESTERN
UNIVERSITY



College Sports Project

Athletes vs. Non-Athletes by Selectivity of Colleges

Sophomore GPA Differences for Athletes: Middle Selectivity



Mean Athlete GPA MINUS Mean Nonathlete GPA



NORTHWESTERN
UNIVERSITY



College Sports Project

Athletes vs. Non-Athletes by Selectivity of Colleges

Sophomore GPA Differences for Athletes: Lower Selectivity



Mean Athlete GPA MINUS Mean Nonathlete GPA

Group with greater underperformance:

- Male recruited athletes at high selectivity institutions
- Female recruited athletes at high selectivity institutions
- Male recruited athletes at medium selectivity institutions

