Grade-Based Performance Prediction

Ben Koester

Outline

- Where is prediction useful at the University?
- Predicting grades at Michigan
  - Example: Physics 140
- Replication outside of Umich
- Building, testing, and improving grade-based models
- Conclusions

Prediction

- Improving advising
  - Meet students' needs/expectations
- Identifying irregularities
  - Outlier terms, groups
- Research in teaching
  - Does a "treatment" affect learning?
- Grades are one metric of performance

Different Predictors

- Grade-based, e.g. GPA
  - Tim McKay: grade penalty
  - Becky Matz: inter-institutional grade penalty (next SLAM seminar)
- Demographics, High School information
- Exams
  - Standardized tests, diagnostic tests, placement exams
- Other validated Instruments
  - Affective characteristics: attitude, self-concept (e.g. CSCI)

Why grades?

- History of grades
  - Schinske and Tanner (*CBE—Life Sciences Education,* 2014)
  - “The result of this investigation [letter grades] is that the experiment started by the faculty five years ago must be pronounced complete failure”(M. Meyer, *Science,* 1908)
- Incentives produced are problematic (Achen and Courant, *J. Economic Perspectives,* 2009)
- Flawed but still informative, meticulously recorded and abundant

Grades at Michigan

Top 10 courses by enrollment in Departments:

Science/Engineering, *Social Sciences*, *Humanities*, and *Other*

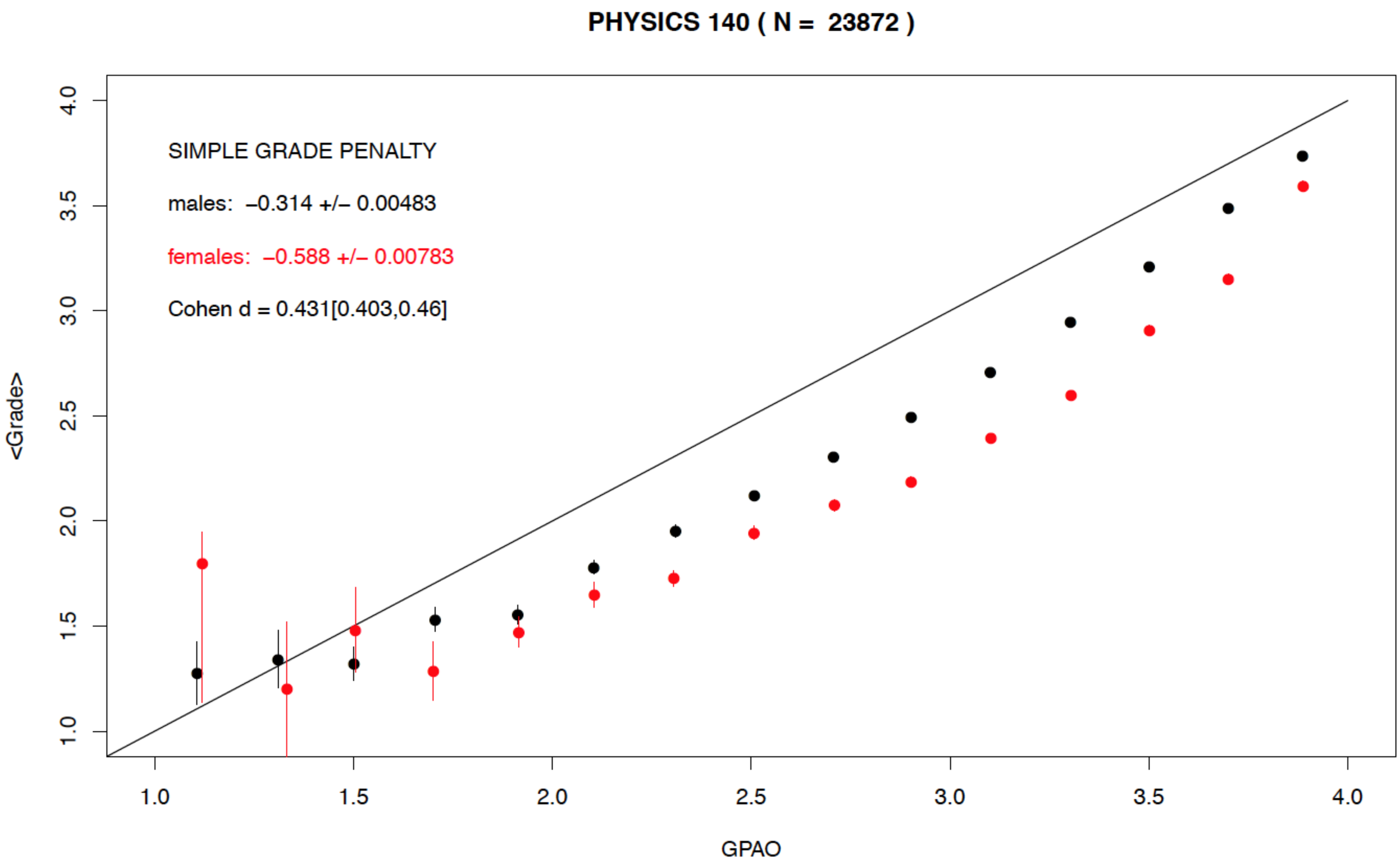
1	2	3	4	5	6	7	8	9	10	Department
104	105	106	107	150	151	254	256	270	280	LS&A First Year Seminars
118	304	310	362	391	392	401	402	406	490	School Of Education
139	139	140	149	344	345	346	347	348	349	School of Music, Theatre and Dance
211	221	231	321	331	418	419	450	458	499	Biomedical Engineering
230	240	350	351	354	363	368	453	459	468	Office of International Programs
100	101	101	102	102	122	201	202	281	331	Near Eastern Studies Department
101	102	221	231	232	243	322	325	326	386	Germanic Languages & Lit Dept
100	201	204	205	206	209	240	301	374	399	American Culture Program
102	111	200	209	210	211	272	315	370	375	Department of Linguistics
220	240	253	270	295	300	324	375	400	483	Women's Studies Department
101	102	125	126	201	202	220	225	226	230	Asian Languages And Cultures
100	110	120	121	130	150	151	220	231	300	School Of Art And Design
111	120	230	240	250	270	280	303	370	401	Psychology Department
122	210	245	252	254	354	356	358	454	456	School Of Nursing
101	101	102	191	192	222	231	232	372	385	Classical Studies Department
102	105	110	111	139	201	211	232	302	360	Program in the Environment
124	125	223	225	239	240	313	317	325	367	English Language & Literature Dept
100	101	103	110	151	195	280	390	455	490	Engineering Undergraduate Educ
101	111	140	160	300	314	353	389	489	496	Political Science Department
122	201	202	230	280	296	310	312	381	481	Studies In Religion
200	236	236	272	290	350	360	366	366	370	Screen Arts and Cultures
101	161	272	285	298	330	344	364	365	368	Anthropology Department
103	111	111	340	358	450	451	458	490	495	Department of Afro−American and Af
101	102	112	212	222	250	251	271	272	394	History Of Art Department

1	2	3	4	5	6	7	8	9	10	Department
300	306	310	400	418	422	427	428	429	436	Molecular, Cellular, and Dev
211	212	260	303	325	351	360	402	421	431	Civil & Environmental Engr
101	102	103	104	106	111	112	115	127	142	Astronomy Department
409	410	411	412	431	432	434	462	485	486	College Of Pharmacy
100	210	239	256	301	306	337	375	418	438	Sch Of Nat Resources & En
101	102	111	211	351	361	371	381	439	458	Communication Studies
100	101	102	210	303	305	310	344	345	368	Sociology Department
215	225	245	285	305	315	325	335	345	405	Aerospace Engineering
271	272	300	300	300	300	301	312	350	471	School of Business Adminis
180	181	196	201	202	232	303	355	359	361	Philosophy Department
220	242	250	330	350	360	412	420	480	489	Materials Science & Enginee
230	330	341	342	343	344	360	460	466	487	Chemical Engineering Depa
312	313	314	315	316	317	322	323	326	425	College of Architecture & Ur
125	126	127	128	135	136	140	141	240	241	Physics Department
211	235	240	250	350	360	382	395	450	495	Mech Eng & Applied Mech D
125	126	130	210	211	215	216	230	241	260	Chemistry Department
201	202	265	310	316	333	334	366	373	425	Industrial−Operations Engr
100	250	350	401	402	408	412	425	426	470	Statistics Department
183	203	215	270	280	281	314	370	482	496	Electrical Engr & Computer
101	102	103	103	231	231	232	232	275	276	Romance Languages Depar
118	162	171	172	173	207	225	226	305	310	Biology Department
101	102	310	340	395	398	401	402	404	435	Economics Department
100	181	183	198	270	280	303	370	380	482	Program In Computer Scien
100	102	103	105	106	107	110	111	113	222	Earth and Environmental Sc
103	105	115	116	156	215	216	217	417	425	Mathematics Department

101	102	221	231	232	243	322	325	326	386	Germanic Languages & Lit Dept
183	203	215	270	280	281	314	370	482	496	Electrical Engr & Computer Sci

Light yellow: 3.85 mean course grade → dark red: 2.65 mean course grade

Predicting Grades: Physics 140

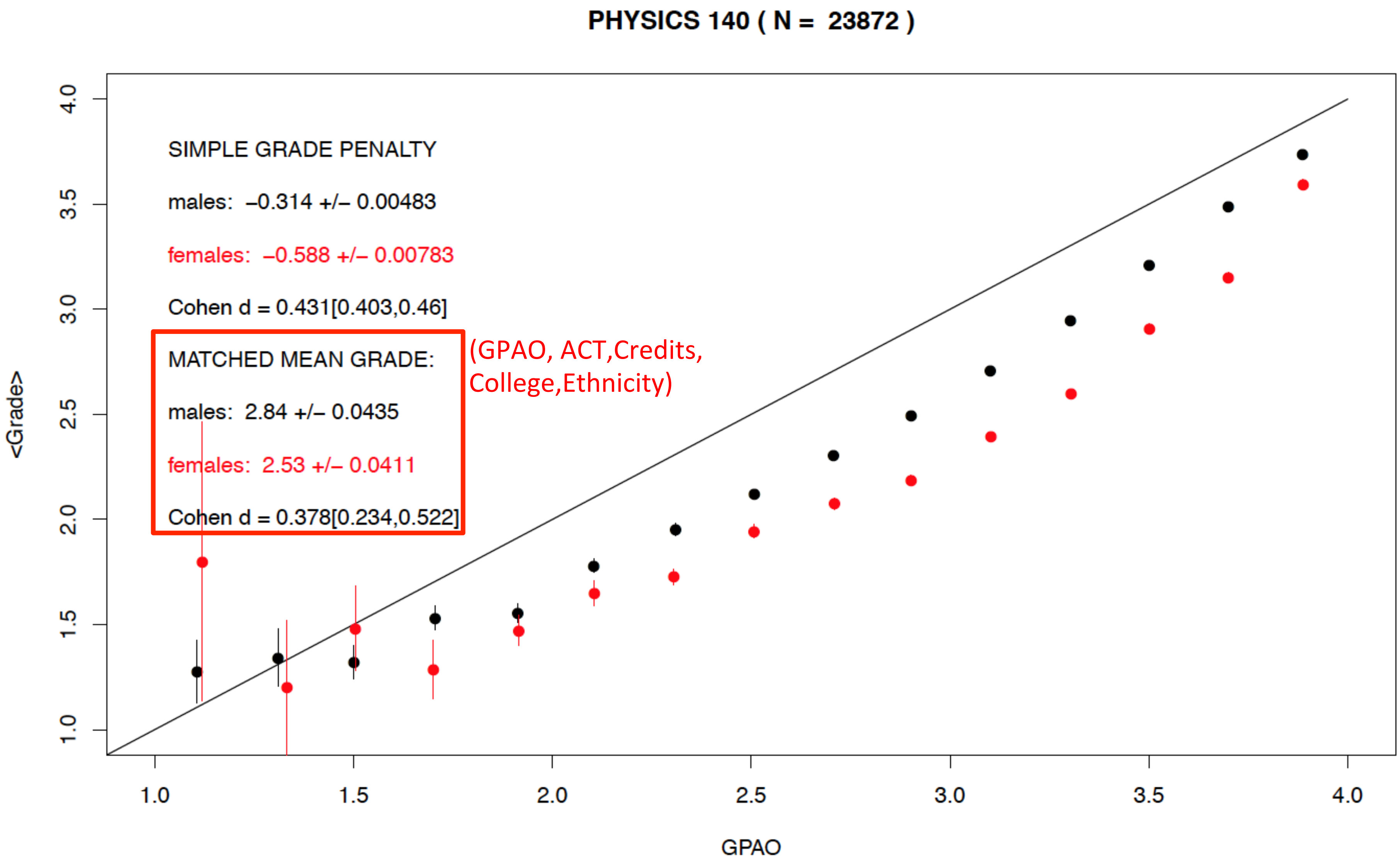


Predicting Grades: Physics 140

	Estimate	Std. Error	t value	Pr(>ltl)
(Intercept)	−1.54	0.095	−16.2	3.7e−58
GPAO	1.09	0.0138	79	0
SEXM	0.228	0.0161	14.1	9.96e−45
INC	−0.00139	0.00178	−0.779	0.436
TOT_TAKEN_GPA	−0.00279	0.00052	−5.36	8.48e−08
LAST_ACT_COMP_SCORE	0.0226	0.0027	8.39	5.6e−17
ETHNIC_GROUP_DESCRSHORTAsian	0.173	0.0413	4.19	2.81e−05
ETHNIC_GROUP_DESCRSHORTBlack	−0.0907	0.0578	−1.57	0.117
ETHNIC_GROUP_DESCRSHORTHawaiian	−0.592	0.445	−1.33	0.184
ETHNIC_GROUP_DESCRSHORTHispanic	0.0414	0.0509	0.814	0.416
ETHNIC_GROUP_DESCRSHORTNative Amr	−0.187	0.185	−1.01	0.312
ETHNIC_GROUP_DESCRSHORTNot Indic	0.14	0.0474	2.95	0.00316
ETHNIC_GROUP_DESCRSHORTWhite	0.0442	0.0388	1.14	0.254

The story is the same in other courses: grades are the best predictors of grades in future classes! But they aren't the whole story…

Predicting Grades: Physics 140



Do these predictors replicate elsewhere?

- Compare to CIC institutions
  - Similar to Umich
- Idea: Standardize student data, keep it *local*, do the same analyses (i.e., run the same code)
  - Build standard tables like this:

Student Record (SR) Table

This contains student demographic and background information, as well as other single-record information such as date and type of degree.

ID	GENDER	ETHNICITY	FIRST_TERM	DEGREE_TERM	TRANSFER	MAJOR1	MAJOR2	MAJOR1_LONG	MAJOR2_LONG	FIRST_DECLARE	FIRST_DECLARE_LONG	FIRST_DECLARE_TERM	PELL_STATUS
XXXXX	M/F/U	IPEDS (1-9)	YYYYMM	YYYYMM	Y/N	CIP code (XX.XXXX)	CIP code (XX.XXXX)	CIP title	CIP title	CIP code	CIP title	YYYYMM	Y/N

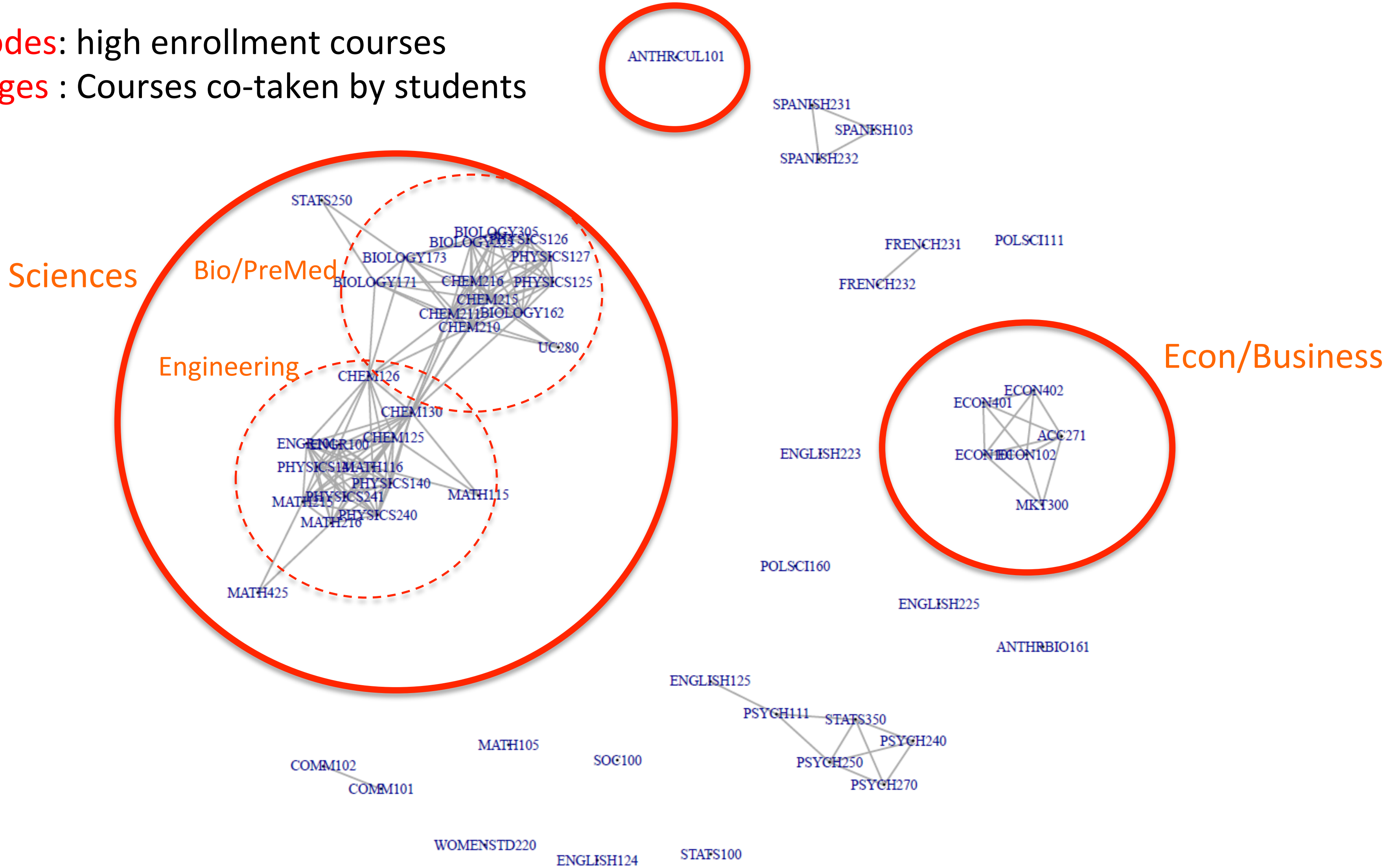
- Difficulties: Heterogeneity among institutions, expectations of the participants
- Ongoing work (see Becky Matz)

How far can we go with grades to predict?

- GPA
  - Baseline grade predictor
  - Credit hour weighted-mean of grades
  - But: agnostic to subject, student performance, term, etc.
- GPAR (grade points above replacement, e.g. Caulkins et al., 1996)
  - For a student in a course, compute mean course grade, subtract that student's grade
  - Created credit hour weighted average of this
  - But maybe you're in a class full of awesome students!
- Student "fixed-effect" (SFE, Murdock et al., 2015)
  - In each course, use the global performances of your peers to improve accuracy

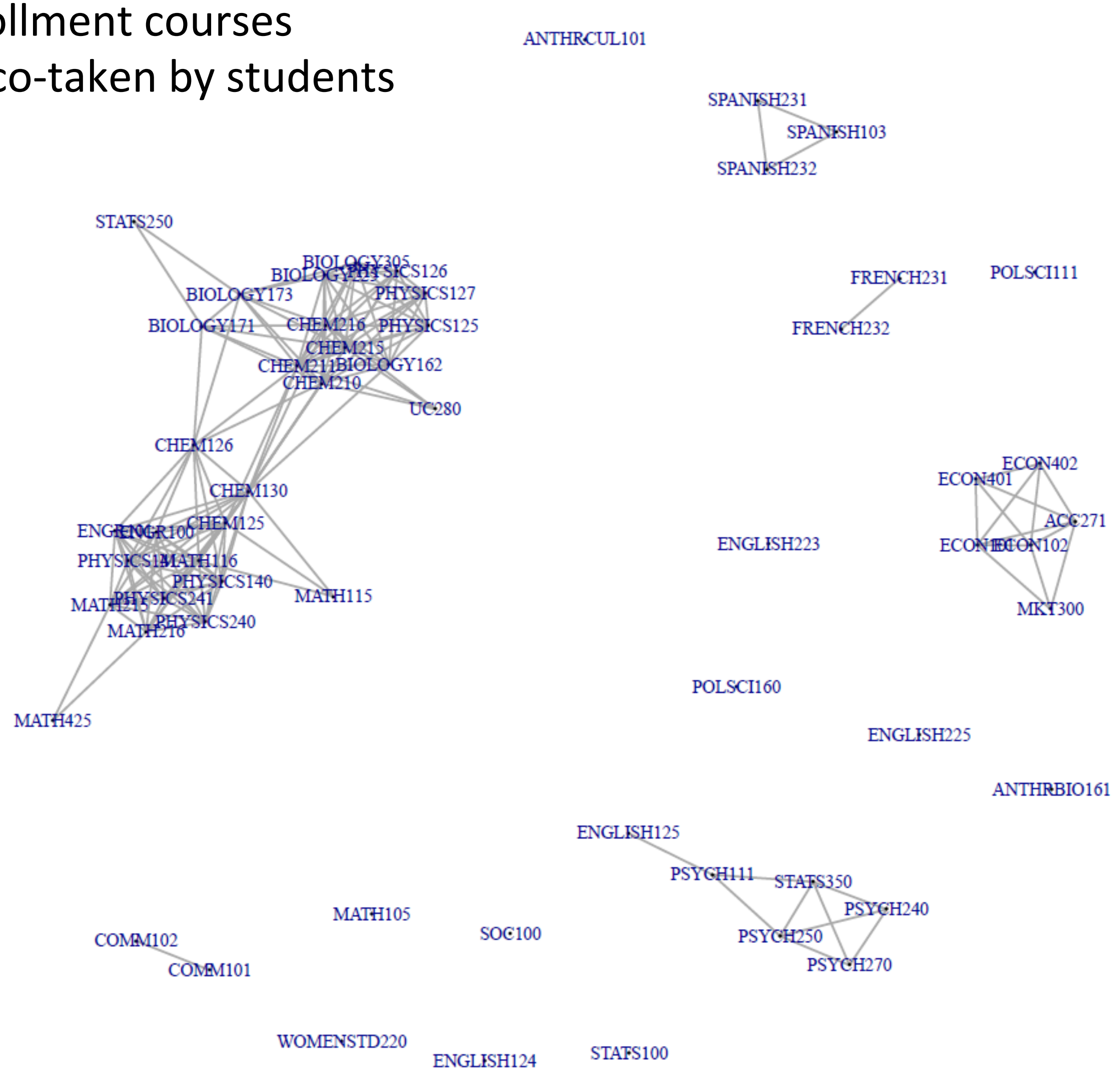
Course-Taking at Michigan

Nodes: high enrollment courses
Edges : Courses co-taken by students



Course-Taking at Michigan

Nodes: high enrollment courses
Edges : Courses co-taken by students



Student Fixed-Effect

- Each of grade at U of M is a linear combination of a course/term-invariant student effect, and a student-invariant course effect

$$Grade_{ict} = StudentFE_i + ClassFE_{ct} + \epsilon_{ict}$$

  – In practice this creates a matrix of ~6,000,000 grades x 150,000 student/course effects, LS solution not tractable, but matrix is sparse
  – Solve for the fixed effects using Arcidiacono (2012)

Which of GPA, GPAR, and SFE predicts grades better?

- Compare grades for ALL undergrad courses in Winter 2014 to our predictions
- Best predictor has the highest correlation coefficient

Correlation coefficients between predicted and actual Winter 2014 grades

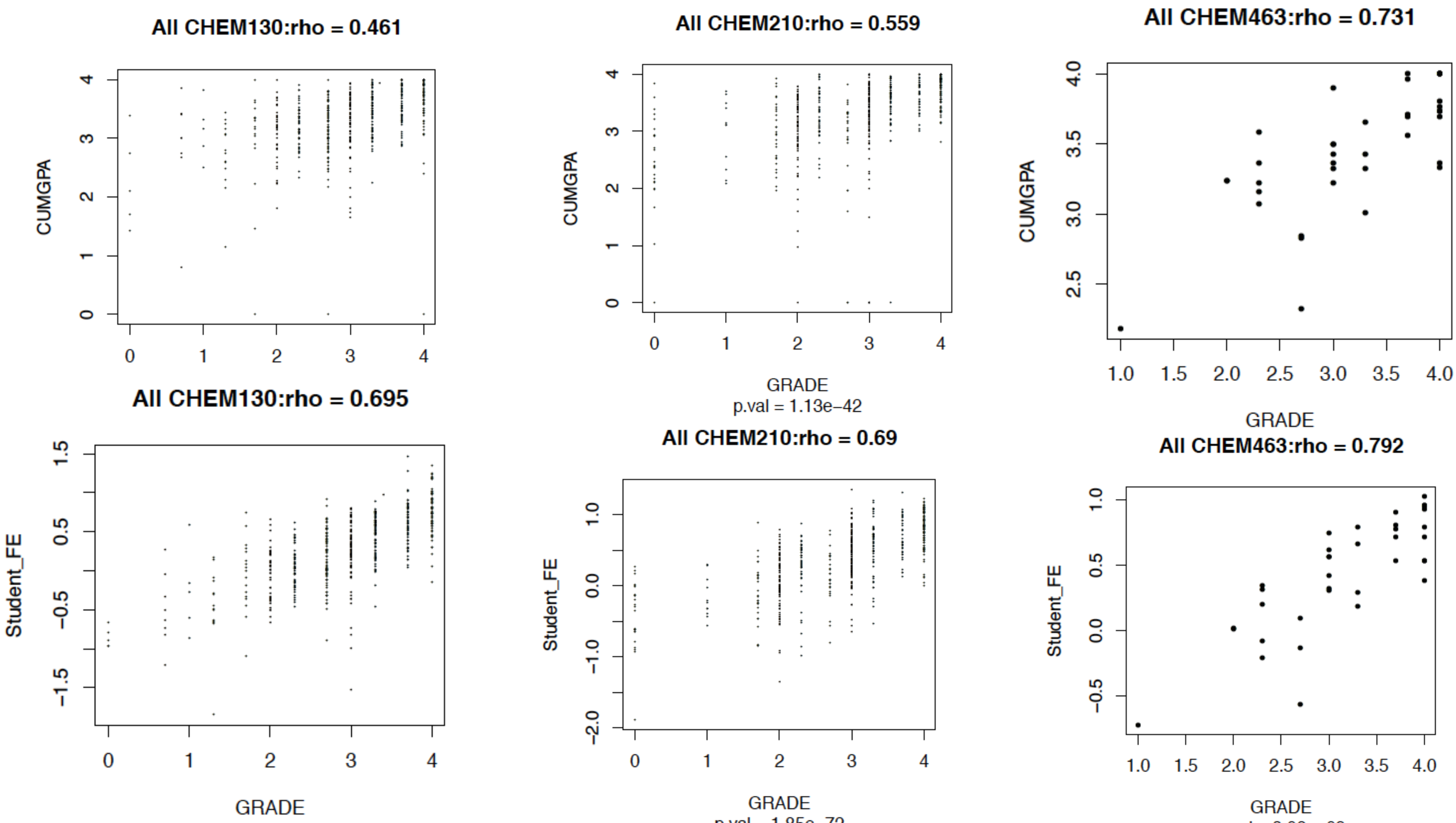
	(1)
	$\rho_{pearson}$
SFE	0.734^{***}
GPA	0.403^{***}
GPAR	0.423^{***}
GPAR2	0.465^{***}
N	~ 150,000

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Which of GPA, GPAR, and SFE predicts better?

- Chemistry, Winter 2014



Which of GPA, GPAR, and SFE predicts ranks better?

- Still depends on course size, subject, level

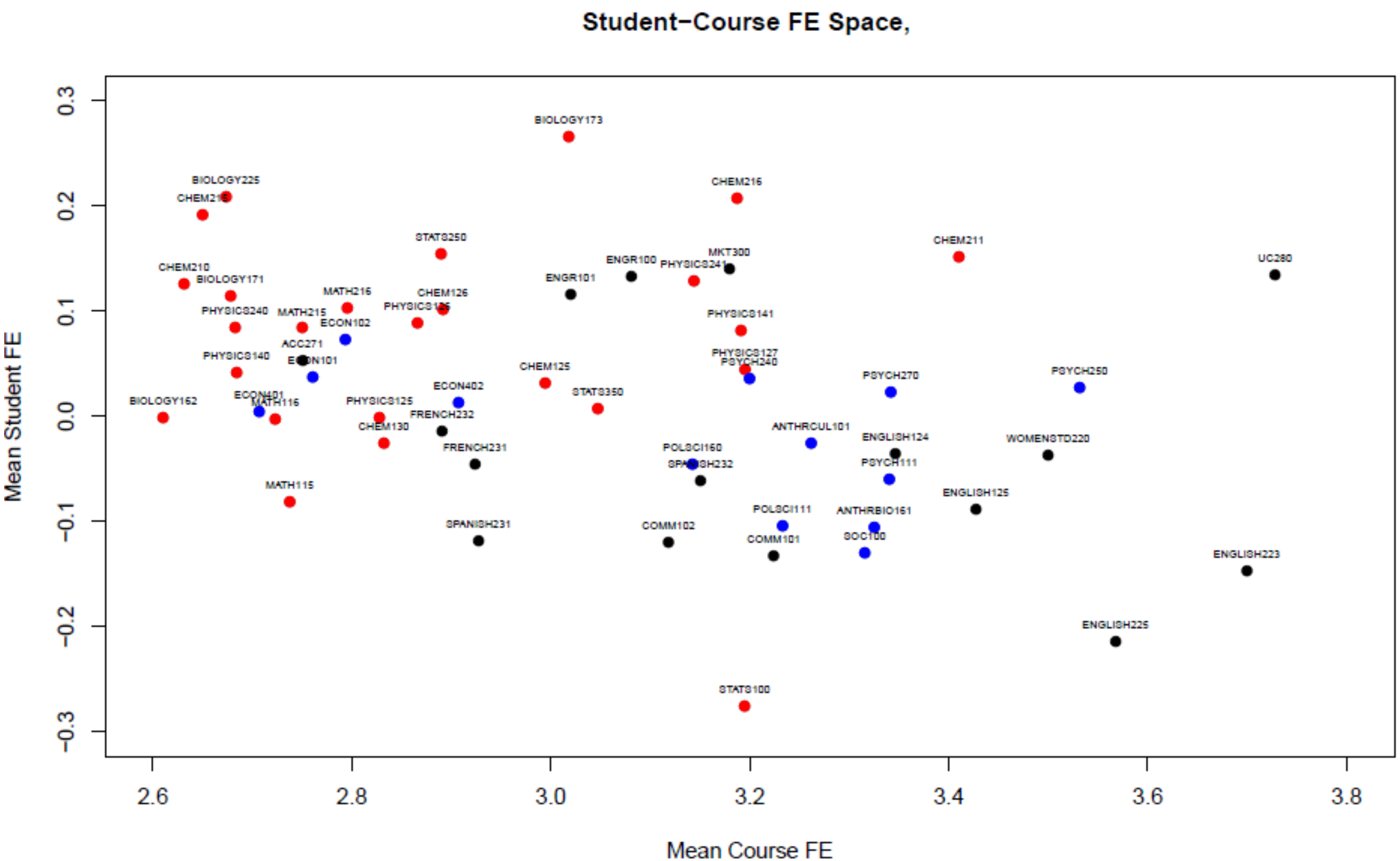
SUBJECT	CATALOG	N	GPA	GPAAR1	SFE
MATH	116	705	0.679***	0.708***	0.71***
MATH	433	16	0.649*	0.652*	0.738*
PHYSICS	140	618	0.632***	0.62***	0.649***
PHYSICS	340	27	0.787***	0.701***	0.687***
PHYSICS	453	23	0.917***	0.904***	0.885***
BIOLOGY	171	535	0.636***	0.68***	0.714***
BIOLOGY	305	328	0.738***	0.749***	0.762***
MCDB	427	51	0.782***	0.803***	0.784***
CHEM	130	480	0.554***	0.633***	0.697***
CHEM	210	494	0.653***	0.677***	0.692***
CHEM	463	38	0.785***	0.805***	0.792***
PSYCH	111	861	0.569***	0.629***	0.658***
PSYCH	458	177	0.532***	0.58***	0.584***
ENGLISH	125	954	0.396***	0.426***	0.419***

- None of these methods work well for courses with narrow grade distributions (i.e. everyone gets an A)

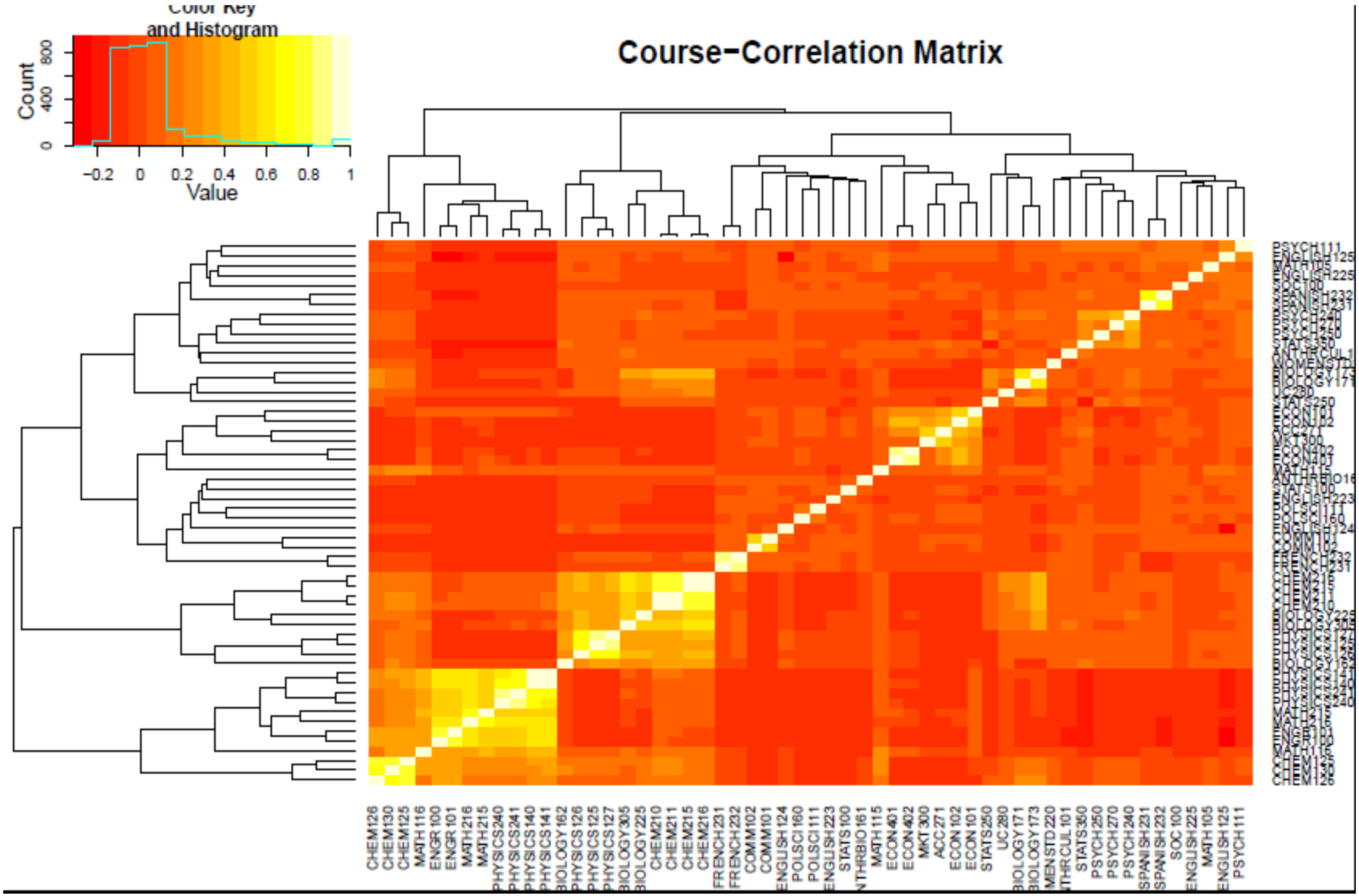
Conclusions/Next Steps

- Grades predict other grades well
- Student/course relationships allow us to improve predictive power
  - On a final transcript, final GPA and final SFE will tell different stories
- Advising: which metrics for which students in which context?
- Continue to reconsider the meaning of grades and the GPA students carry with them upon graduation

A Sidebar of the Modeling



Course-Taking at Michigan



Example: STEM interest at UMich