Equilibrium Grade Inflation with Implications for Female Interest in STEM Majors

Thomas Ahn  
University of Kentucky

Peter Arcidiacono  
Duke University & NBER

Amy Hopson  
Duke University

James Thomas  
Duke University
Motivation

- Perceived shortage of STEM majors despite these fields paying well.
- *The Obama Administration stands committed to providing students at every level with the skills they need to excel in the high-paid, highly-rewarding fields of science, technology, engineering, and math (STEM).* (White House website)
- 100K in 10 program wants 100,000 more STEM teachers (primary and secondary) within the next ten years.
- The lack of STEM majors is not driven by a lack of interest:
  - 48% of those pursuing a bachelor’s degree in a STEM field leave, with half going to a non-STEM field and half dropping out (NCES 2013)
MAJOR CHOICES profiles more than 65 recent Princeton graduates who followed their intellectual passions to major in subjects they loved, subjects that in most cases bore very little obvious connection to the careers they have subsequently pursued. Its purpose is to encourage undergraduates to follow their intellectual passions and study what they love, with confidence in the fulfilling lives that lie ahead and the knowledge that in no way will their choice of major limit the career choices they may wish to make in the future.

The book focuses on many of the smaller departments in order to encourage undergraduates to be imaginative and open-minded about their choices and to take the fullest advantage of the many intellectual opportunities available to them at Princeton.

(emphasis added)
Universities may not want to see more students in STEM fields if it comes at the expense of other departments.

Princeton’s push for enrollments in smaller departments has been somewhat successful, with the largest percentage shifts in classics, music, Slavic languages and literature, comparative literature, and religion. (Princeton website)

But universities (or their professors....) may engage in other activities that encourage more even representation across majors:
- differences in grading standards
- differences in workloads
Grading standards and workloads differ substantially across departments

- Unadjusted difference in grades is around 0.3.
- Students in STEM classes study an extra 45 minutes per week, off a base of 2.3 hours a week.

These differences occur despite STEM classes drawing students with higher test scores.

Suggests universities are actually subsidizing students to go into low paying majors...
Why are STEM classes different?

We want to separate out how much of the differences across departments are driven by:

- Student demand for courses
- Professor preferences
Why does it matter?

- Harsher grading in STEM classes affects enrollment; many come in intending to major in STEM but then attrit.
- Who switches is predictable: those with the least preparation and worst performance within a school are much more likely to switch.
  - Controlling for academic background virtually eliminates racial differences in STEM persistence (at least at Duke...).
- How courses are graded affects not only the number of STEM majors but also their composition.
Women and STEM

- But women are both less likely to begin in STEM and more likely to switch out of STEM. This is surprising because:
  - Women are just as academically prepared as men.
  - Women study significantly more than men (about 2 hours a week at both Duke and Berea).

- Higher study times by women may reflect both:
  - Lower costs of studying
  - Valuing grades more

where the first makes STEM more attractive, the second makes STEM less attractive.
Suppose women both have lower costs of studying and value grades more than men.

A policy that restricts average grades to be the same across classes would result in:

- Differences in average grades between STEM and non-STEM falling
- STEM classes requiring even more work relative to non-STEM classes

With average grades and workload going up in STEM, STEM becomes relatively more attractive to women.
Our approach

Using transcript and course evaluation data from the University of Kentucky

- Estimate grade production functions
- Estimate student choices of courses and effort conditional on the grade production functions
- Estimate professor preference parameters given:
  - Estimates of the grade production parameters for all courses
  - Estimates of student preferences
Administrative data from the University of Kentucky - Lexington (UK)

Full data: Fall, 2008 - Spring, 2013
Our Sample: Fall, 2012
- 20,343 unique students
- 100,811 student-course observations

Also have data on course evaluations which we use to get average study time
## Descriptive Statistics by Gender

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school GPA</td>
<td>3.13</td>
<td>3.34</td>
</tr>
<tr>
<td></td>
<td>(1.20)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>ACT Score</td>
<td>25.2</td>
<td>24.4</td>
</tr>
<tr>
<td></td>
<td>(4.42)</td>
<td>(4.18)</td>
</tr>
<tr>
<td>Fall 2012 GPA</td>
<td>3.02</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>(0.713)</td>
<td>(0.665)</td>
</tr>
<tr>
<td>Fall 2012 Credits</td>
<td>11.7</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>(4.29)</td>
<td>(4.22)</td>
</tr>
<tr>
<td>STEM Major</td>
<td>38.0%</td>
<td>23.8%</td>
</tr>
</tbody>
</table>

Note: Fall 2012 University of Kentucky undergraduate students, 9,729 men, 9,798 women. Standard deviations in parentheses.
## Descriptive Statistics by Course Type

<table>
<thead>
<tr>
<th></th>
<th>STEM</th>
<th>Non-STEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Size</td>
<td>78.1</td>
<td>46.3</td>
</tr>
<tr>
<td></td>
<td>(101.1)</td>
<td>(64.0)</td>
</tr>
<tr>
<td>Average Grade</td>
<td>3.03</td>
<td>3.31</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Average Grade</td>
<td>Female</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Study Hours</td>
<td>3.61</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(1.12)</td>
</tr>
<tr>
<td>Percent Female</td>
<td>37.0%</td>
<td>55.9%</td>
</tr>
</tbody>
</table>

Note: Fall 2012 University of Kentucky courses with enrollments of 16 or more students, 379 STEM courses, 1,164 non-STEM courses. For study hours, 293 STEM courses and 793 non-STEM courses. Standard deviations in parentheses.
Regressions of Grades and Study Time on Characteristics of the Individual and/or Class

<table>
<thead>
<tr>
<th>Dependent Var.</th>
<th>Grade</th>
<th>Study hours per week</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEM Class</td>
<td>-0.325</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.148)</td>
</tr>
<tr>
<td>Female</td>
<td>0.140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Percent Female</td>
<td>0.395</td>
<td>0.547</td>
</tr>
<tr>
<td></td>
<td>(0.203)</td>
<td>(0.191)</td>
</tr>
<tr>
<td>Average Grade</td>
<td>-0.635</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td></td>
</tr>
<tr>
<td>ln(Class Size)</td>
<td>-0.116</td>
<td>-0.396</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Observations</td>
<td>72,449</td>
<td>1,085</td>
</tr>
</tbody>
</table>

Note: Additional controls for grades regression include, minority status, freshman, STEM major, pell grant, in-state student, ACT score, HS gpa, percent minority. Additional controls in study hours regression include percent freshmen, percent STEM major, percent pell grant, percent in-state, average ACT score, average HS gpa, percent minority.
Grade Production

- \( j = [1, \ldots, J] \) indexes courses, each course \( j \) belongs to an area of study \( k \).
- \( A_{ij} \) is \( i \)'s preparation for course \( j \).
- Grades for \( i \) in course \( j \), \( g_{ij} \), are given by:

\[
g_{ij} = \beta_j + \gamma_j \left( A_{ij} + \ln(s_{ij}) \right) + \eta_{ij}
\]

where \( s_{ij} \) refers to study time and \( \eta_{ij} \) is noise.
- \( \beta_j \) and \( \gamma_j \) are set by the professor (restricted to linear grading policies)
- \( s_{ij} \) is set by the student.
Utility for choosing course $j$ is given by:

$$U_{ij} = \phi_i \mathbb{E} [g_{ij}] - \psi_i s_{ij} + \delta_{ij}$$

Students then solve the following maximization problem when choosing their optimal course bundle:

$$\max_{d_{i1}, \ldots, d_{iJ}} \sum_{j=1}^{J} d_{ij} U_{ij}$$

subject to: $\sum_{j=1}^{J} d_{ij} = n, \ d_{ij} \in \{0, 1\} \forall j$
The optimal study effort in course \( j \) can be found by differentiating \( U_{ij} \) with respect to \( s_{ij} \):

\[
0 = \frac{\phi_i \gamma_j}{s_{ij}} - \psi_i
\]

\[
S^*_{ij} = \frac{\phi_i \gamma_j}{\psi_i}
\] (2)

Substituting the optimal choice of study time into the utility function yields:

\[
U_{ij} = \phi_i \left( \beta_j + \gamma_j \left[ A_{ij} + \ln(\phi_i) + \ln(\gamma_j) - \ln(\psi_i) - 1 \right] \right) + \delta_{ij}
\] (3)
Reduced form grade equation

Substituting the expression for optimal study time into the grade process equation yields:

\[ g_{ij} = \beta_j + \gamma_j (A_{ij} + \ln(\phi_i) + \ln(\gamma_j) - \ln(\psi_i)) + \eta_{ij} \]  \hspace{1cm} (4)

Professors who set relatively higher values of \( \gamma_j \) see more study effort because higher \( \gamma_j \)'s induce more effort and because higher \( \gamma_j \)'s attract students with lower study costs.
The key equations for estimation are given by:

(i) the solution to the students maximization problem after substituting in optimal study effort,

(ii) the reduced form grade production process, and

(iii) the optimal study effort.
Parameterizations

\[ A_{ij} = w_i \alpha_{1k(j)} + X_i \alpha_{2k(j)} \]

Preparation in field \( k(j) \)

\[ \delta_{ij} = \delta_0 + w_i \delta_{1k(j)} + Z_{ij} \delta_{2k(j)} + \epsilon_{ij} \]

Preference for course \( j \)

\[ \psi_i = \exp (\psi_0 + w_i \psi_1 + X_i \psi_2) \]

Study costs

\[ \phi_i = \phi_0 + w_i \phi_1 \]

Preferences for grades

where:

- \( w_i \) is an indicator for female
- \( X_i \) includes ACT scores, other family background characteristics
- \( Z_{ij} \) includes year of the student interacted with level of the course as well as female cross female instructor
- \( \epsilon_{ij} \) i.i.d. Type 1 extreme value preference shock
Grade parameters

Estimate the following by NLLS:

\[ g_{ij} = \theta_{0j} + \gamma_j \left( w_i \theta_{1k(j)} + X_i \theta_{2k(j)} \right) + \eta_{ij} \] (5)

where:

\[ \theta_{0j} = \beta_j + \gamma_j (\ln(\phi_0) + \ln(\gamma_j) - \psi_0) \]
\[ \theta_{1k(j)} = \alpha_{1k(j)} + \ln(\phi_0 + \phi_1) - \ln(\phi_0) - \psi_1 \]
\[ \theta_{2k(j)} = \alpha_{2k(j)} - \psi_2 \]

- The variation in the data used to identify \( \{\theta_{1k(j)}, \theta_{2k(j)}\} \)
  comes from the relationship between student characteristics and grades.
- The \( \gamma_j \)’s are identified from how these characteristics translate into grades relative to the normalized course.
We assume that the relationship between log study hours and log study effort is linear; and that log study hours is reported with measurement error $\zeta_{ij}$:

$$
\ln(h_{ij}) = \mu \ln(s^*_{ij}) + \zeta_{ij}
$$

$$
= \mu \left( \ln(\phi_i) + \ln(\gamma_j) - \ln(\psi_i) \right) + \zeta_{ij}
$$

$$
= \kappa_0 + w_i \kappa_1 + X_i \kappa_2 + \ln(\gamma_j) + \zeta_{ij}
$$

where:

$$
\kappa_0 = \mu(\psi_0 + \ln(\phi_0))
$$

$$
\kappa_1 = \mu(\ln(\phi_0 + \phi_1) - \ln(\phi_0) - \psi_1)
$$

$$
\kappa_2 = -\mu \psi_2
$$
Needed to normalize one $\gamma$ to one for each department in the grade estimation. We can now undo this normalization. Namely, we estimated $\gamma_j^N = \gamma_j / C_{k(j)}$ where $C_{k(j)}$ is a group-specific constant.

Substituting in for $\ln(\gamma)$ with $\ln(\gamma^N) + \ln(C_{k(j)})$ yields our estimating equation:

$$\ln(h_{ij}) = \tilde{\kappa}_0 + w_i\kappa_1 + X_i\kappa_2 + \kappa_{3k(j)} + \mu\ln(\hat{\gamma}^N) + \zeta_{ij}$$ (6)

where $\kappa_{3k(j)} = \mu\ln(C_{k(j)}/C_1)$ and $\tilde{\kappa}_0 = \kappa_0 + \ln(C_1)$.

We can then partially undo the normalization on the $\gamma$’s, with $\hat{\gamma}_j^P = \hat{\gamma}_j^N \exp(\hat{\kappa}_{2k(j)}/\hat{\mu})$ now normalized with respect to one course.

We don’t actually observe linked course evaluation records but we do know year in school so our estimating equation is formed from averaging across individuals by year of school.
Utility parameters

\[
E[ g_{ij} | s^*_{ij} ] = \hat{\theta}_0 + \hat{\gamma}_j^N \left( w_i \hat{\theta}_{1k(j)}^N + X_i \hat{\theta}_{2k(j)}^N \right)
\]

\[
U_{ij} = \delta_{0j} + w_i \delta_{1k(j)} + Z_{ij} \delta_{2k(j)} + E[ g_{ij} | s^*_{ij} ](\phi_0 + w_i \phi_1) + C_1 \hat{\gamma}_j^P (\phi_0 + w_i \phi_1) + \epsilon_{ij}
\]

- \( \phi_0 \) and \( \phi_1 \) are identified from sorting into classes that reward a student's abilities more: lower ability students have a relative preference for courses with low \( \gamma \)'s.
- With these estimates we can recover all remaining structural parameters.
Even with Type 1 extreme value errors estimation is complicated because students are choosing *bundles* of courses.

Denote $K_i$ as the set of courses chosen by $i$.

Denote $M_i$ as the highest payoff associated with any of the non-chosen courses:

$$M_i = \max_{j \notin K_i} \delta_{0j} + w_i \delta_{1k(j)} + Z_{ij} \delta_{2k(j)} + \left( \hat{E}[g_{ij}] - \hat{\gamma}_j \right) \left( \phi_0 + w_i \phi_1 \right) + \epsilon_{ij}$$

Suppose $K_i$ consisted of courses $\{1, 2, 3\}$ and that the values for all the preference shocks, the $\epsilon_{ij}$’s, were known with the exception of those for $\{1, 2, 3\}$. 
The probability of choosing \{1, 2, 3\} could then be expressed as

\[
Pr(d_i = \{1, 2, 3\}) = Pr(\overline{U}_{i1} > M_i, \overline{U}_{i2} > M_i, \overline{U}_{i3} > M_i)
= Pr(\overline{U}_{i1} > M_i)Pr(\overline{U}_{i2} > M_i)Pr(\overline{U}_{i3} > M_i)
= (1 - G(M_i - \overline{U}_{i1}))(1 - G(M_i - \overline{U}_{i2}))
\times (1 - G(M_i - \overline{U}_{i3}))
\]

where \(G(\cdot)\) is the extreme value cdf and \(\overline{U}_{ij}\) is the flow payoff for \(j\) net of \(\epsilon_{ij}\).
Simulated maximum likelihood procedure

- Since the $\epsilon_{ij}$’s for the non-chosen courses are not observed, we integrate them out of the likelihood function and approximate the integral by simulating their values from the Type I extreme value distribution.

- Denoting $M_{ir}$ as the value of $M_i$ at the $r$th draw of the non-chosen $\epsilon_{ij}$’s and $R$ as the number of simulation draws, estimates of the reduced form payoffs come from solving:

$$
\max_{\phi, \delta} \sum_i \ln \left( \frac{ \sum_{r=1}^{R} \prod_{j=1}^{J} \left( 1 - G(M_{ir} - \bar{U}_{ij}) \right)^{d_{ij}} }{R} \right) \quad (7)
$$
Unobserved heterogeneity

- Assume individuals are one of $S$ unobserved types.
  - In practice, two.
- Unobserved type is a component of ability and through ability affects course selection.
- Estimation of grades is easy, but the choice model is hard.
- Following Arcidiacono and Miller (2011), use the EM algorithm with a reduced-from choice model that gives estimates of:
  1. the grade parameters
  2. the conditional probabilities of being each unobserved type
- Use the conditional type probabilities as weights to recover the study and choice parameters.
Demand-side estimates allow us to decompose differences in course choices, grades, and study effort between males and females into parts due to:

(i) differences in preferences (both over departments and over having a female professor),
(ii) differences in value of grades ($\phi_i$),
(iii) differences in study costs ($\psi_{ij}$),
(iv) differences in preparation ($A_{ik(j)}$),
(v) differences in grading practices ($\{\beta_j, \gamma_j\}$)
Professor payoffs are assumed to be a function of:

(i) the total amount of learning in the course: \( a(\beta, \gamma) \),
(ii) total enrollment: \( b(\beta, \gamma) \), and
(iii) student study time: \( c(\beta, \gamma) \).

where learning for individual \( i \) in course \( j \) is given by:

\[
L_{ij}(\gamma_j) = A_{ik}(j) + \ln(s_{ij}^*) \\
= A_{ik}(j) + \ln(\phi_i) + \ln(\gamma_j) - \ln(\psi_i)
\]  \( (8) \)
Professor Objective Function

\[ V_j(\beta, \gamma) = \lambda_{0j} a_j(\beta, \gamma) - \lambda_{1j} b_j(\beta, \gamma) - \lambda_{2j} c_j(\beta, \gamma) \]

\[ = \lambda_{0j} \left[ \sum_i P_{ij}(\beta, \gamma) \left( A_{ij} + \ln(\phi_i) + \ln(\gamma_j) - \ln(\psi_i) \right) \right] \]

\[ - \lambda_{1j} \left[ \sum_i P_{ij}(\beta, \gamma) \right]^2 - \lambda_{2j} \left[ \sum_i P_{ij}(\beta, \gamma) \left( \frac{\phi_i \gamma_j}{\psi_i} \right)^2 \right] \]

where we normalize \( \lambda_{0j} \) to one.
The choice of $\beta_j$ and $\gamma_j$ satisfy the two first order conditions:

\[
\frac{\partial V_j}{\partial \beta_j} = 0 = \frac{\partial a(\beta, \gamma)}{\partial \beta_j} - \lambda_1 j \frac{\partial b(\beta, \gamma)}{\partial \beta_j} - \lambda_2 j \frac{\partial c(\beta, \gamma)}{\partial \beta_j} \tag{9}
\]

\[
\frac{\partial V_j}{\partial \gamma_j} = 0 = \frac{\partial a(\beta, \gamma)}{\partial \gamma_j} - \lambda_1 j \frac{\partial b(\beta, \gamma)}{\partial \gamma_j} - \lambda_2 j \frac{\partial c(\beta, \gamma)}{\partial \gamma_j} \tag{10}
\]

The solution to this system is then:

\[
\begin{bmatrix}
\lambda_1 j \\
\lambda_2 j
\end{bmatrix} = \begin{bmatrix}
\frac{\partial b(\beta, \gamma)}{\partial \beta_j} & \frac{\partial c(\beta, \gamma)}{\partial \beta_j} \\
\frac{\partial b(\beta, \gamma)}{\partial \gamma_j} & \frac{\partial c(\beta, \gamma)}{\partial \gamma_j}
\end{bmatrix}^{-1} \begin{bmatrix}
\frac{\partial a(\beta, \gamma)}{\partial \beta_j} \\
\frac{\partial a(\beta, \gamma)}{\partial \gamma_j}
\end{bmatrix} \tag{11}
\]

By substituting in the estimates of the professor’s own grading practices, the estimates of the grading practices of the other professors, and the estimates of student preferences, the right hand side is known.
Counterfactuals

- How course choices, study time, and average grades would change if STEM professors had the same preferences as non-STEM professors.
- How grading practices would change in response to restriction on grading policies such as capping the fraction of high grades.
### Preference Parameters

<table>
<thead>
<tr>
<th>Preference for:</th>
<th>Coeff.</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected grades ($\phi$)</td>
<td>0.4157</td>
<td>(0.0199)</td>
</tr>
<tr>
<td>Female $\times$ expected grades</td>
<td>0.0759</td>
<td>(0.0157)</td>
</tr>
<tr>
<td>Female $\times$ female professor</td>
<td>0.1455</td>
<td>(0.0186)</td>
</tr>
<tr>
<td>$C_1$ (normalizing constant)</td>
<td>0.9581</td>
<td>(0.2593)</td>
</tr>
</tbody>
</table>

**Female preferences for Departments**

<table>
<thead>
<tr>
<th>Departments</th>
<th>Coeff.</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineering</td>
<td>-1.0615</td>
<td>(0.0737)</td>
</tr>
<tr>
<td>Econ., Fin., Acct.</td>
<td>-0.5091</td>
<td>(0.0593)</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>-0.2862</td>
<td>(0.0545)</td>
</tr>
<tr>
<td>Communication</td>
<td>-0.1528</td>
<td>(0.0537)</td>
</tr>
<tr>
<td>Chemistry &amp; Physics</td>
<td>-0.1482</td>
<td>(0.0599)</td>
</tr>
<tr>
<td>Languages</td>
<td>-0.1033</td>
<td>(0.0582)</td>
</tr>
<tr>
<td>Mathematics</td>
<td>-0.0072</td>
<td>(0.0688)</td>
</tr>
<tr>
<td>Mgmt. &amp; Mkting.</td>
<td>0.1153</td>
<td>(0.0662)</td>
</tr>
<tr>
<td>Regional Studies</td>
<td>0.2216</td>
<td>(0.0698)</td>
</tr>
<tr>
<td>Biology</td>
<td>0.2546</td>
<td>(0.0638)</td>
</tr>
<tr>
<td>Education &amp; Health</td>
<td>0.3287</td>
<td>(0.0581)</td>
</tr>
<tr>
<td>Psychology</td>
<td>0.3758</td>
<td>(0.0659)</td>
</tr>
<tr>
<td>English</td>
<td>0.4167</td>
<td>(0.0769)</td>
</tr>
</tbody>
</table>

Note: Agriculture normalized to zero.
### Results: Study Effort and Returns in Median Courses

<table>
<thead>
<tr>
<th></th>
<th>Coeff. ($-\psi$)</th>
<th>Std. Error</th>
<th>Department</th>
<th>Median Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-0.074</td>
<td>-0.080</td>
<td>Mathematics</td>
<td>3.59</td>
</tr>
<tr>
<td>ACTR (0’s)</td>
<td>-0.002</td>
<td>-0.010</td>
<td>Engineering</td>
<td>3.11</td>
</tr>
<tr>
<td>ACTM (0’s)</td>
<td>-0.018</td>
<td>-0.011</td>
<td>Biology</td>
<td>2.00</td>
</tr>
<tr>
<td>HS GPA</td>
<td>-0.004</td>
<td>-0.087</td>
<td>English</td>
<td>1.80</td>
</tr>
<tr>
<td>Black</td>
<td>0.234</td>
<td>-0.172</td>
<td>Chem &amp; Phys</td>
<td>1.78</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.317</td>
<td>-0.252</td>
<td>Psychology</td>
<td>1.64</td>
</tr>
<tr>
<td>Other Min.</td>
<td>-0.269</td>
<td>-0.261</td>
<td>Econ., Fin., Acct.</td>
<td>1.49</td>
</tr>
<tr>
<td>First Gen</td>
<td>0.107</td>
<td>-0.115</td>
<td>Regional Studies</td>
<td>1.49</td>
</tr>
<tr>
<td>Unobs. Type</td>
<td>0.198</td>
<td>-0.084</td>
<td>Communication</td>
<td>1.37</td>
</tr>
<tr>
<td><strong>Effort elasticity</strong></td>
<td></td>
<td></td>
<td>Languages</td>
<td>1.34</td>
</tr>
<tr>
<td>ln($\gamma$)</td>
<td>0.545</td>
<td>(0.206)</td>
<td>Social Sciences</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Agriculture</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mgmt. &amp; Mkting.</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Education &amp; Health</td>
<td>0.77</td>
</tr>
</tbody>
</table>
## Professor Preferences

<table>
<thead>
<tr>
<th>Level</th>
<th>Category</th>
<th>Enrollment $^2 (\lambda_{1j})$</th>
<th>Study Time $^2 (\lambda_{1j})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Lower</td>
<td>non-STEM</td>
<td>0.039</td>
<td>0.023</td>
</tr>
<tr>
<td>Level</td>
<td>STEM</td>
<td>0.023*</td>
<td>0.017</td>
</tr>
<tr>
<td>Upper</td>
<td>non-STEM</td>
<td>0.048*,†</td>
<td>0.022</td>
</tr>
<tr>
<td>Level</td>
<td>STEM</td>
<td>0.035*,†</td>
<td>0.016</td>
</tr>
</tbody>
</table>
Examine how the STEM gap changes given five changes to the environment:

1. No differences between men and women in their preferences for grades ($\phi_1 = 0$)
2. No differences in unobserved abilities ($\alpha_1 = 0$)
3. No differences in observed abilities ($\overline{X}_f = \overline{X}_m$)
4. No differences in tastes for departments
5. No differential preference for female professors
6. Equalized expected grades across courses for the average student
PE and GE responses to changing female tastes and abilities on share STEM

<table>
<thead>
<tr>
<th>Turn off...</th>
<th>PE Female Increase Over Base</th>
<th>GE Female Increase Over Base</th>
<th>GE/PE</th>
<th>GE Male Increase Over Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) grade prefs</td>
<td>0.94%</td>
<td>0.26%</td>
<td>27.5%</td>
<td>-0.43%</td>
</tr>
<tr>
<td>(2) gender ability</td>
<td>3.78%</td>
<td>1.52%</td>
<td>40.3%</td>
<td>-1.79%</td>
</tr>
<tr>
<td>(3) obs ability</td>
<td>0.50%</td>
<td>0.37%</td>
<td>74.4%</td>
<td>-0.15%</td>
</tr>
<tr>
<td>(4) tastes</td>
<td>1.54%</td>
<td>0.78%</td>
<td>50.2%</td>
<td>-0.92%</td>
</tr>
<tr>
<td>(5) female prof pref</td>
<td>0.40%</td>
<td>0.16%</td>
<td>39.1%</td>
<td>-0.22%</td>
</tr>
</tbody>
</table>

Baseline female: 28.40%
Baseline male: 40.02%
Supply-side Counterfactuals

Examine how changing grading practices affects the STEM gap by:

1. Setting expected grades to be the same across classes for the average student by shifting course intercepts

2. Adjust STEM professor preferences so that the averages are similar across STEM and non-STEM professors and solving for the new equilibrium grading policies
Supply-side Counterfactual Results on Share STEM

<table>
<thead>
<tr>
<th></th>
<th>Male Share</th>
<th>Female Share</th>
<th>Gender Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.400</td>
<td>0.284</td>
<td>0.710</td>
</tr>
<tr>
<td>(6) Equalize exp grade for average student</td>
<td>0.440</td>
<td>0.334</td>
<td>0.760</td>
</tr>
<tr>
<td>(7) Change STEM prof prefs</td>
<td>0.235</td>
<td>0.126</td>
<td>0.536</td>
</tr>
</tbody>
</table>

Gender ratio = Women/Men
Preliminary evidence suggests men and women respond differently to grading practices.

A cheap way of changing the STEM gap may be to change grading practices.

May be important to take into account how professors respond to university regulations on grading policies.