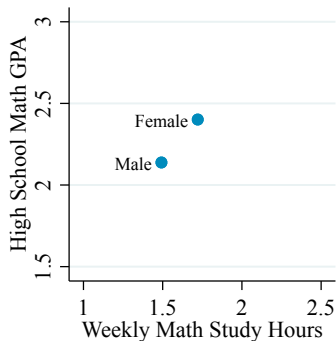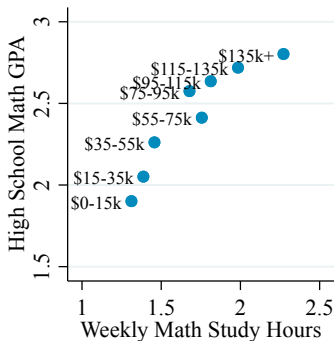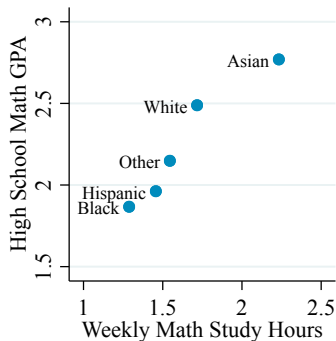# Beliefs, Preferences, and Student Effort

John J. Conlon   Spencer Kwon   William Murdock III   Dev Patel

NBER Education Program Meeting
December 3rd, 2021

Achievement Gaps Correspond to Large Gaps in Education "Inputs"

"During a typical school week, how many hours do you spend working on math homework and studying for math class?"

Source: Authors' calculations based on data from High School Longitudinal Study of 2009 ⊖ *Decomposition*

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:
1. Beliefs about the productivity of studying

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:

1. Beliefs about the productivity of studying
2. Opportunity cost of studying

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:

1. Beliefs about the productivity of studying
2. Opportunity cost of studying
3. Value of grades

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:
1. Beliefs about the productivity of studying
2. Opportunity cost of studying
3. Value of grades

These reduced-form beliefs & preferences are certainly functions of many other factors (e.g., teachers, schools, peers, home environment)

What Drives Gaps in Student Effort?

Research Question: What drives these differences in student effort?

We consider three dimensions:
1. Beliefs about the productivity of studying
2. Opportunity cost of studying
3. Value of grades

These reduced-form beliefs & preferences are certainly functions of many other factors (e.g., teachers, schools, peers, home environment)

What we do:
- We administered surveys to 2,500 US high school students
  - Elicited preferences over grades and effort
  - Elicited beliefs about the returns (in better grades) to effort

Background and Related Work

- Growing evidence of the importance of effort in school, and its causal effect on achievement (e.g., Romer 1993; Stinebrickner and Stinebrickner 2003, 2004, 2008; Metcalfe et al. 2019)

- Has spurred interest in increasing effort through incentives and information provision (e.g., Fryer 2011; Bettinger et al. 2012; de Paola et al. 2012; Fryer and Holden 2013; Blimpo 2014; Gneezy et al. 2019; Cotton et al. 2020; Rury et al. 2020; Oreopoulos et al. 2020; Ersoy 2019)

- And in understanding what underlies students studying decisions in different contexts (e.g., Delavande et al. 2020; Cotton et al. 2020)

Today's Talk

Research Question: What drives differences in student effort?

- Explore reduced-form heterogeneity in preferences and beliefs

- Decompose gaps in effort into variation in perceived efficacy of studying, value for grades, and dislike for studying

Today's Talk

Research Question: What drives differences in student effort?

- Explore reduced-form heterogeneity in preferences and beliefs

- Decompose gaps in effort into variation in perceived efficacy of studying, value for grades, and dislike for studying

- Simulate how our sample of students would respond to various incentive schemes

Today's Talk

Research Question: What drives differences in student effort?

- Explore reduced-form heterogeneity in preferences and beliefs

- Decompose gaps in effort into variation in perceived efficacy of studying, value for grades, and dislike for studying

- Simulate how our sample of students would respond to various incentive schemes

- Will mostly focus on race today, but also have results for gender and SES

Toy Model

Assume individual utility is quasi-linear in cash, M, and separable in hours and grades:

$$u_i(M, G, H) = M + \psi_i(G) - v_i(H), \qquad (1)$$

where G is grade and H is hours of studying.

Toy Model

Assume individual utility is quasi-linear in cash, M, and separable in hours and grades:

$$u_i(M, G, H) = M + \psi_i(G) - \nu_i(H), \tag{1}$$

where G is grade and H is hours of studying.

Students choose H for a specific class to maximize:

$$\hat{H}_i = \arg\max_H \sum_{G \in \mathcal{G}} \overbrace{\psi_i(G)}^{\text{value of grade}} \cdot \underbrace{\pi_i(G|H)}_{\text{prob. of grade}} \overbrace{- \nu_i(H)}^{\text{cost of studying}} \tag{2}$$

Surveying Students

- Partner with Character Lab Research Network to survey high schoolers from a large urban public school district in Florida.
    - Take 25-minute online surveys *during school hours*

- **Survey 1, February 2020**: Main preferences and belief elicitation. 2,501 complete survey and pass attention checks

- **Survey 2, October 2020**: Additional covariates. 523 repeat respondents (major attrition due to pandemic)

- Administrative data on demographics (race, gender, FRPL) and transcripts

⊖ *Sample Across Waves*

Survey Sample

	Survey Sample	HSLS Sample
	(1)	(2)
Grades	9-11	11
Female	0.50	0.50
Underrep. Minority	0.63	0.42
Free Lunch	0.51	0.48
Math GPA	2.39	2.22
Typical Math Study Hours	1.83	1.51
N	2,501	18,247

$$\text{URM} = \{\text{American Indian, Black, Hispanic, Multiracial}\}$$
$$\text{Non-URM} = \{\text{Asian, White}|\text{Non-Hispanic}\}$$

Expected Studying Next Semester

"During a typical school week next semester, what do you think is the most likely number of hours you will spend working on homework and studying for math (English) class?"

- We framed everything as next semester to avoid capturing any semester-specific information

- Otherwise identical to studying question from HSLS

- Each student always asked about either math or English. For today's talk, we pool across subjects

Expected Academic Gaps by Race in Our Sample
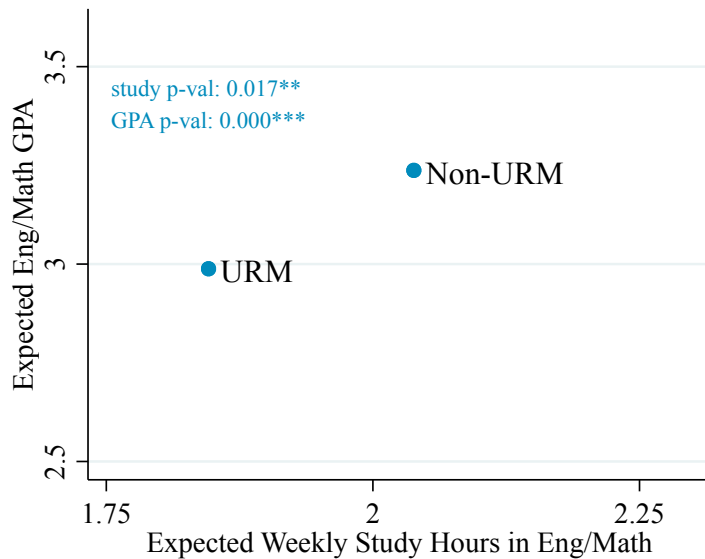
Measuring Beliefs and Preferences

We will now walk through the survey measures we used to capture the three components of our model:

1. Beliefs about the productivity of studying
2. Perceived opportunity cost of studying
3. Perceived value of grades

Along the way, we will summarize the differences across races we observe for each measure

Beliefs Elicitation: Hypothetical Studying Scenarios

"Imagine that you had to spend exactly *X* hours per week studying and working on homework for your math (English) class. This could be either in one continuous *X* hour block each week, or you could break up the time across the days however you want."

	0 hours	2 hours
A- to A+	45 %	%
B- to B+	35 %	%
C- to C+	20 %	%
D+ or less	0 %	%
Total	100 %	0 %

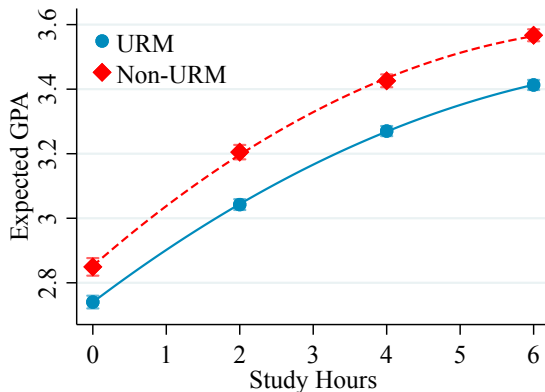Note that because there is a 100 percent chance that you would receive some grade, your answers in each column must add to 100.

⊖ *Probability Practice*

How Much Does Studying Improve Grades?
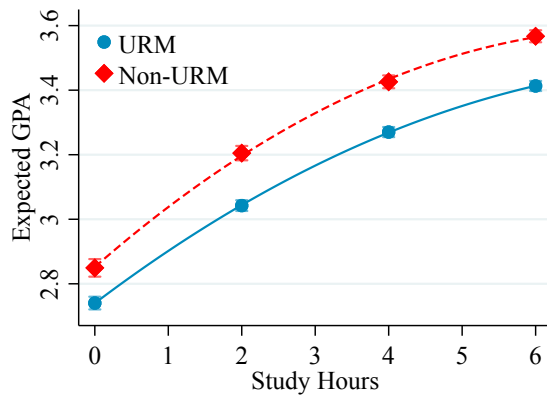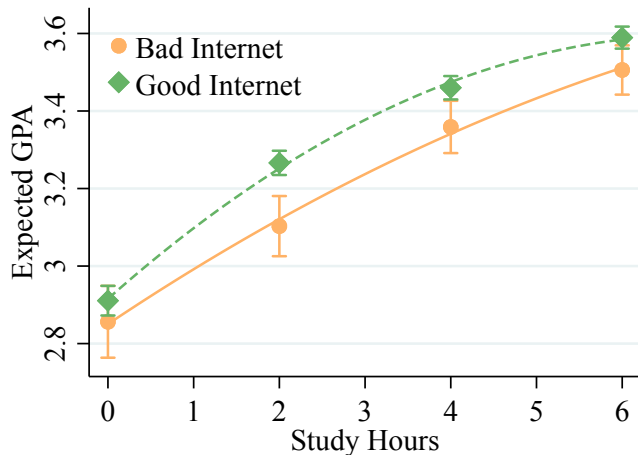

Bands show standard errors.

How Much Does Studying Improve Grades?
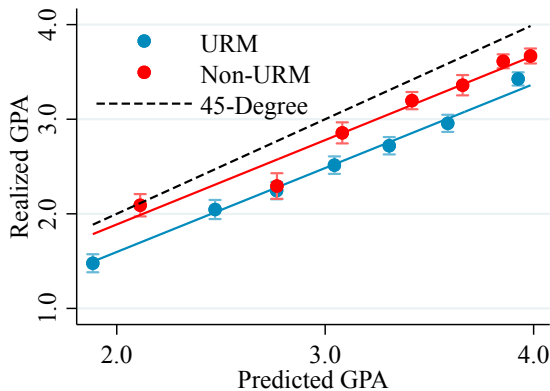

Bands show standard errors.

- We found these curves surprisingly flat: if a student wants to increase their grade from a C+ to a B+, they think they need to study more than 10 hours *per week* to do this

- Slopes, or marginal returns to studying, *are* different across racial groups

Students' Beliefs Intuitively Correlate with Other Measures

Bands show standard errors. ⊖ *More*

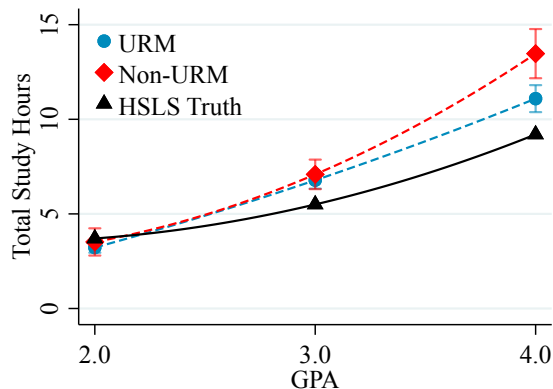# Students Can Predict Their Own Future Grades

Bands show standard errors.

- Predicted and realized grades are strongly correlated, but evidence of overestimation (below the 45-degree line), with URMs having more than non-URMs

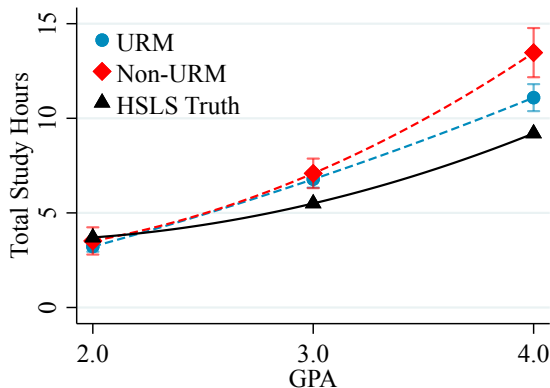- Students are willing to report to researchers that they expect low grades

Students May Be Underestimating True Returns to Studying

Bands show standard errors. ⊖ *Instructions*

Students May Be Underestimating True Returns to Studying

- Our students correctly perceive that students who get higher grades in the HSLS also study more
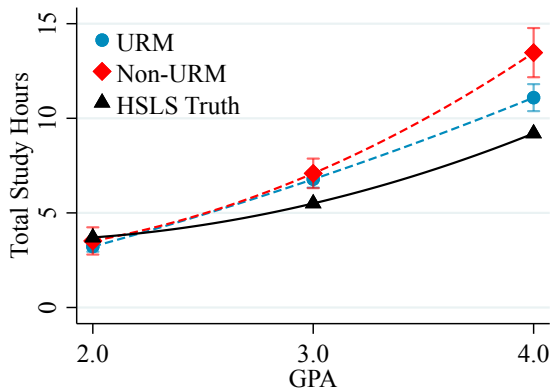- But they appear to overstate this relationship

Bands show standard errors. ⊖ *Instructions*

Students May Be Underestimating True Returns to Studying

- Our students correctly perceive that students who get higher grades in the HSLS also study more
- But they appear to overstate this relationship
- If naively taken to be causal, this suggests that students in our sample may be underestimating the returns to studying
- Differential underestimation by race $\Rightarrow$ true differences in studying effectiveness may be larger than reported

Bands show standard errors. ⊖ *Instructions*

Eliciting Preferences for Study Hours: Multiple Price Lists

Trying to capture $v_i(H)$: the intrinsic disutility of extra study hours abstracting away from effect on grades

Which of the following would you prefer?

Be paid **\$39** per week, **2** extra hours of math schoolwork per week ○ ○ Be paid **\$20** per week, **0** extra hours of math schoolwork per week
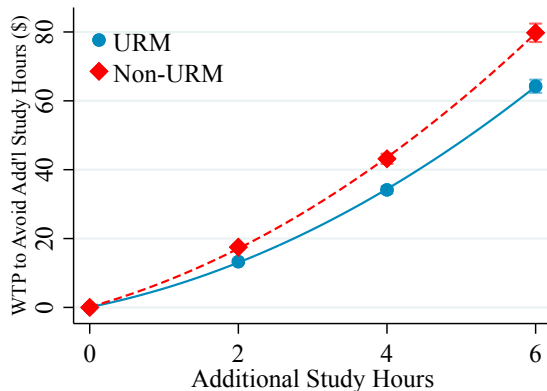
- Dollar amounts change dynamically to zoom into students' indifference points

- Time spent on non-graded extra schoolwork

- Told grades and study hours for *all* classes would remain the same to capture intrinct disutility from studying another hour

URMs Perceive Study Hours As Less Costly
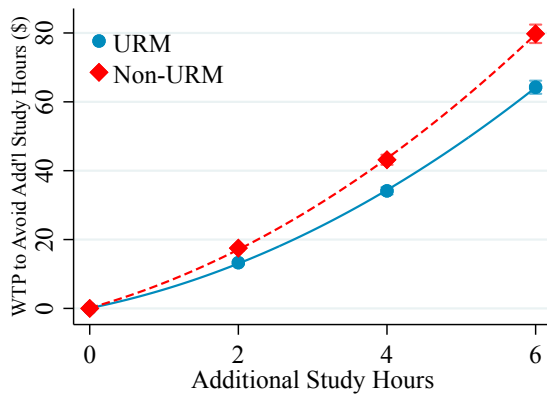
Bands show standard errors.
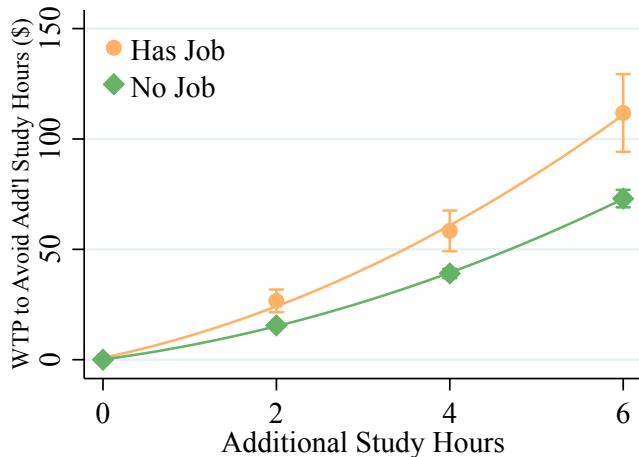
URMs Perceive Study Hours As Less Costly

- Students are WTP ≈ $7.5 per hour to avoid first two additional study hours

- Federal minimum wage = $7.25

- Convexity ⇒ Students WTP ≈$12.50 per hour to avoid six additional study hours

Bands show standard errors.

Students' Preferences Intuitively Correlate with Other Measures

Bands show standard errors.

Eliciting Preferences for Grades: Multiple Price Lists

Trying to capture $\psi_i(G)$: the intrinsic value of grades for the class in question

Which of the following would you prefer?

Study **5** hours per week for math, Get a **B**   ○ ○   Study **1** hour per week for math, Get a **C**
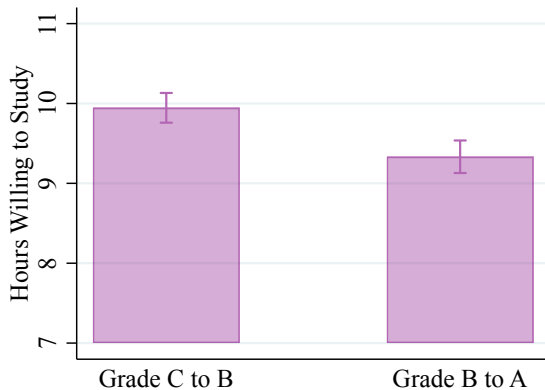
- Hours changed dynamically to zoom in on students' indifference points

- Told grades and study hours for *other* classes would remain the same to isolate value of that grade

Average Student Willing to Study *A Lot* To Increase Grades

- Students are willing to study about 9.5 hours *per week* to increase their class grade by one letter
- So why don't they study that much?

Average Student Willing to Study *A Lot* To Increase Grades
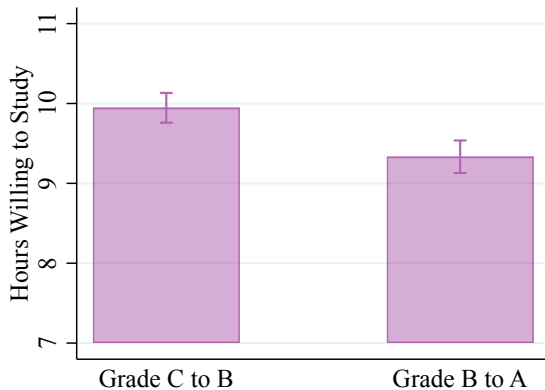
Bands show standard errors.

- Students are willing to study about 9.5 hours *per week* to increase their class grade by one letter
- So why don't they study that much?
- They just don't think studying is that effective. Recall, the average student thinks it will take more than 10 hours to increase grades by a letter grade

Average Student Willing to Pay *A Lot* To Increase Grades
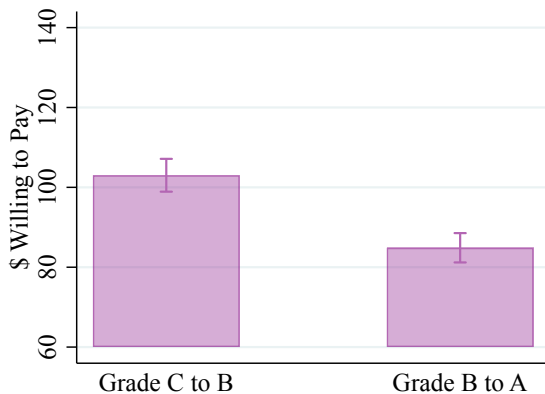
Bands show standard errors.

- This translates to about $90 *per week*

- Or ≈$2,700 per academic year

- Correlational estimates suggest a one-point increase in overall high school GPA is associated with a 12-14% bump in annual adulthood earnings (French et al., 2015)

URMs Willing to Pay Less For Grades
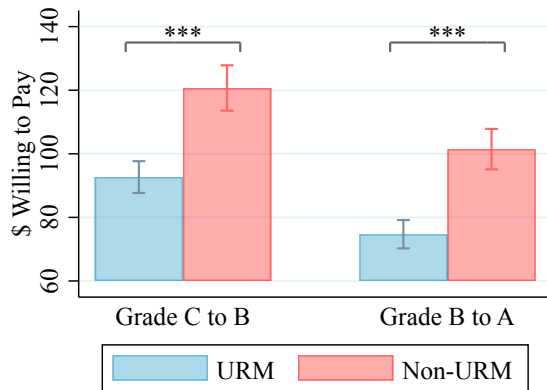
Bands show standard errors.

- URMs willing to pay about $25 less, per week, for each letter grade

- This is ≈ $800 per semester

Students' Preferences Intuitively Correlate with Other Measures
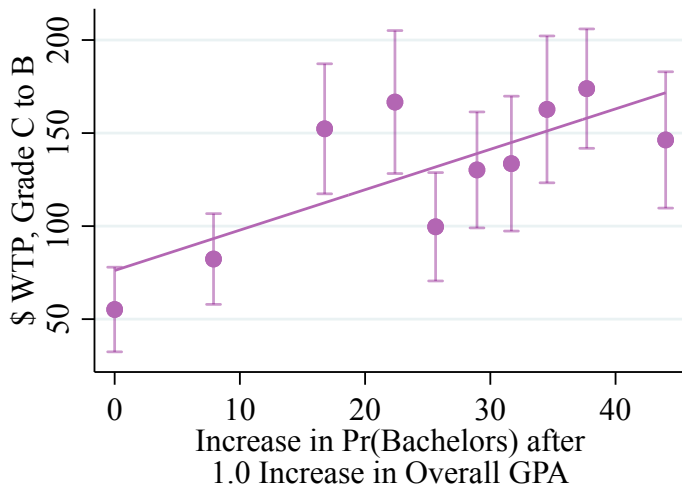
Bands show standard errors. ⊖ *Instructions* ⊖ *More* ⊖ *Pr(Bachelors) Dist.*

Summary of Differences

URMs, relative to non-URMS:

- Perceive studying to be **less** effective
- Perceive study hours as **less** costly
- Willing to study **less** to increase grades

⊖ *SES + Gender*

Combining Our Measures in Toy Model to Predict Behavior

Combining Our Measures in Toy Model to Predict Behavior

- Recall, students solve:

$$\hat{H}_i = \arg \max_H \sum_{G \in \mathcal{G}} \pi_i(G|H) \cdot \psi_i(G) - \nu_i(H)$$

- Get $\pi_i(G|H)$ and $\nu_i(H)$ directly from elicitations

- Our elicitation of grade valuation was "hours wiling to study to increase grade from G_1 to G_2"
    - Normalize $\psi_i(C) = 0$
    - Use $v_i(H)$ to convert to \$'s

- We elicited $\pi_i(G|H)$, $v_i(H)$ for $H \in \{0, 2, 4, 6\}$ and $\psi_i(G)$ for $G \in \{B, A\}$

- Interpolate to values of H and G we don't directly elicit

Predicting Studying Behavior

We now have:

- $\pi_i(G|H)$: beliefs about grades conditional on study hours

- $\nu_i(H)$: disutility from study hours

- $\psi_i(G)$: utility from grades

In our toy model, students solve:

$$\hat{H}_i = \arg \max_H \sum_{G \in \mathcal{G}} \pi_i(G|H) \cdot \psi_i(G) - \nu_i(H)$$

We solve it too!

Model is Predictive of Study Habits, but with Attenuation
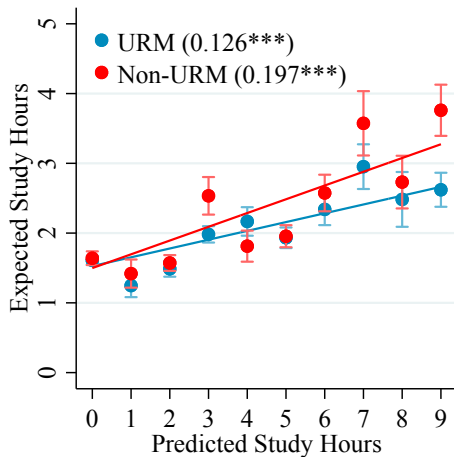
Model is Predictive of Study Habits, but with Attenuation
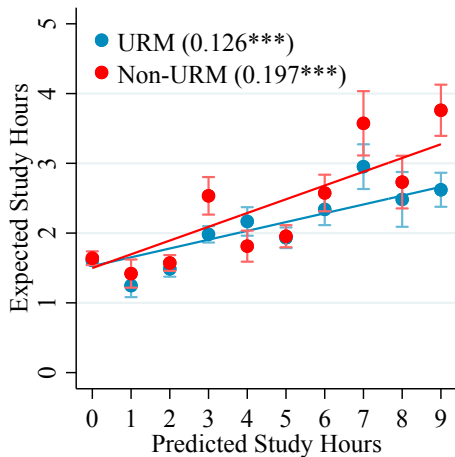
Bands show standard errors.

We find two major drivers of prediction error

- Measurement error – Finding the survey difficult is associated with greater prediction error

- Agency over studying choices – Parents desiring more studying is associated with greater prediction error

Unpacking the Study Effort Gap

Unpacking the Study Effort Gap

- How much are gaps in study behavior driven by gaps in:
    - **beliefs** about the returns to studying
    - **preferences** for marginal study hours
    - **preferences** for grades

- **Conceptual exercise:** Give URMs the preferences/beliefs of non-URMs, re-solve toy model

Unpacking the Study Effort Gap

- How much are gaps in study behavior driven by gaps in:
    - **beliefs** about the returns to studying
    - **preferences** for marginal study hours
    - **preferences** for grades

- **Conceptual exercise:** Give URMs the preferences/beliefs of non-URMs, re-solve toy model

- Rank-rank distribution matching: assign p^{th} percentile URM the pth percentile value of non-URMs

What is Driving Differences Across Race in Study Effort?
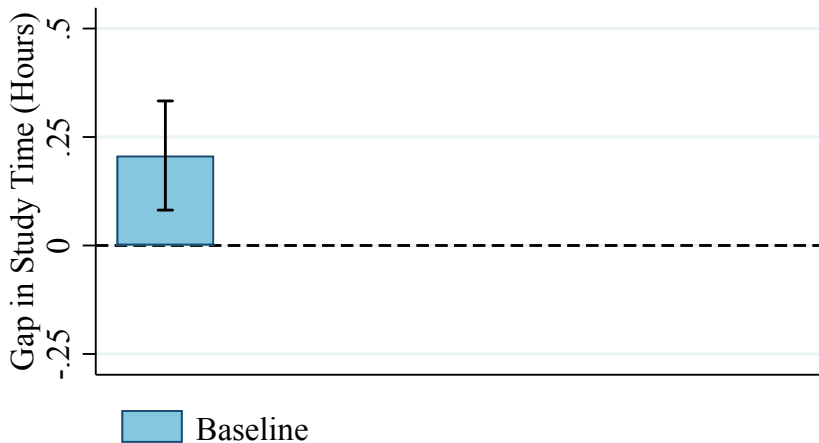
What is Driving Differences Across Race in Study Effort?
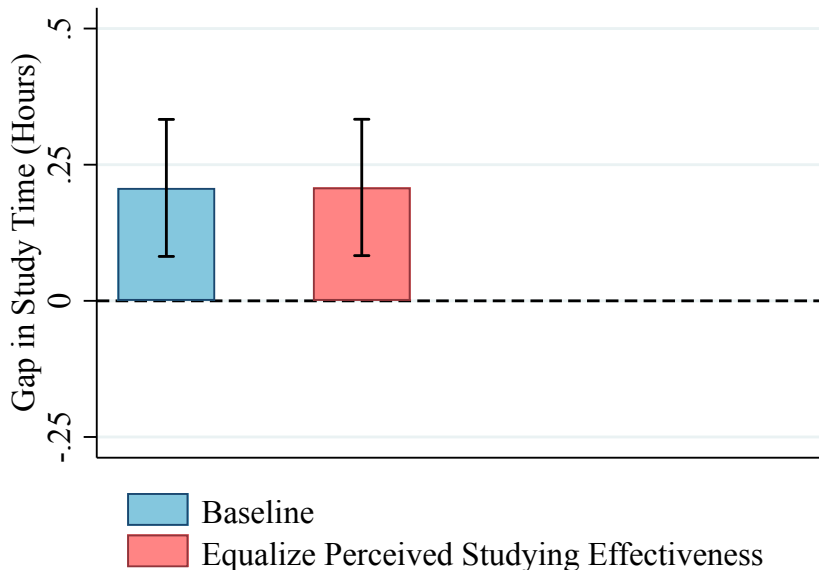
What is Driving Differences Across Race in Study Effort?
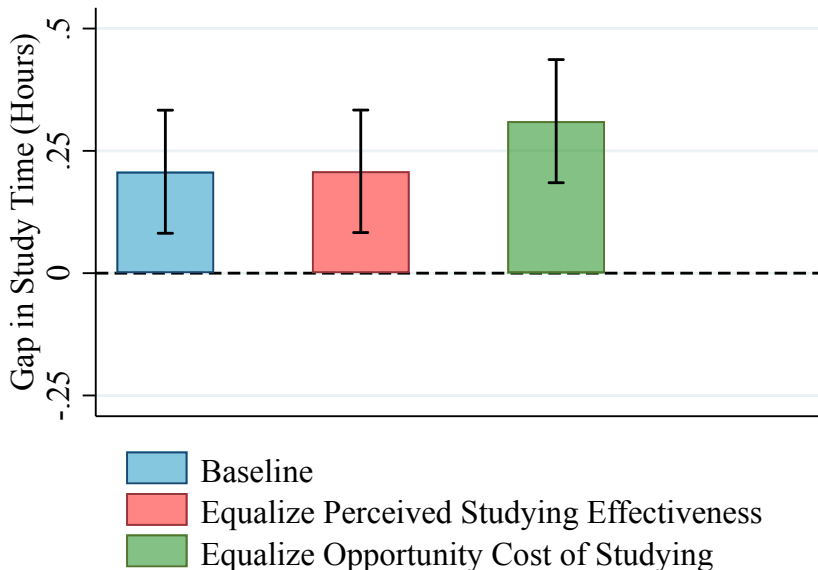
What is Driving Differences Across Race in Study Effort?
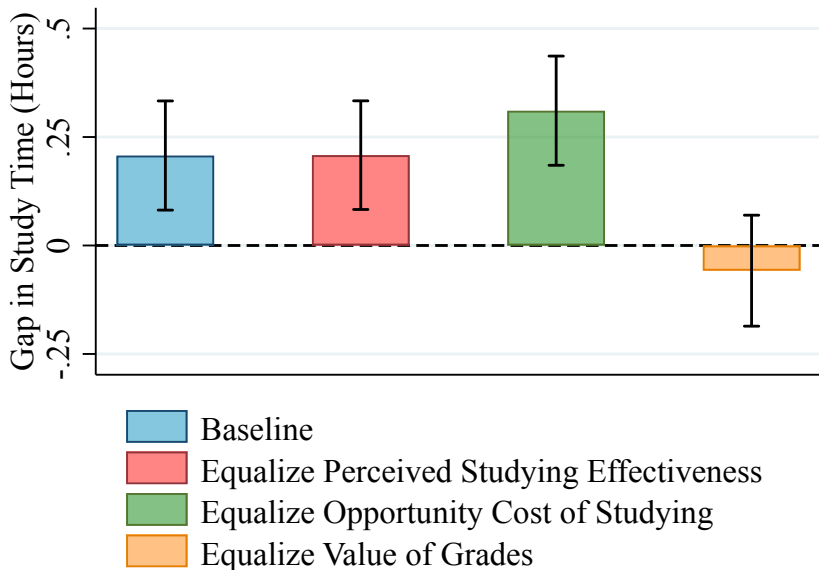
What Would It Take to Close the Achievement Gap? Paying for Grades

What Would It Take to Close the Achievement Gap? Paying for Grades

- Fryer (2011) paid 9th graders from Chicago Public Schools $80 for a C, $140 for a B, and $200 for an A, per semester per class

	Fryer (2011)	Our Setup
Grades	0.08σ	
Effort	0.15σ	

What Would It Take to Close the Achievement Gap? Paying for Grades

- Fryer (2011) paid 9th graders from Chicago Public Schools $80 for a C, $140 for a B, and $200 for an A, per semester per class

	Fryer (2011)	Our Setup
Grades	0.08σ	0.07σ
Effort	0.15σ	0.10σ

What Would It Take to Close the Achievement Gap? Paying for Grades

- Fryer (2011) paid 9th graders from Chicago Public Schools $80 for a C, $140 for a B, and $200 for an A, per semester per class

	Fryer (2011)	Our Setup
Grades	0.08σ	0.07σ
Effort	0.15σ	0.10σ

In our sample, targeting just URMs:

- These incentives would completely close the study gap. But minimal impact on grades

- Need to increase these incentives 8-fold in order to fully close the grade gap $\Rightarrow$ In expectation, would need to spend $\approx \$1,200$ per student per semester per class

What Would It Take to Close the Achievement Gap? Paying for Effort

- **Incentive Scheme:** Pay each URM student X for each hour of studying they do

What Would It Take to Close the Achievement Gap? Paying for Effort

- **Incentive Scheme:** Pay each URM student $X for each hour of studying they do

- To close grade gap in a specific class, we'd need to pay URMs $6.25 per hour of studying

- This would increase URMs' study hours by 2.8 hours and cost $500 per student per semester

What Would It Take to Close the Achievement Gap? Paying for Effort

- **Incentive Scheme:** Pay each URM student $*X* for each hour of studying they do

- To close grade gap in a specific class, we'd need to pay URMs $6.25 per hour of studying

- This would increase URMs' study hours by 2.8 hours and cost $500 per student per semester

- The increase in study hours needed to close the achievement gap is similar in magnitude to high-dosage tutoring interventions which typically find large impacts on achievement (Fryer, 2017)

What Would It Take to Close the Achievement Gap? Paying for Effort

- **Incentive Scheme:** Pay each URM student $X for each hour of studying they do

- To close grade gap in a specific class, we'd need to pay URMs $6.25 per hour of studying

- This would increase URMs' study hours by 2.8 hours and cost $500 per student per semester

- The increase in study hours needed to close the achievement gap is similar in magnitude to high-dosage tutoring interventions which typically find large impacts on achievement (Fryer, 2017)

- Inframarginal students drive up the cost of these simple pay-for-effort and pay-for-grade policies. If we could target the marginal study hours / grades perfectly, both would cost closer to $225

What Would It Take to Close the Achievement Gap? Returns to Effort

- The incentive simulations took students beliefs about the efficacy of their studying at face value

- Large increase in study hours needed due to low (perceived) returns to studying

What Would It Take to Close the Achievement Gap? Returns to Effort

- The incentive simulations took students beliefs about the efficacy of their studying at face value

- Large increase in study hours needed due to low (perceived) returns to studying

- This suggests increasing the perceived (or real) returns to effort could be especially successful policies to explore

- Remains to be seen whether students hold accurate beliefs, but if they are underestimating, information interventions could be very cost effective

Conclusion

- Administer a new survey to high school students to carefully elicit
    - **beliefs** about the returns to studying
    - **preferences** for marginal study hours
    - **preferences** for grades

- We find substantial differences by race along these dimensions

- Simulations suggest that differential perceived benefits of receiving higher grades is the primary driver of the studying gap

Thank you!

wmurdock@g.harvard.edu

References I

Bettinger, E. P., B. Terry Long, P. Oreopoulos, and L. Sanbonmatsu, "Information in College Decisions : Results From," *The Quarterly Journal of Economics*, 2012, pp. 1205–1242.

Blimpo, Moussa P., "Team incentives for education in developing countries: A randomized field experiment in Benin," *American Economic Journal: Applied Economics*, 2014, *6* (4), 90–109.

Cotton, Christopher, Brent R. Hickman, John A. List, Joseph Price, and Sutanuka Roy, "Producticity Versus Motivation in Adolescent Human Capital Production: Evidence from a Structurally-Motivated Field Experiment," *NBER Working Paper 27995*, 2020.

de Paola, Maria, Vincenzo Scoppa, and Rosanna Nisticò, "Monetary incentives and student achievement in a depressed labor market: Results from a randomized experiment," *Journal of Human Capital*, 2012, *6* (1), 56–85.

Delavande, Adeline, Emilia Del Bono, and Angus Holford, "Academic and non-academic investments at university: The role of expectations, preferences and constraints," *Journal of Econometrics*, 2020.

Ersoy, Fulya, "Effects of Perceived Productivity on Study Effort: Evidence from a Field Experiment," Technical Report 2019.

French, Michael T., Jenny F. Homer, Ioana Popovici, and Philip K. Robins, "What You Do in High School Matters: High School GPA, Educational Attainment, and Labor Market Earnings as a Young Adult," *Eastern Economic Journal*, 2015, *41*, 370–386.

Fryer, Roland G., "Financial incentives and student achievement: Evidence from randomized trials," *Quarterly Journal of Economics*, 2011, *126* (4), 1755–1798.

Fryer, Roland G, "The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments," *In: Handbook of Field Experiments*, 2017, *2*, 95–322.

__ **and Richard T Holden**, "Multitasking, Dynamic Complementarities, and Incentives: A Cautionary Tale," 2013.

Gneezy, Uri, John A. List, Jeffrey A. Livingston, Xiangdong Qin, Sally Sadoff, and Yang Xu, "Measuring Success in Education: The Role of Effort on the Test Itself," *American Economic Review: Insights*, 2019, *1* (3), 291–308.

References III

Metcalfe, Robert, Simon Burgess, and Steven Proud, "Students' effort and educational achievement: Using the timing of the World Cup to vary the value of leisure," *Journal of Public Economics*, 2019, *172*, 111–126.

Oreopoulos, Philip, Richard W Patterson, Uros Petronijevic, and Nolan G Pope, "Lack of Study Time is the Problem, but What is the Solution? Unsuccessful Attempts to Help Traditional and Online College Students," *Journal of Human Resources*, 2020.

Romer, David, "Do Students Go to Class? Should They?," *Journal of Economic Perspectives*, 1993, *7* (3), 167–174.

Rury, Derek, U C Davis, Scott Carrell, and U C Davis, "Knowing What It Takes: The Effect of Information About Returns to Studying on Study Effort and Achievement," 2020.

Stinebrickner, Ralph and Todd R. Stinebrickner, "Working during School and Academic Performance," *Journal of Labor Economics*, 2003, *21* (2), 473–491.

__ **and** __ , "Time-use and college outcomes," *Journal of Econometrics*, 2004, *121* (1-2), 243–269.

__ **and** __ , "The causal effect of studying on academic performance," *B.E. Journal of Economic Analysis and Policy*, 2008, *8* (1).

Achievement Gap Decomposition ⊖ *Back*
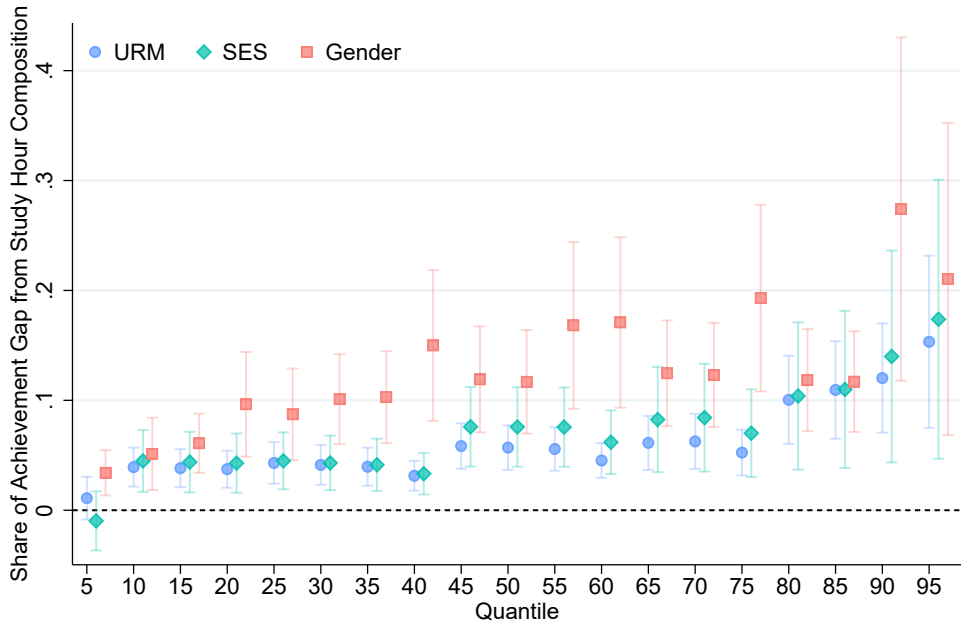
Students' Beliefs Intuitively Correlate with Other Measures

Efficacy of studying is...

- positively correlated with students' home internet quality

- negatively correlated with receiving tutoring, students' assessments of neighborhood safety, hours spent socializing with family members

Expected GPA without studying is...

- positively correlated with parental education, biological parents being married, size of home

- negatively correlated with having a job, hours worked per week

WTP to avoid additional study hours is...

- positively correlated with having trouble focusing in class, parental education, college affordability, having a job

- negatively correlated with frequency parents help with homework, college desirability

WTP for higher grades is...

- positively correlated with perceived earnings returns to GPA, perceived BA probability returns to GPA, parental education, lack of obstacles to college attendance

- negatively correlated with hours spent socializing with friends, being frustrated with homework, having trouble focusing in class/at home

Instructions 1

In the Spring of 2012, the U.S. Department of Education conducted a survey of American high school students. Students were recruited for this survey such that survey participants were representative of the entire country - their responses can be used to think about the average high schooler in the United States. In the questions below, we will ask you about this sample of students.

The national survey asked students when they were juniors how many hours they spend during a typical week studying for math class, for science class, and for all other classes. We add the answers to these three questions together to get the **total hours a student studies in a typical week**.

Instructions 3 ⊖ *Back*

The same survey also followed these students for multiple years and observed all the grades they received in high school.

What do you think was the **total number of hours studied in a typical week** for the average high school junior who ended high school with a cumulative GPA of...

2.0? [   ] hours

3.0? [   ] hours

4.0? [   ] hours

Recall that a student with a cumulative GPA of 2.0 on average got C's, a 3.0 GPA on average got B's, and a 4.0 GPA on average got A's.

Survey Question ⊖ *Back*

What do you think is the percent chance that you would complete a Bachelor's (4-year college) degree if you ended high school with a cumulative GPA of...

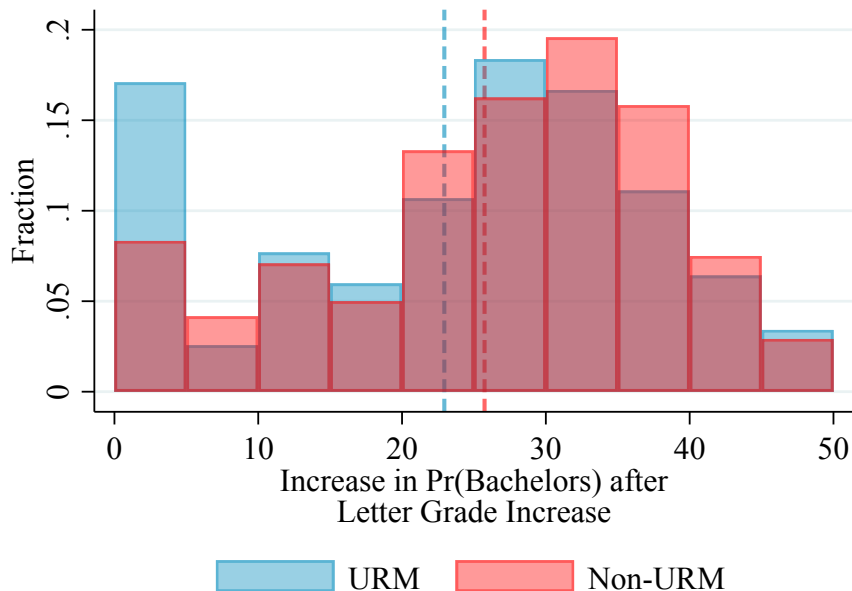| 0	10	20	30	40	50	60	70	80	90	100

2.0?

3.0?

4.0?

Attrition Across Survey Waves ⊖ *Back*

	Feb. Sample	Oct. Sample
	(1)	(2)
Female	0.50	0.54*
Underrep. Minority	0.63	0.51***
Free Lunch	0.51	0.39***
GPA	2.61	2.92***
Study Hours	1.72	1.76
N	2,501	523

Summary of Differences

	URM	Low SES	Male
Perceive studying to be ____ effective	less	less	equally
Perceive study hours as ____ costly	less	less	more
Willing to study ____ to increase grades	less	less	less

Wait. Can High Schoolers Report Probabilities?

- Practice modules introducing students to every concept throughout the survey

- 80-90% submitted practice problems correctly on first attempt

- Explanation of percentages from NY Fed's Survey of Consumer Expectations

- 90% reported that giving percentages about their own future was not difficult

- 90+% of students gave percentages other than 0 and 100 when multiple bins