SLIDES: Productivity Versus Motivation: Using Field Experiments and Structural Econometrics to Better Understand Adolescent Human Capital Production. *(manuscript coming soon)*

Brent Hickman (University of Chicago)

John List (University of Chicago)

Joseph Price (Brigham Young University)

Christopher Cotton (Queen's University)

The researcher would often like to answer questions that depend on individual characteristics that are either fundamentally unobservable or hard to quantify. **EXAMPLES:**

- **What are the factors underlying academic success?**
  - Learning Productivity
  - Motivation
  - Environmental Variables
- *Related questions:*
  - How best to intervene (at the individual level)?
  - Where do demographic academic disparities come from (at the group level)?

FIELD EXPERIMENTAL RESEARCH DESIGN: We do not aim to test a policy effect, but rather, *to generate observables which will identify a structural model of decision-making based on agent heterogeneity*:

- θ_E "Efficiency/Productivity Type"
  - Indexes some aspect of ability
- θ_L "Leisure Preference/Motivation Type"
  - Measures incentives required to induce a certain level of output, holding θ_E fixed

General Model Overview

- **PRINCIPAL:** values some output O by agents
  - Accomplishment of tasks $O = Q$
  - Accumulated proficiency $O = S$
- **PRODUCTION TECHNOLOGY:** $\tau(O; \theta_E)$
  - $t = \tau(o; \theta_E)$ time units required for output o, given θ_E
- **PRODUCTION COSTS:** $C(O; \theta_E) = \theta_L c(t)$
- **LABOR CONTRACT:** $(O, P(O))$
- **AGENT DECISION PROBLEM:**

$$\max_{o \in \mathbb{R}_+} \{P(o) - \theta_L c [\tau(o; \theta_E)]\}$$

Research Goal: To develop an experimental design to track ***total output*** *AND* ***time inputs*** which, combined with **exogenous contract variation**, are enough to identify τ, c, and the joint distribution of (θ_E, θ_L)

HOW TO TELL PRODUCTIVITY/MOTIVATION APART?

IDEAL EXPERIMENT: Consider an avg agent named Suzie.

1. Clone Suzie
2. Lock the two clones in separate, identical observation rooms for a month. Each room contains:
   - A couch and entertainment center with an Xbox and TV,
   - A desk with a set of work tasks to accomplish.
3. Offer different incentives to spend time on work instead of leisure
   - Offer Clone #1 $5 per unit of output (tasks accomplished)
   - Clone #2 gets $10 per unit.
4. Monitor Suzies'
   - Time choices,
   - Rate of progress,
   - Total output.

HOW TO TELL PRODUCTIVITY AND MOTIVATION APART?

So we can't exactly do that, but we can come close!

1. "Clone" a large group of agents through randomization
   - Offer each group different incentives to work
2. Place working materials *online*
   - Suzie can take the observation room with her...
   - More natural, less invasive
3. Progress/time tracking using built-in online capability

STRUCTURAL IDENTIFICATION SKETCH

NEEDED INGREDIENTS:

(1) Joint Distribution of inputs (time) and outputs (completed tasks) AND

(2) exogenous contract variation

STRUCTURAL IDENTIFICATION SKETCH

NEEDED INGREDIENTS:

(1) Joint Distribution of inputs (time) and outputs (completed tasks) AND

(2) exogenous contract variation

1. θ_E
   - Standard panel data idea: by observing individual cumulative production histories, we get a *WITHIN-CHILD TIME SERIES*
   - *e.g.*, Specify Cobb-Douglas production technology

$$\tau(q;\theta_E) = \delta_1 \left(\frac{q}{\theta_E}\right)^{\delta_2} \times u$$

   and estimate θ_E as a child-specific fixed effect

STRUCTURAL IDENTIFICATION SKETCH

NEEDED INGREDIENTS:

(1) Joint Distribution of inputs (time) and outputs (completed tasks) AND

(2) exogenous contract variation

1. θ_E
   - Standard panel data idea: by observing individual cumulative production histories, we get a *WITHIN-CHILD TIME SERIES*
   - *e.g.*, Specify Cobb-Douglas production technology

$$\tau(q;\theta_E) = \delta_1 \left(\frac{q}{\theta_E}\right)^{\delta_2} \times u$$

     and estimate θ_E as a child-specific fixed effect

2. θ_L and $c(t)$
   - Conditional on θ_E, *EXPERIMENTAL INCENTIVES VARIATION* serves as an instrument
   - Can non-parametrically separate idiosyncratic costs (θ_L) from common baseline costs (c)

OUTLINE: STRUCTURAL IDENTIFICATION OF A 2-DIMENSIONAL PRINCIPAL-AGENT MODEL

1. One-dimensional identification based on
   - D'Haultfeouille and Fevrier (DF2015)
   - Torgovitsky (T2015)
2. Extension: 2-D identification with distributions of outputs AND inputs

ONE-DIMENSIONAL MODEL

1-D PRINCIPAL-AGENT MODEL

(motivation heterogeneity only)

- **TYPES:** $\theta \in \mathbb{R}_{++}, \;\; \theta \sim F_j(\theta), \; j = \mathcal{A}, \mathcal{D}, \mathcal{K}$
- **COSTS:**
  $C(q;\theta) = \theta c(q), \; c'(q) > 0, \; c''(q) > 0 \quad Q \sim G_j(q), \; j = 1, 2, \ldots, J$
- **LABOR CONTRACTS:** $(Q, P(Q))$
  - $P_j(Q) = b_j + w_j Q, \; j = 1, 2, \ldots, J$
  - Exogenous contract assignment
  - Linearity is for simplicity; not crucial here...

1-D MODEL

A agent's objective function is

$$\max_{q \in \mathbb{R}_+} \left\{ \pi_j(\theta) = P_j(q) - \theta c(q) \right\},$$

FOCs defining optimal production choice:

$$\theta_j(q) = \frac{\left(P_j\right)'(q)}{c'(q)} \quad \text{(monotone decreasing)} \tag{1}$$

Also, let $Q_j(\theta) \equiv \theta_j^{-1}$

STRUCTURAL PRIMITIVES: $F_j(\theta)$, and $c(q)$
OBSERVABLES: $P_j^r(q)$, $G_j(q)$

1-D IDENTIFICATION *à la* DF2015/T2015

EXCLUSION RESTRICTIONS:

1. $c_j(q) = c(q), \ \forall j = 1, \ldots, J.$
2. $F_j(\theta) = F(\theta), \ j = 1, \ldots, J.$

NORMALIZATION:
for some $(\theta^*, q^*) \in [\underline{\theta}, \overline{\theta}] \times \mathbb{R}_+$ let $\theta_1(q^*) = \theta^*$.

This is innocuous because

$$C(q; \theta) = \theta c(q) \quad \Leftrightarrow \quad \widetilde{C}(q; \theta) = (\iota\theta)\,(c(q)/\iota)\,, \ \forall \iota \in \mathbb{R}_+$$

- *I.e.*, normalization simply fixes total cost units
- θ^* measured relative to baseline θ^* type.

1-D IDENTIFICATION *à la* DF2015/T2015

“HORIZONTAL TRANSFORM” OPERATOR:

$$\begin{aligned} F(\theta) &= 1 - G_1(Q_1(\theta)) = 1 - G_2(Q_2(\theta)) \\ \Rightarrow Q_1(\theta) &= (G_1)^{-1}\left[(G_2)\left(Q_2(\theta)\right)\right] \equiv H_{1,2}\left[Q_2(\theta)\right]. \end{aligned} \tag{2}$$

$\Rightarrow$ Tell me (θ, q) under contract 1 and I can figure out q' for that same θ under contract 2.

1-D IDENTIFICATION *à la* DF2015/T2015

"HORIZONTAL TRANSFORM" OPERATOR:

$$F(\theta) = 1 - G_1(Q_1(\theta)) = 1 - G_2(Q_2(\theta))$$
$$\Rightarrow Q_1(\theta) = (G_1)^{-1}\left[(G_2)(Q_2(\theta))\right] \equiv H_{1,2}\left[Q_2(\theta)\right]. \tag{2}$$

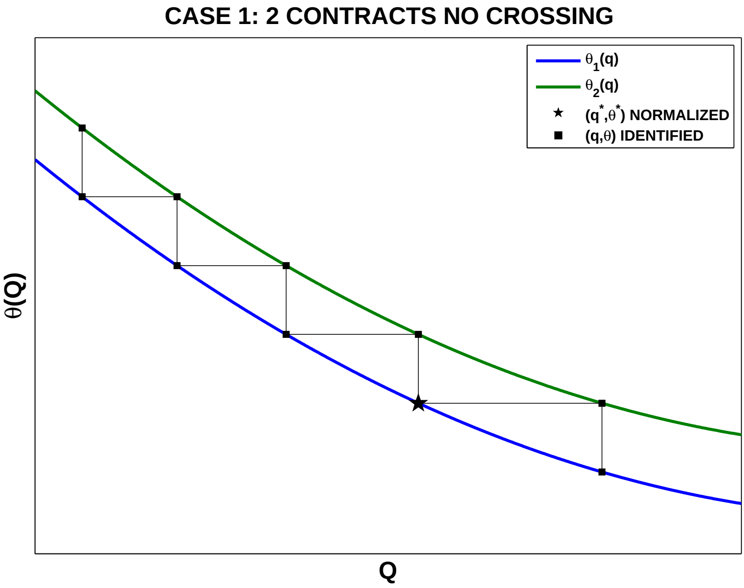
$\Rightarrow$ Tell me (θ, q) under contract 1 and I can figure out q' for that same θ under contract 2.

"VERTICAL TRANSFORM" OPERATOR:

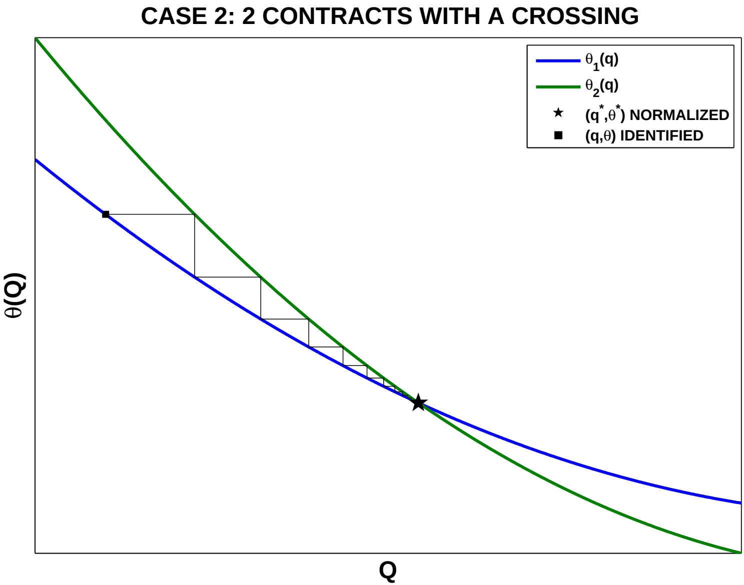
$$\theta_j(q) = \left(P_j\right)'(q)/c'(q)$$
$$\theta_1(q) = \frac{(P_1)'(q)}{(P_2)'(q)}\theta_2(q) \equiv V_{1,2}\left[\theta_2(q)\right] \tag{3}$$

$\Rightarrow$ Tell me (θ, q) under contract 1 and I can figure out θ' for that same q under contract 2.

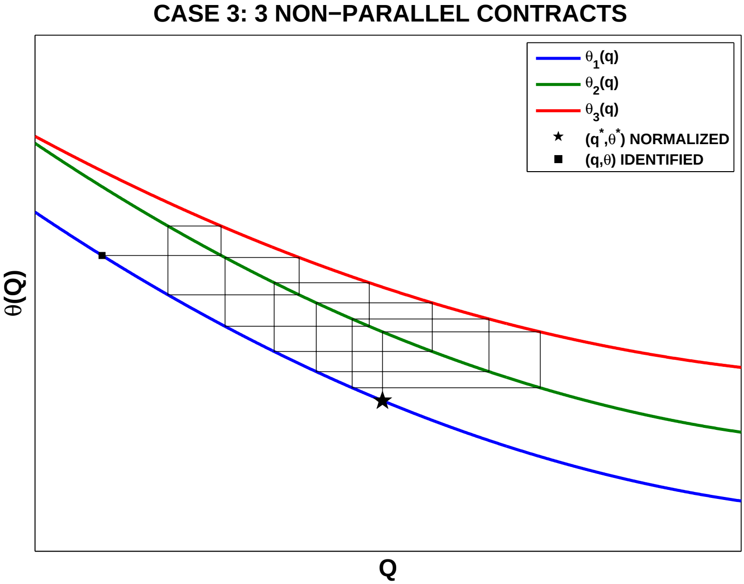
1-D IDENTIFICATION *à la* DF2015/T2015



1-D IDENTIFICATION *à la* DF2015/T2015



1-D IDENTIFICATION *à la* DF2015/T2015



TWO-DIMENSIONAL MODEL EXTENSION

- **TYPES:** $\theta = (\theta_E, \theta_L) \in \mathbb{R}^2_{++}$
- **PRODUCTION TECHNOLOGY:**
  - Let $\tau(q; \theta_E) = t$ denote qty of time input t required to achieve output level q, given type θ_E
- **PRODUCTION COSTS:** $C(t; \theta_L) = \theta_L c(t)$
  - Costs are convex in q: $\frac{d^2 c[\tau(q;\theta_e)]}{dq^2} > 0$
  - **Identifying assumptions:** separability and convexity
- **STRUCTURAL PRIMITIVES:** $\tau(q; \theta_E)$, $c(t)$, and $F_{\theta_E,\theta_L}(\theta_E, \theta_L)$
- **CONTRACTS USED:**
  - $P_k(Q) = b_k + w_k Q,\ j = 1, \ldots, J$

IDENTIFICATION OF θ_E : Cobb-Douglas Production

$$\tau(q;\theta_E) = \left(\frac{q}{\theta_E}\right)^{\delta_2} \times u \quad (*)$$

OBSERVABLES: for each agent i w/positive payout (*i.e.*, total output $q_i \geq 2$ units), and for each individual unit produced p_{ik}, $k = 1, \dots, q_i$,

- $\tau_{p_{ik}} \equiv$ total observed time to achieve p_{ik} units of output

$$(*) \Rightarrow \log(\tau_{p_{ik}}) = \delta_2 \log(p_{ik}) - \delta_2 \log(\theta_{Ei}) + \varepsilon_{p_{ik}}, \quad (**)$$

- Assume $\mathrm{E}\left[[1\ p_{ik}\ \theta_{Ei}]^\top \varepsilon_{p_{ik}}\right] = \mathbf{0}$ across i and k

IDENTIFICATION OF θ_E, τ: Cobb-Douglas Production

IDENTIFICATION OF $(\delta_2, \{\theta_{Ei}\}_{i=1}^{I})$:

- **STEP 1:** For each i and $p_{ik} = 2, \ldots, q_i$ we difference to get

$$(**) \Rightarrow \log\left(\frac{\tau_{p_{ik}}}{\tau_{p_{i,k-1}}}\right) = \delta_2 \log\left(\frac{p_{ik}}{p_{i,k-1}}\right) + (\varepsilon_{p_{ik}} - \varepsilon_{pi,k-1})$$

  which identifies δ_2.

IDENTIFICATION OF θ_E, τ: Cobb-Douglas Production

IDENTIFICATION OF $\left(\delta_2, \{\theta_{Ei}\}_{i=1}^{I}\right)$**:**

- **STEP 1:** For each i and $p_{ik} = 2, \ldots, q_i$ we difference to get

$$(**) \Rightarrow \log\left(\frac{\tau_{p_{ik}}}{\tau_{p_{i,k-1}}}\right) = \delta_2 \log\left(\frac{p_{ik}}{p_{i,k-1}}\right) + (\varepsilon_{p_{ik}} - \varepsilon_{pi,k-1})$$

  which identifies δ_2.

- **STEP 2:** For each i and $p_{ik} = 1, \ldots, q_i \geq 2$, we have

$$\log(\tau_{p_{ik}}) - \delta_2 \log(p_{ik}) = \gamma_i D_i + v_{p_{ik}},$$

  where D_i is a dummy for student i.

IDENTIFICATION OF θ_E, τ: Cobb-Douglas Production

IDENTIFICATION OF $(\delta_2, \{\theta_{Ei}\}_{i=1}^{I})$**:**

- **STEP 1:** For each i and $p_{ik} = 2, \ldots, q_i$ we difference to get

$$(**) \Rightarrow \log\left(\frac{\tau_{p_{ik}}}{\tau_{p_{i,k-1}}}\right) = \delta_2 \log\left(\frac{p_{ik}}{p_{i,k-1}}\right) + (\varepsilon_{p_{ik}} - \varepsilon_{pi,k-1})$$

  which identifies δ_2.

- **STEP 2:** For each i and $p_{ik} = 1, \ldots, q_i \geq 2$, we have

$$\log(\tau_{p_{ik}}) - \delta_2 \log(p_{ik}) = \gamma_i D_i + v_{p_{ik}},$$

  where D_i is a dummy for student i.

- **STEP 3:** $(**)$ implies that

$$\theta_{Ei} = \exp\left(-\frac{\gamma_i}{\delta_2}\right)$$

IDENTIFICATION OF θ_L, c:

Agent's Obj. Fn.: $\max_{q \in \mathbb{R}_+} \left\{ P_j(q) - \theta_L c \left[\tau(q;\theta_E) \right] \right\}, \; j = 1,2,3$

$$\text{FOCs:} \quad \theta_{Lj}(q;\theta_E) = \frac{\left(P_j\right)'(q)}{c'\left[\tau(q;\theta_E)\right]\tau'(q;\theta_E)}. \tag{4}$$

IDENTIFICATION OF θ_L, c:

Agent's Obj. Fn.: $\max_{q\in\mathbb{R}_+}\left\{P_j(q)-\theta_L c\left[\tau(q;\theta_E)\right]\right\},\ j=1,2,3$

$$\text{FOCs:}\quad \theta_{Lj}(q;\theta_E)=\frac{\left(P_j\right)'(q)}{c'\left[\tau(q;\theta_E)\right]\tau'(q;\theta_E)}. \tag{4}$$

EXCLUSION RESTRICTIONS, NORMALIZATION:

1. $c_j(t)=c(t),\ j=1,2,3.$
2. $\tau_j(q;\theta_E)=\tau(q;\theta_E),\ j=1,2,3.$
3. $F_j(\theta_L|\theta_E)=F(\theta_L|\theta_E),\ j=1,2,3.$

   (*i.e.*, randomization implies no selective entry, and work tasks are the same across incentive groups)
4. For some $(\theta_L^*,\theta_E^*,q^*)\in[\underline{\theta}_L,\overline{\theta}_L]\times[\underline{\theta}_E,\overline{\theta}_E]\times\mathbb{R}_+$ we fix

$$\theta_{L1}(q^*;\theta_E^*)=\theta_L^*.$$

IDENTIFICATION OF θ_L, c:

HORIZONTAL TRANSFORM: Recall that

$$F(\theta_L|\theta_E) = 1 - G_j\left[Q_j(\theta_L;\theta_E)\middle|\theta_E\right]$$
$$= 1 - G_{j'}\left[Q_{j'}(\theta_L;\theta_E)\middle|\theta_E\right],\ j \neq j'$$

IDENTIFICATION OF θ_L, c:

HORIZONTAL TRANSFORM: Recall that

$$
\begin{aligned}
F(\theta_L|\theta_E) &= 1 - G_j\left[Q_j(\theta_L;\theta_E)\middle|\theta_E\right] \\
&= 1 - G_{j'}\left[Q_{j'}(\theta_L;\theta_E)\middle|\theta_E\right],\ j \neq j'
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\Rightarrow Q_j(\theta_L;\theta_E) &= \left(G_j\right)^{-1}\left[\left(G_{j'}\right)^{-1}\left(Q_{j'}[\theta_L;\theta_E]\middle|\theta_E\right)\middle|\theta_E\right] \\
&\equiv H_{j,j'}\left[Q_{j'}(\theta_L;\theta_E)\middle|\theta_E\right],\ j = 1,2,3,\ j \neq j'
\end{aligned}
\tag{5}
$$

IDENTIFICATION OF θ_L, c:

VERTICAL TRANSFORM: Recall that

$$\theta_{Lj}(q;\theta_E) = \frac{\left(P_j\right)'(q)}{c'\left[\tau(q;\theta_E)\right]\tau'(q;\theta_E)}, \; j = 1,2,3$$

IDENTIFICATION OF θ_L, c:

VERTICAL TRANSFORM: Recall that

$$\theta_{Lj}(q;\theta_E) = \frac{\left(P_j\right)'(q)}{c'\left[\tau(q;\theta_E)\right]\tau'(q;\theta_E)},\ j = 1,2,3$$

Therefore,

$$\begin{aligned}\Rightarrow \theta_{Lj}(q;\theta_E) &= \frac{\left(P_j\right)'(q)}{\left(P_{j'}\right)'(q)}\theta_{Lj'}(q;\theta_E) \\ &\equiv V_{j,j'}\left[\theta_{Lj'}(q;\theta_E)\middle|\theta_E\right],\ j = 1,2,3,\ j \neq j'.\end{aligned} \tag{6}$$

IDENTIFICATION OF θ_L, c:

IDENTIFICATION ARGUMENT (assuming no mass points in G_j):

1. **STEP 1:** $\theta_E = \frac{q}{t}$ directly identifies the joint distribution

$$F_{\theta_E,Q,j}(\theta_E, q) \quad \Rightarrow G_j(Q|\theta_E)$$

IDENTIFICATION OF θ_L, c:

IDENTIFICATION ARGUMENT (assuming no mass points in G_j):

1. **STEP 1:** $\theta_E = \frac{q}{t}$ directly identifies the joint distribution

$$F_{\theta_E,Q,j}(\theta_E, q) \quad \Rightarrow G_j(Q|\theta_E)$$

2. **STEP 2:** given $(\theta_L^*, \theta_E^*, q^*)$, we know that $G_j(q|\theta_E^*)$, HT, VT identify $\theta_{Lj}(q;\theta_E^*)$, $j = 1, 2, 3$ for $q \geq 2$

IDENTIFICATION OF θ_L, c:

<u>IDENTIFICATION ARGUMENT</u> (assuming no mass points in G_j):

1. **STEP 1:** $\theta_E = \frac{q}{t}$ directly identifies the joint distribution

$$F_{\theta_E,Q,j}(\theta_E, q) \quad \Rightarrow G_j(Q|\theta_E)$$

2. **STEP 2:** given $(\theta_L^*, \theta_E^*, q^*)$, we know that $G_j(q|\theta_E^*)$, HT, VT identify $\theta_{Lj}(q; \theta_E^*)$, $j = 1, 2, 3$ for $q \geq 2$

3. **STEP 3:** Plugging $\theta_{L1}(\cdot; \theta_E^*)$ into the FOC identifies $c(t)$

IDENTIFICATION OF θ_L, c:

IDENTIFICATION ARGUMENT (assuming no mass points in G_j):

1. **STEP 1:** $\theta_E = \frac{q}{t}$ directly identifies the joint distribution

$$F_{\theta_E,Q,j}(\theta_E, q) \quad \Rightarrow G_j(Q|\theta_E)$$

2. **STEP 2:** given $(\theta_L^*, \theta_E^*, q^*)$, we know that $G_j(q|\theta_E^*)$, HT, VT identify $\theta_{Lj}(q;\theta_E^*)$, $j = 1, 2, 3$ for $q \geq 2$

3. **STEP 3:** Plugging $\theta_{L1}(\cdot;\theta_E^*)$ into the FOC identifies $c(t)$

4. **STEP 4:** Knowing $c(\cdot)$ identifies $\theta_{L2}(q;\theta_E)$, $\theta_{L3}(q;\theta_E)$ for arbitrary (θ_E) using the FOCs

IDENTIFICATION OF θ_L, c:

IDENTIFICATION ARGUMENT (assuming no mass points in G_j):

1. **STEP 1:** $\theta_E = \frac{q}{t}$ directly identifies the joint distribution
$$F_{\theta_E,Q,j}(\theta_E, q) \quad \Rightarrow G_j(Q|\theta_E)$$
2. **STEP 2:** given $(\theta_L^*, \theta_E^*, q^*)$, we know that $G_j(q|\theta_E^*)$, HT, VT identify $\theta_{Lj}(q;\theta_E^*)$, $j = 1,2,3$ for $q \geq 2$
3. **STEP 3:** Plugging $\theta_{L1}(\cdot;\theta_E^*)$ into the FOC identifies $c(t)$
4. **STEP 4:** Knowing $c(\cdot)$ identifies $\theta_{L2}(q;\theta_E)$, $\theta_{L3}(q;\theta_E)$ for arbitrary (θ_E) using the FOCs
5. **STEP 5:** Inverting q for a given θ_E, then $\{\theta_E, \theta_L\}$ the joint distribution $F_{\theta_E,\theta_L}(\theta_E, \theta_L)$ is nonparametrically identified.

APPLICATION: "*Motivation Vs Productivity: A Study of HC Investment Using Field Experiments and Structural Modeling*" Cotton, Hickman, List, Price (2015)

MOTIVATION: Measures of childhood human capital (*e.g.*, SAT scores, educational attainment) are known to be linked to important outcomes:

- College attendance/GPA/Graduation Probability
- Employment, Productivity, Income
- Even Marital/Family Stability
- Early development also drives socioeconomic inequality

Raw Ability vs. Work Ethic

Consider Suzie, a 7th grader who scores highly on a math exam. We ask WHY?

TWO COMPETING EXPLANATIONS (both hinge on Suzie's unobservable characteristics):

1. **LEARNING PRODUCTIVITY:** relatively small amounts of Suzie's study time translate into big gains in her math understanding...

2. **MOTIVATION:** Suzie loves learning math or doesn't care as much about leisure activities, so she chooses to spend more time studying than others...

Raw Ability vs. Work Ethic (cont'd)

POLICY IMPLICATIONS:

Knowing more about Suzie's characteristics and time choices can inform us on how best to intervene if she struggles later on.

1. If Suzie is highly motivated but slow to learn,
   - Resource problem? (*e.g.*, tutoring)
   - Study Skills deficiency? (*e.g.*, help her to use time more efficiently)
2. If Suzie is a fast learner but lacking motivation,
   - Inspiration problem? (*e.g.*, feeling unengaged by her coursework)
   - Incentives problem?
3. At the group level, we could decompose gender/race gaps into differences in unobservables...
   - Mainly a productivity difference?
   - Mainly a motivational difference?
   - Some mixture?

Research Design:

- Experimental variation in short-term time investment incentives.
- Observations of time investment into learning in a controlled, but natural field setting

Research Design:

- Experimental variation in short-term time investment incentives.
- Observations of time investment into learning in a controlled, but natural field setting
    - Online Mathematics portal w/ 5th-8th grade students & Common Core curriculum

Research Design:

- Experimental variation in short-term time investment incentives.
- Observations of time investment into learning in a controlled, but natural field setting
  - Online Mathematics portal w/ 5th-8th grade students & Common Core curriculum
  - 3 randomly assigned wage contracts for completing math learning tasks on website

Research Design:

- Experimental variation in short-term time investment incentives.
- Observations of time investment into learning in a controlled, but natural field setting
  - Online Mathematics portal w/ 5th-8th grade students & Common Core curriculum
  - 3 randomly assigned wage contracts for completing math learning tasks on website
  - **Automated Monitoring of time use decisions**

Research Design:

- Experimental variation in short-term time investment incentives.
- Observations of time investment into learning in a controlled, but natural field setting
  - Online Mathematics portal w/ 5th-8th grade students & Common Core curriculum
  - 3 randomly assigned wage contracts for completing math learning tasks on website
  - **Automated Monitoring of time use decisions**
  - Progress in math proficiency measured in classroom, working with students' regular teachers

CHICAGO HUMAN CAPITAL STUDY

- Mathematics study held at public schools in Chicago suburbs
  - VERY heterogeneous population, racially, culturally, socioeconomically

CHICAGO HUMAN CAPITAL STUDY

- Mathematics study held at public schools in Chicago suburbs
  - VERY heterogeneous population, racially, culturally, socioeconomically
- Adjacent grades combined for the study (5/6, 7/8)
  - 5/6=GROUP 2
  - 7/8=GROUP 3

WWW.CHICAGOMATHGAME.ORG

- MC questions taken from state standardized testing materials over the past decade
  - CA, FL, IL, MN, NY, New England, TX, WI, OH, NAEP (National)
  - All questions similar to curriculum in use at subjects' schools
  - All questions separated into 5 subject categories, based on Common Core
  - Questions arranged into "quizzes"
    - ★ 6 questions each
    - ★ 5 subject-specific quizzes per category
    - ★ 55 general quizzes
    - ★ Passing criterion: at least 5/6 correct
    - ★ Practice run for missed problems
    - ★ Random ordering each attempt
- Supplementary Instructional Materials (glossary, web videos, practice problems w/ solutions)

Experimental Design (grades 5-8):

- Pre-test from same pool of questions

Experimental Design (grades 5-8):

- Pre-test from same pool of questions
- Direct wage contracts for completing quizzes: $(Q, P(Q))$, where

$$P_j(Q) = \left[b_j + m_j Q\right] \cdot \mathbb{1}\{Q \geq 2\},\ j = 1, 2, 3$$

  - **CONTRACT 1:** $(b_1, m_1) = (\$15, \$0.75)$
  - **CONTRACT 2:** $(b_2, m_2) = (\$10, \$1)$
  - **CONTRACT 3:** $(b_3, m_3) = (\$5, \$1.25)$

Experimental Design (grades 5-8):

- Pre-test from same pool of questions
- Direct wage contracts for completing quizzes: $(Q, P(Q))$, where

$$P_j(Q) = \left[b_j + m_j Q\right] \cdot \mathbb{1}\{Q \geq 2\},\ j = 1,2,3$$

  - **CONTRACT 1:** $(b_1, m_1) = (\$15, \$0.75)$
  - **CONTRACT 2:** $(b_2, m_2) = (\$10, \$1)$
  - **CONTRACT 3:** $(b_3, m_3) = (\$5, \$1.25)$
- Individual randomassignment w/in each school/grade

Experimental Design (grades 5-8):

- Pre-test from same pool of questions
- Direct wage contracts for completing quizzes: $(Q, P(Q))$, where

$$P_j(Q) = \left[b_j + m_j Q\right] \cdot \mathbb{1}\{Q \geq 2\},\ j = 1,2,3$$

  - **CONTRACT 1:** $(b_1, m_1) = (\$15, \$0.75)$
  - **CONTRACT 2:** $(b_2, m_2) = (\$10, \$1)$
  - **CONTRACT 3:** $(b_3, m_3) = (\$5, \$1.25)$
- Individual randomassignment w/in each school/grade
- Students informed of rules on an information sheet
  - Their own score from practice test
  - Quiz passing criterion
  - Wage contract/total earnings potential
  - 10-day work period + login for website
- Post-test (with 1/4 overlap from pre)

Experimental Design:

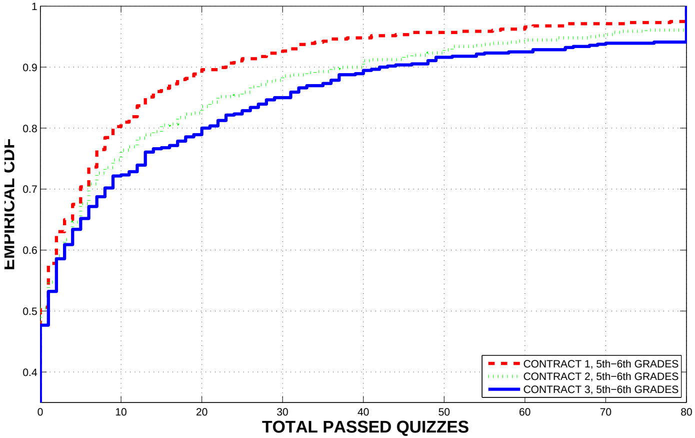
OTHER OBSERVABLES:

- Student Pre/Post Survey
  - Extracurricular/leisure activites, gaming consoles
  - Math preferences
  - Intrinsic vs Extrinsic Motivation
- Student note taking on Test booklet
- Parent Post Survey
  - Family Structure, housing, income
  - Education background for both parents, STEM ability
  - Parenting Styles, time use
  - Prize Preferences
- Home values (County tax records)

Preliminary Descriptive Results

- WEBSITE USAGE:
  - **1565 students (48%)** used website at least once
  - For site users, avg time spent was **2 hours**
  - **Over 16 hours** in come cases
- QUIZ OUTPUT:
  - Avg # of quizzes passed was **21.5 out of 80** possible
  - Avg # of quiz attempts was **39**

Preliminary Descriptive Results



Preliminary Descriptive Results

- **“Struggling Workhorse” (SW):**
  - Bottom half on pre-test, but top half (w/in incentive group) for minutes spent on website
  - **434 students out of 2,380** (18.2% overall)
- **“Struggling SuperWorkhorse” (SSW):**
  - Bottom half/top quarter OR bottom quarter/top half
  - **287 students out of 2,380** (12.1% overall)
- **“Low Motivation” (LM):**
  - Top half/bottom half
  - **460 students out of 2,380** (19.3% overall)

ESTIMATION

STAGE 1 (PRODUCTION TECH): Estimate $\widehat{\delta}_2, \{\widehat{\theta}_{Ei}\}_{i=1}^{I}$ using standard methods for (unbalanced) panel data.

ESTIMATION

STAGE 1 (PRODUCTION TECH): Estimate $\widehat{\delta}_2, \{\widehat{\theta}_{Ei}\}_{i=1}^{I}$ using standard methods for (unbalanced) panel data.

STAGE 2: Now, how to estimate $c(t)$, $\{\widehat{\theta}_{Li}\}_{i=1}^{I}$, and $F_{\theta_E,\theta_L}(\theta_E, \theta_L)$?

- **METHOD 1:** Nonparametric Kernel-based estimator
  - Partial Identification Problem: 50% of students choose zero output
  - Can only bound their 2-D type within a region of (θ_E, θ_L) space
  - Also cannot include covariates in model since high-dimensional nonparametrics are infeasible

- **METHOD 2:** Spline-based semi-parametric estimator
  - Flexible B-spline specification for $c(t)$
  - Joint normality assumption to extrapolate non-identified region of (θ_E, θ_L) space
  - Also allows for inclusion of covariates...

ESTIMATION METHOD 1: Kernel Smoothing

Given STAGE 1, we have (q_{ij}, θ_{Eij}) for each student i in contract j

1. Within each contract sample, estimate the conditional CDF

$$\widehat{G}^{j}_{Q|\theta_E}(q|\theta_E) = \sum_{i=1}^{I} \mathbb{1}(q_{ij} \leq q)\kappa\left(\frac{\theta_E - \theta_{Eij}}{h_j}\right),$$

where κ is a density function and h_j is a bandwidth.

ESTIMATION METHOD 1: Kernel Smoothing

Given STAGE 1, we have (q_{ij}, θ_{Eij}) for each student i in contract j

1. Within each contract sample, estimate the conditional CDF

$$\widehat{G}^{j}_{Q|\theta_E}(q|\theta_E) = \sum_{i=1}^{I} \mathbb{1}(q_{ij} \leq q)\kappa\left(\frac{\theta_E - \theta_{Eij}}{h_j}\right),$$

   where κ is a density function and h_j is a bandwidth.

2. Use the HT and VT operators to estimate $\widehat{\theta}_L(q;\theta^*_E)$

ESTIMATION METHOD 1: Kernel Smoothing

Given STAGE 1, we have (q_{ij}, θ_{Eij}) for each student i in contract j

1. Within each contract sample, estimate the conditional CDF

$$\widehat{G}^{j}_{Q|\theta_E}(q|\theta_E) = \sum_{i=1}^{I} \mathbb{1}(q_{ij} \leq q)\kappa\left(\frac{\theta_E - \theta_{Eij}}{h_j}\right),$$

   where κ is a density function and h_j is a bandwidth.

2. Use the HT and VT operators to estimate $\widehat{\theta}_L(q;\theta_E^*)$

3. Plug $\widehat{\theta}_L(q;\theta_E^*)$ into FOC to get $\widehat{c}(t)$, which applies to all θ_E

ESTIMATION METHOD 1: Kernel Smoothing

Given STAGE 1, we have (q_{ij}, θ_{Eij}) for each student i in contract j

1. Within each contract sample, estimate the conditional CDF

$$\widehat{G}^{j}_{Q|\theta_E}(q|\theta_E) = \sum_{i=1}^{I} \mathbb{1}(q_{ij} \leq q)\kappa\left(\frac{\theta_E - \theta_{Eij}}{h_j}\right),$$

   where κ is a density function and h_j is a bandwidth.

2. Use the HT and VT operators to estimate $\widehat{\theta}_L(q; \theta_E^*)$

3. Plug $\widehat{\theta}_L(q; \theta_E^*)$ into FOC to get $\widehat{c}(t)$, which applies to all θ_E

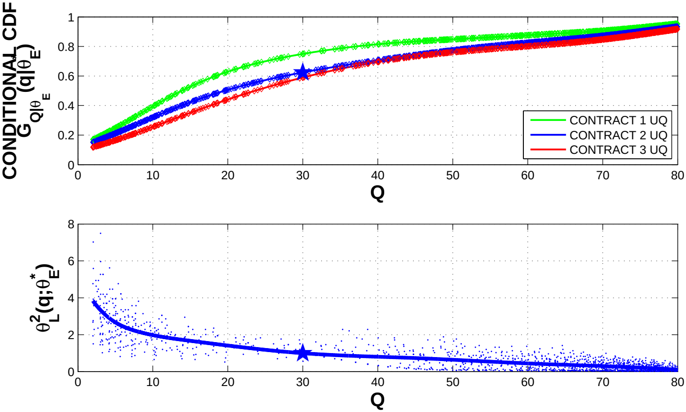
4. Use FOCs and $\widehat{c}(t)$ to recover $\widehat{\theta}_{Lij}(q_{ij}; \widehat{\theta}_{Eij})$ for each student in the data with $2 \leq q_{ij} < 80$

ESTIMATION METHOD 1: Kernel Smoothing

PROS: Simple, method follows directly from identification argument

CONS: Statistically inefficient for two reasons:

1. Kernel smoothing in multiple dimensions very data hungry
2. Method does not directly impose theory on CDF estimates



ESTIMATION METHOD 2: Semiparametric GMM

First, parameterize

$$\widehat{c}(t;\boldsymbol{\alpha}_c) = \sum_{m=1}^{M} \alpha_{cm} C_m(t)$$

where the $C_m(\cdot)$'s are a set of B-spline basis functions.

- Knots in t space chosen as uniform quantiles

ESTIMATION METHOD 2: Semiparametric GMM

First, parameterize

$$\widehat{c}(t;\boldsymbol{\alpha}_c) = \sum_{m=1}^{M} \alpha_{cm} C_m(t)$$

where the $C_m(\cdot)$'s are a set of B-spline basis functions.

- Knots in t space chosen as uniform quantiles

Next, assume joint log-normality of types

$$(\log\theta_E, \log\theta_L) = \mathcal{BVN}([\mu_E, \mu_L], \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma} = [\mu_e, \mu_l, \sigma_e^2, \sigma_l^2, \sigma_{el}]$

Parametric form characterizes **selection equations**:

$$\Pr[zero\ output] = \Pr\left[b_j + 2m_j < \theta_L \widehat{c}\left(\widehat{\tau}(2/\theta_E); \boldsymbol{\alpha}_c\right); \boldsymbol{\alpha}_{el}\right]$$

$$\Pr[full\ output] = \Pr\left[m_j \geq \theta_L \widehat{c}'\left(\widehat{\tau}(80/\theta_{Eij}); \boldsymbol{\alpha}_c\right) \widehat{\tau}'(80/\theta_{Eij}); \boldsymbol{\alpha}_{el}\right]$$

ESTIMATION METHOD 2: Semiparametric GMM

QUESTION: Why not just estimate via Tobit Maximum Likelihood?

ANSWER: The censoring point (extensive margin for work vs not) is a function of the fixed effects θ_{Ei} and θ_{Li}, which are also parameters to be estimated. This violates regularity conditions needed for MLE to be numerically well-behaved and consistent...

ESTIMATION METHOD 2: Semiparametric GMM

Additional Conditions to impose:

- Orthogonality conditions from production technology
- Normalization: $\theta_L^* = \frac{P_1'(q^*)}{\widehat{c}'\left[\widehat{\tau}(q^*;\theta_E^*);\boldsymbol{\alpha}_c\right]\widehat{\tau}'(q^*;\theta_E^*)} = 1$, $\widehat{G}_2(q^*) = 0.75$
- FOCs: $\widehat{\theta}_{Lij} = \frac{\left(P_j\right)'(q_{ij})}{\widehat{c}'\left[\widehat{\tau}(q_{ij};\widehat{\theta}_{Eij});\boldsymbol{\alpha}_c\right]\widehat{\tau}'(q_{ij};\widehat{\theta}_{Eij})}$, $\forall i, j$,

Shape Restrictions from linear constraints (B-splines=magic!):

- Boundary Condition: $\widehat{c}(0;\boldsymbol{\alpha}_c) = 0 \quad \Rightarrow \alpha_{c1} = 0$
- Monotonicity: $\widehat{c}'(t;\boldsymbol{\alpha}_c) > 0 \quad \Rightarrow \alpha_{ck} < \alpha_{c,k+1}$, $\forall k$
- Convexity:

$$\widehat{c}''\left[\widehat{\tau}(q;\widehat{\theta}_E);\boldsymbol{\alpha}_c\right]\left(\widehat{\tau}'(q;\widehat{\theta}_E)\right)^2 + \widehat{c}'\left[\widehat{\tau}(q;\widehat{\theta}_E);\boldsymbol{\alpha}_c\right]\widehat{\tau}''(q;\widehat{\theta}_E) > 0$$

ESTIMATION METHOD 2: Semiparametric GMM

Additional Model Components:

- Decompose (θ_E, θ_L) as functions of student covariates (socioeconomics, leisure opportunities, home environment, subject preferences, etc...)
  - Specify types as a single index of covariates $\mathbf{X}$:

$$\theta_E = x_{e1}^{\beta_{E1}} \cdot x_{e2}^{\beta_{E2}} \cdots x_{ek}^{\beta_{Ek}} \cdot \eta_E$$

$$\theta_L = x_{l1}^{\beta_{L1}} \cdot x_{l2}^{\beta_{L2}} \cdots x_{lk}^{\beta_{Lk}} \cdot \eta_L$$

    where η_E and η_L are the components of a student's type that are orthogonal to observables.
  - Incorporate information from:
    - ★ Organizational skills/habits using test booklet data
    - ★ Survey data (math attitude, motivation, home environment)
    - ★ Census block group data on poverty/stability measures

ESTIMATION METHOD 2: Semiparametric GMM

Additional Model Components:

- Relate estimates to math pre-test score S and score change ΔS using post-test scores
- Cobb-Douglas production:

$$S_i = A_s \theta_E^{\alpha_{sE}} \theta_L^{\alpha_{sL}} U_{si}$$

$$\Delta S_i = A_\Delta \theta_E^{\alpha_{\Delta E}} \theta_L^{\alpha_{\Delta L}} U_{\Delta i}$$

where we assume $\mathrm{E}\left[[1,\ \theta_{Ei},\ \theta_{Li}]^\top \log U_{ji}\right] = \mathbf{0}$ for $j = s, \Delta$.

- This provides a link between our learning by doing production model and traditional outcome measures of human capital.

ESTIMATION METHOD 2: Semiparametric GMM

Additional Model Components:

- Relate estimates to math pre-test score S and score change ΔS using post-test scores
- Cobb-Douglas production:

$$S_i = A_s \theta_E^{\alpha_{sE}} \theta_L^{\alpha_{sL}} U_{si}$$

$$\Delta S_i = A_\Delta \theta_E^{\alpha_{\Delta E}} \theta_L^{\alpha_{\Delta L}} U_{\Delta i}$$

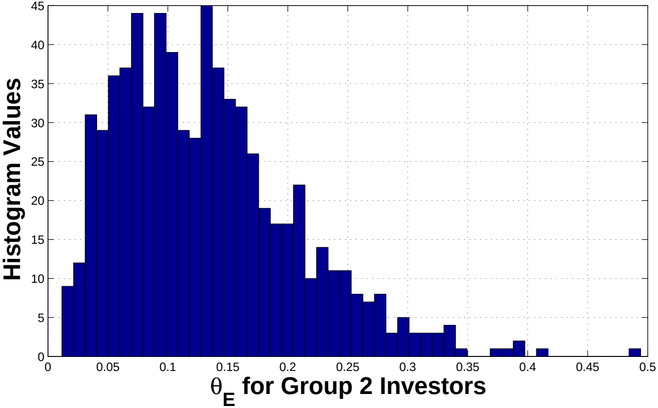
  where we assume $\mathrm{E}\left[[1,\ \theta_{Ei},\ \theta_{Li}]^\top \log U_{ji} \right] = \mathbf{0}$ for $j = s, \Delta$.
- This provides a link between our learning by doing production model and traditional outcome measures of human capital.

IMPORTANT NOTE: involving covariates or test scores in the empirical model requires the joint normality assumption for feasibility and to correct for sample selection (*i.e.*, mass points at zero output).

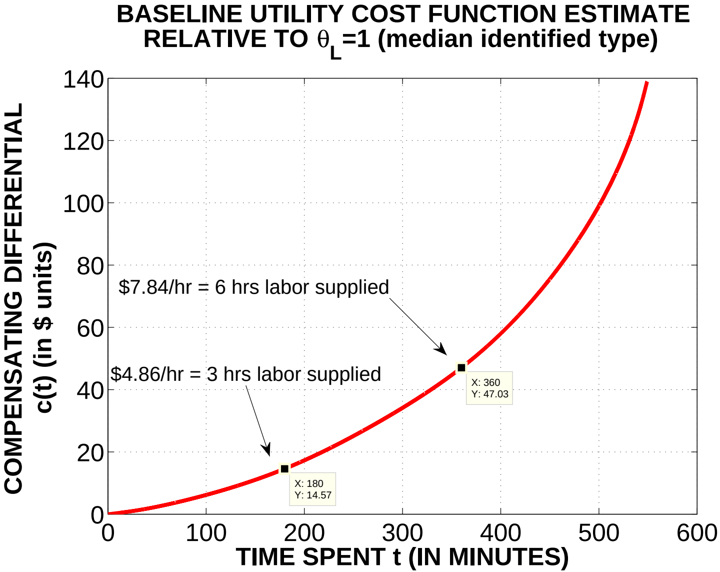
Stage 1 Results:

Production Technology estimate: $\hat{\tau}(q;\theta_e) = \left(\frac{q}{\theta_E}\right)^{\hat{\delta}_2}$

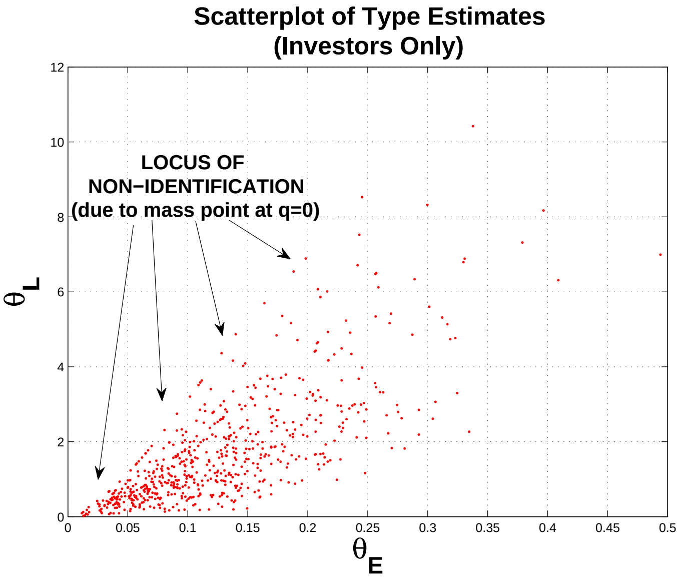
- $\hat{\delta}_2 = 0.9969$ (not statistically different from 1)



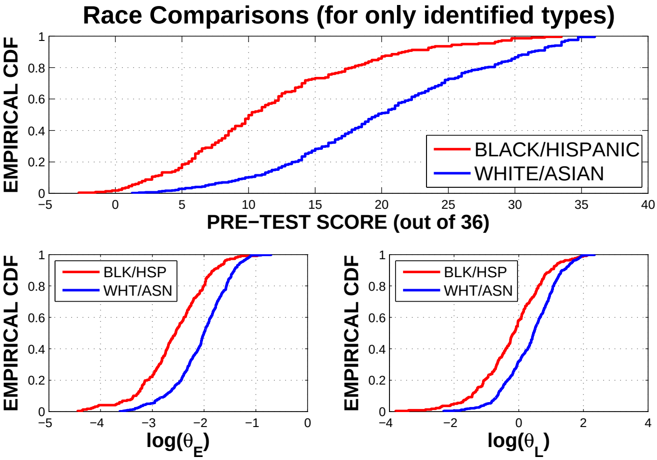
Stage 2 Kernel-Based Results:



Stage 2 Kernel-Based Results:

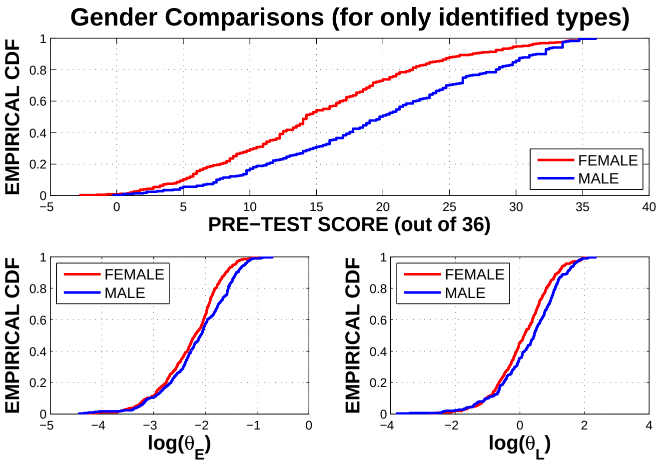


Race Comparison (kernel-based results)



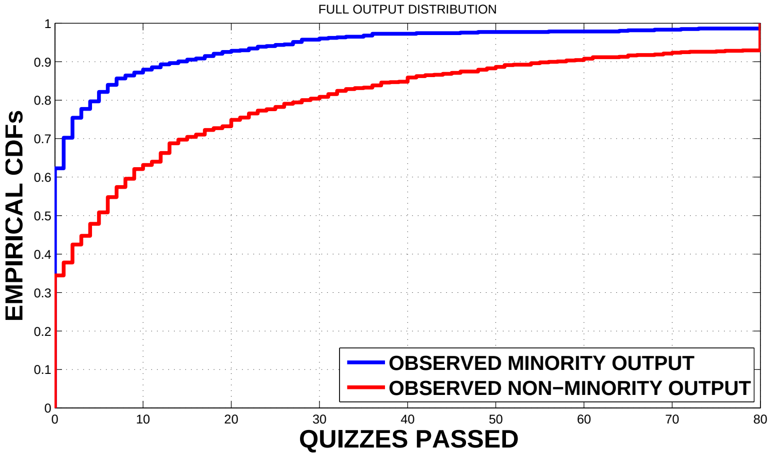
I.e., BLK/HSP students take longer to accomplish learning tasks but are more willing to substitute from leisure time toward work.

Gender Comparison (kernel-based results)

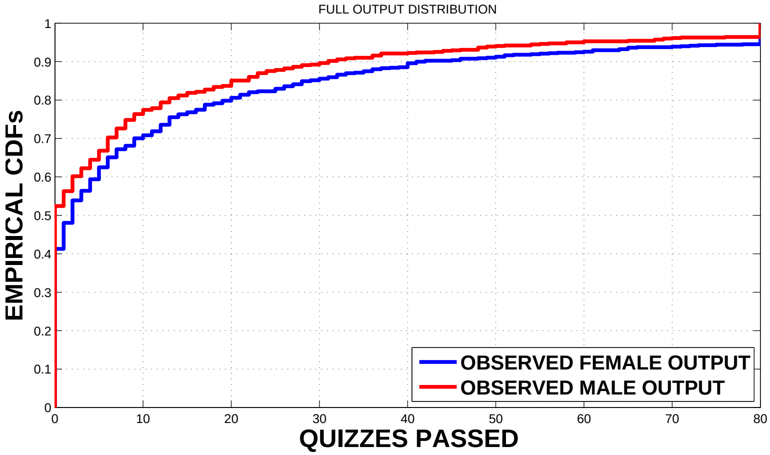


I.e., FEMALE students take longer to accomplish learning tasks but are more willing to substitute from leisure time toward work.

Empirical Output Distributions by Race



Output Distributions by Gender



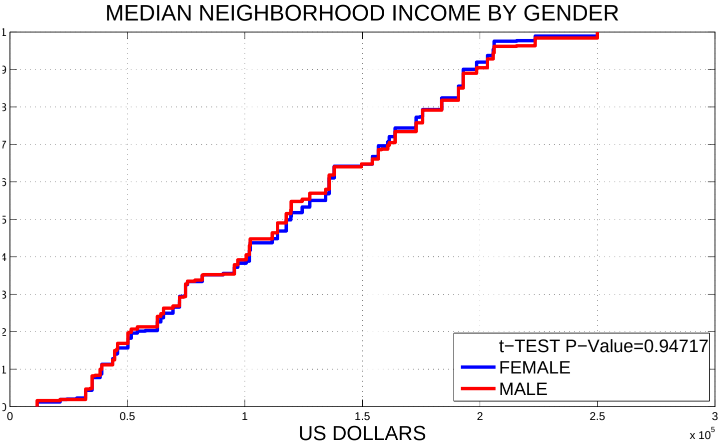
SEMIPARAMETRIC STRUCTURAL ESTIMATES

NEARING COMPLETION AND COMING SOON...

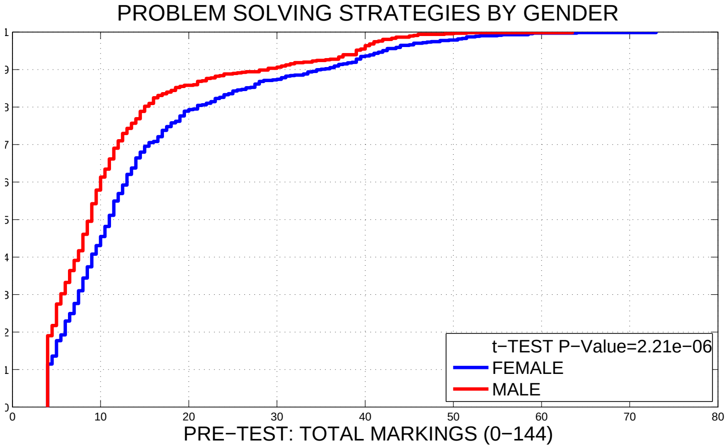
1. Production Counterfactuals
   - Equalize marginal distributions of different characteristics/covariates across demographic groups, one at a time (decomposition)
   - What level of incentives needed to close race/gender gaps in output?

AS A PREVIEW, WE INCLUDE DESCRIPTIVE ANALYSES OF COVARIATES BY RACE AND GENDER HERE:

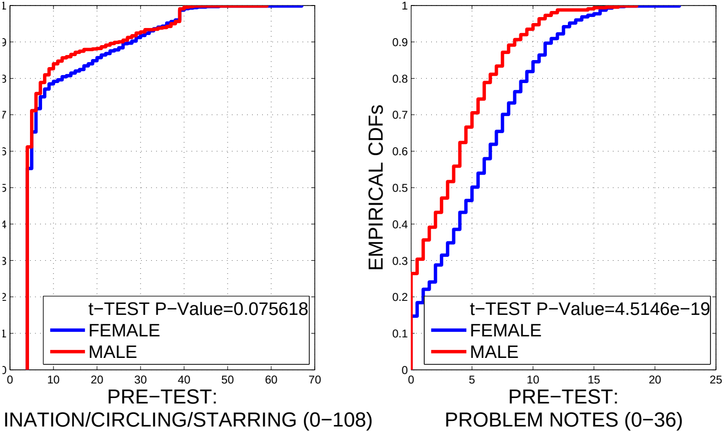
COVARIATES by Gender



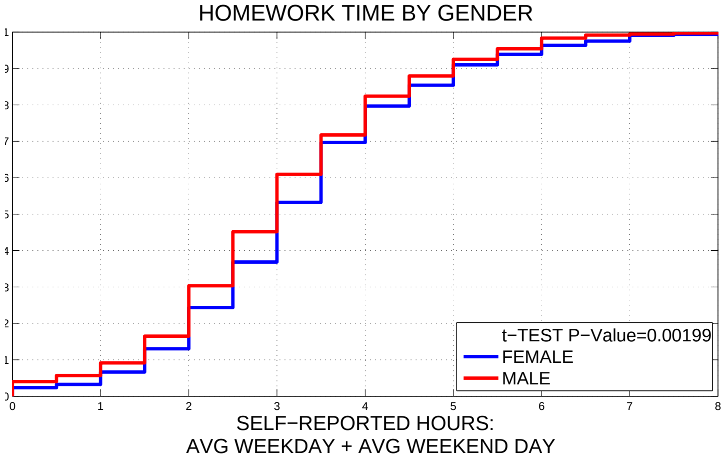
COVARIATES by Gender



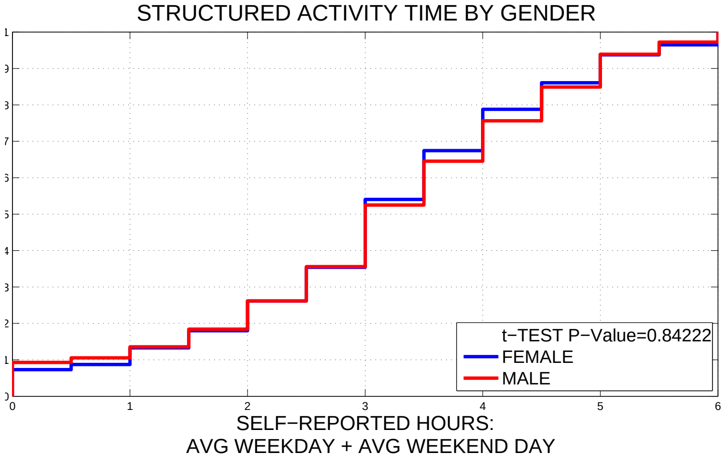
COVARIATES by Gender



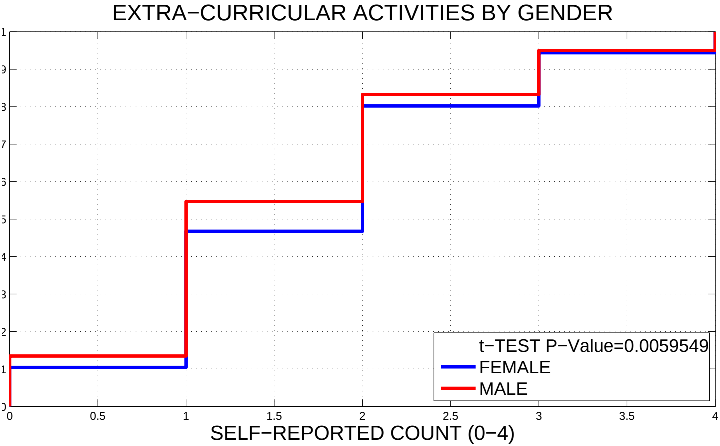
COVARIATES by Gender



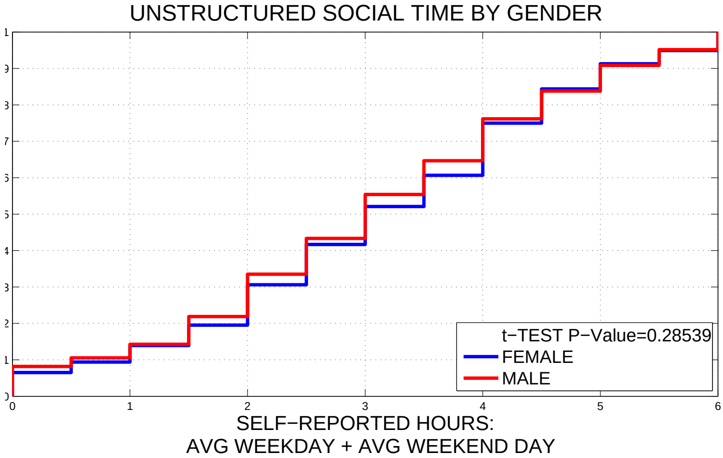
COVARIATES by Gender



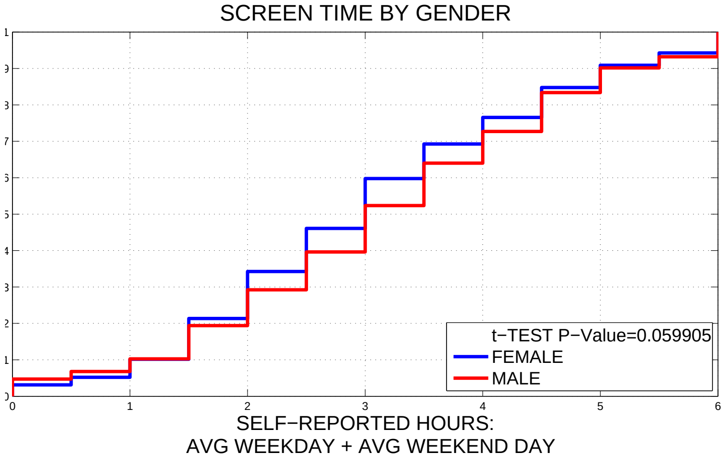
COVARIATES by Gender



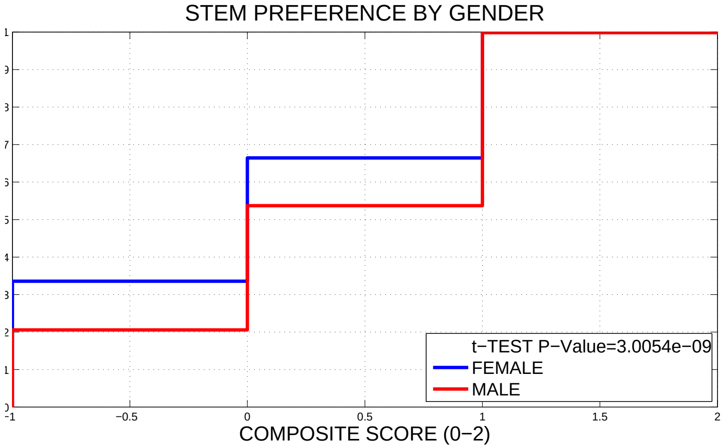
COVARIATES by Gender



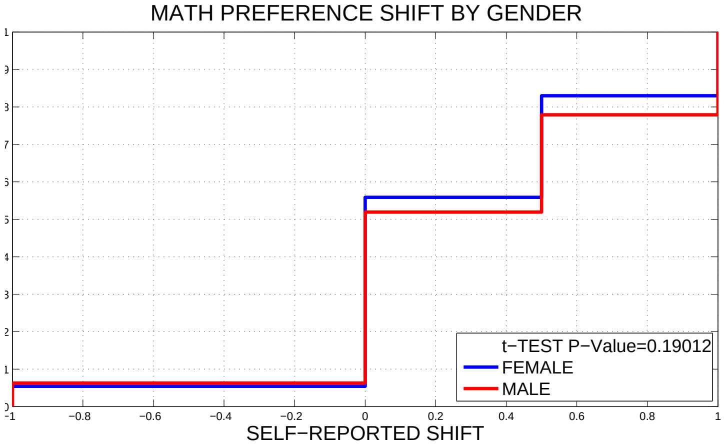
COVARIATES by Gender



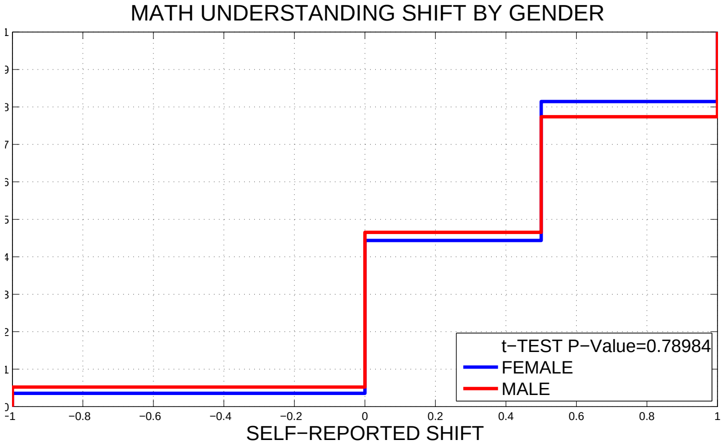
COVARIATES by Gender



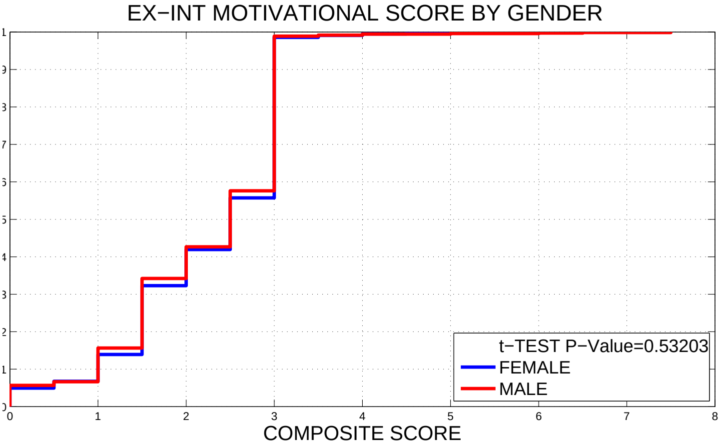
COVARIATES by Gender



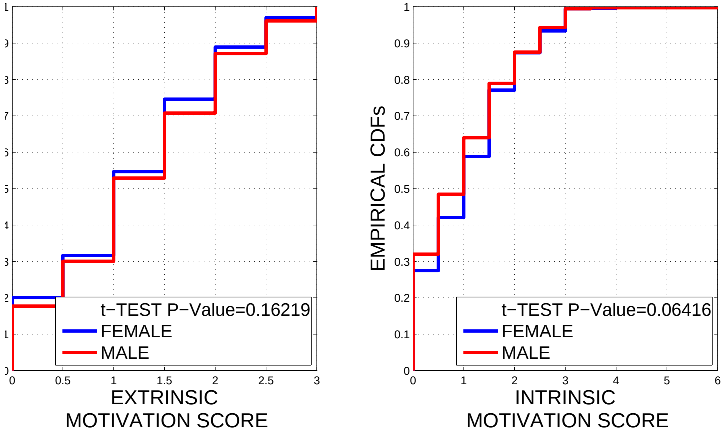
COVARIATES by Gender



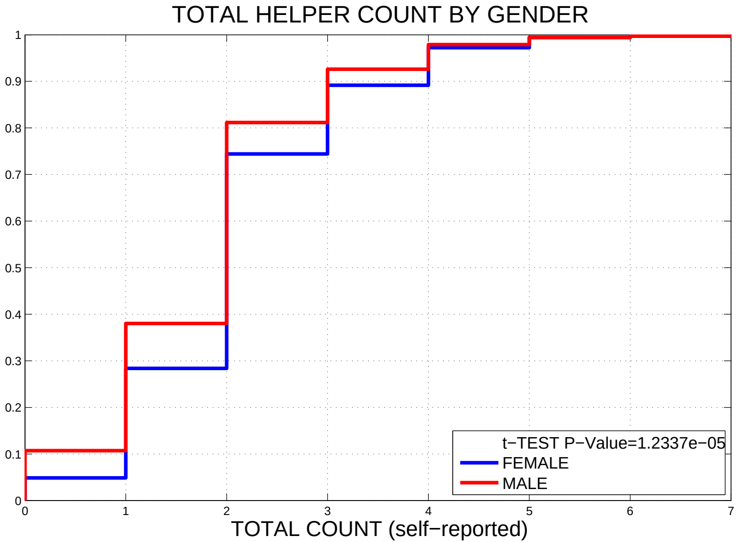
COVARIATES by Gender



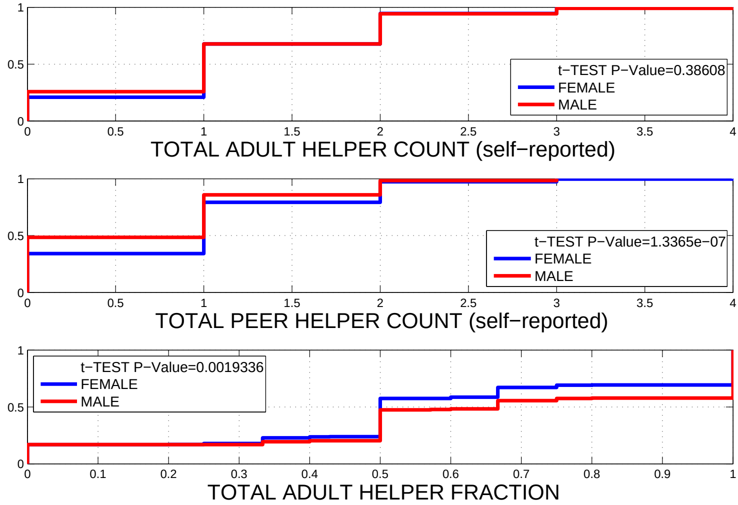
COVARIATES by Gender



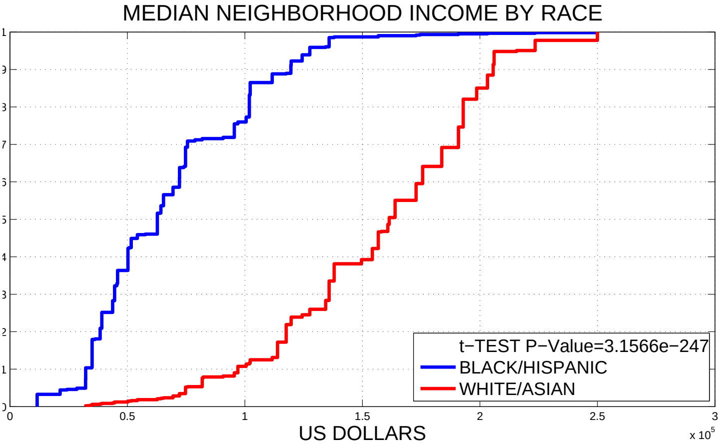
COVARIATES by Gender



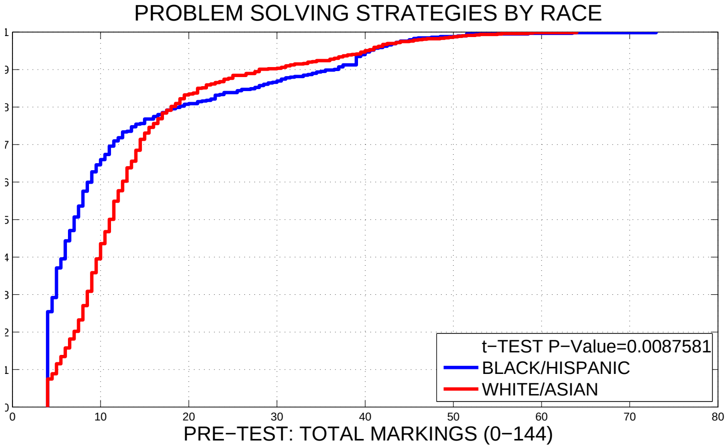
COVARIATES by Gender



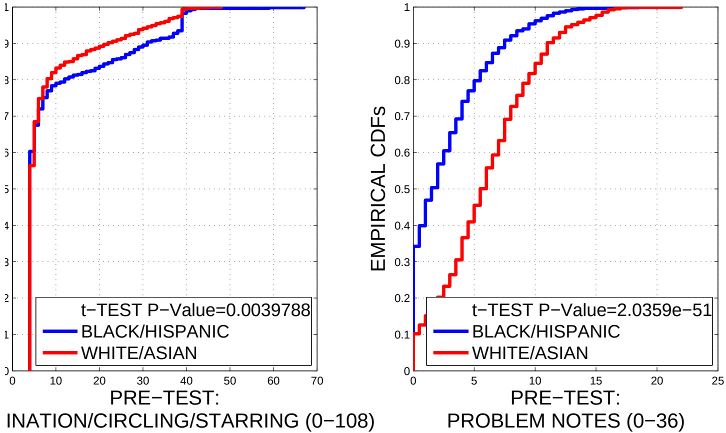
COVARIATES by Race



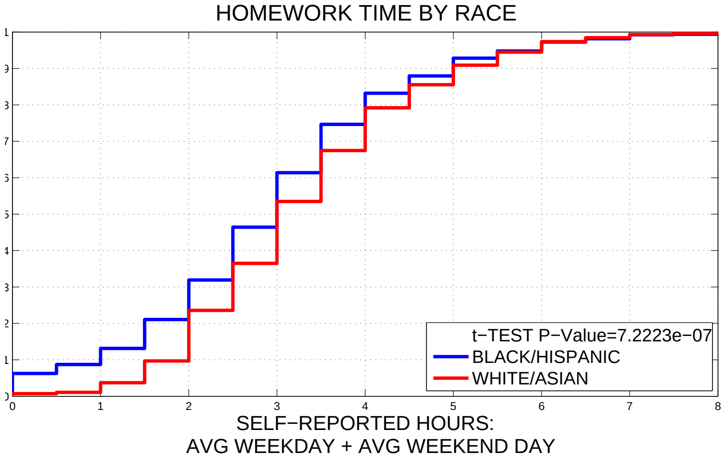
COVARIATES by Race



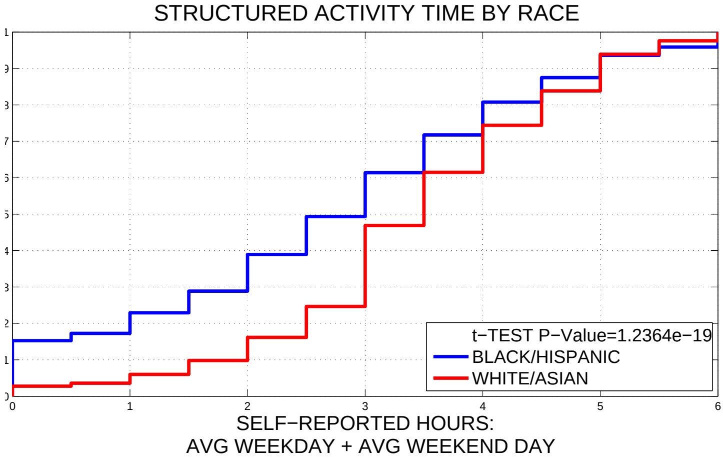
COVARIATES by Race



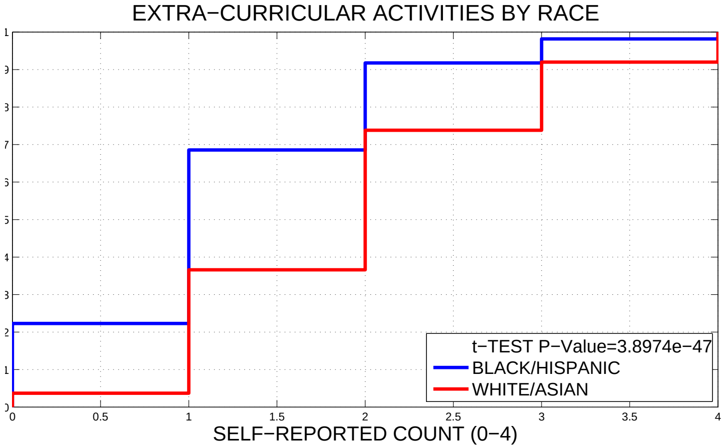
COVARIATES by Race



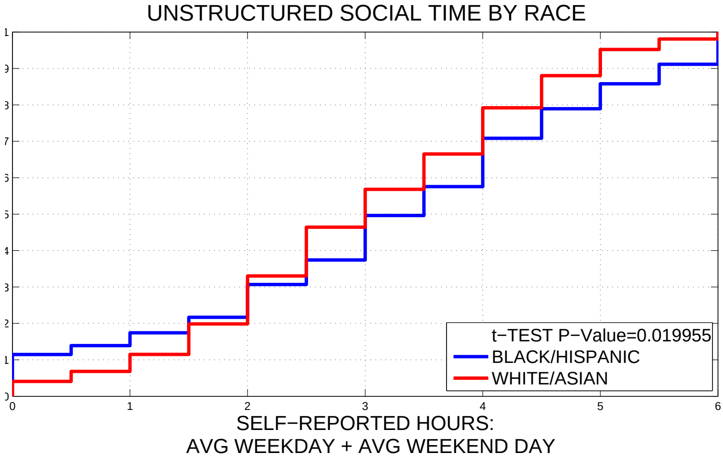
COVARIATES by Race



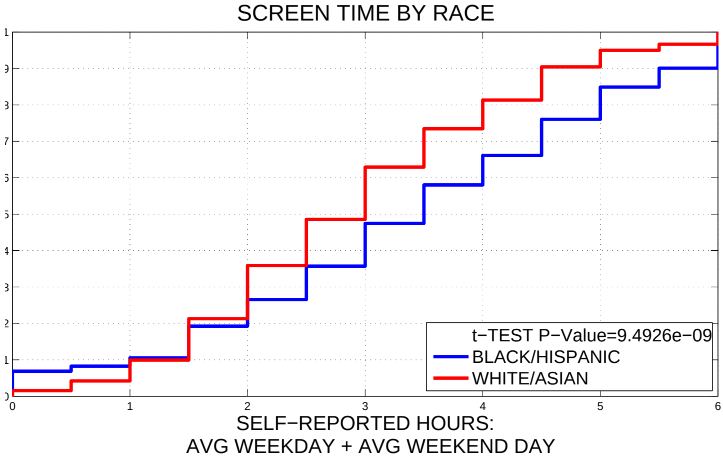
COVARIATES by Race



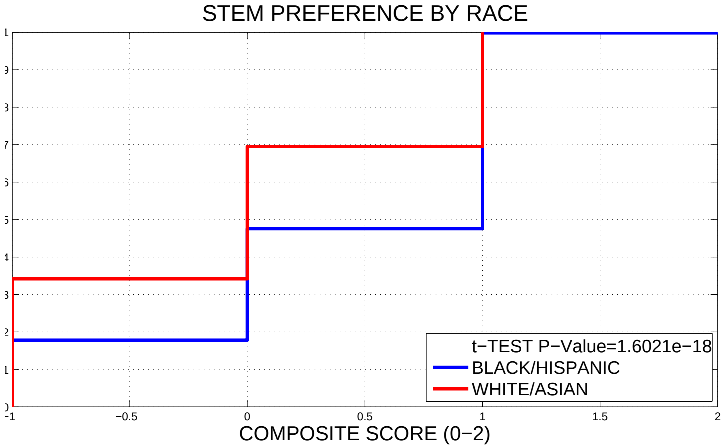
COVARIATES by Race



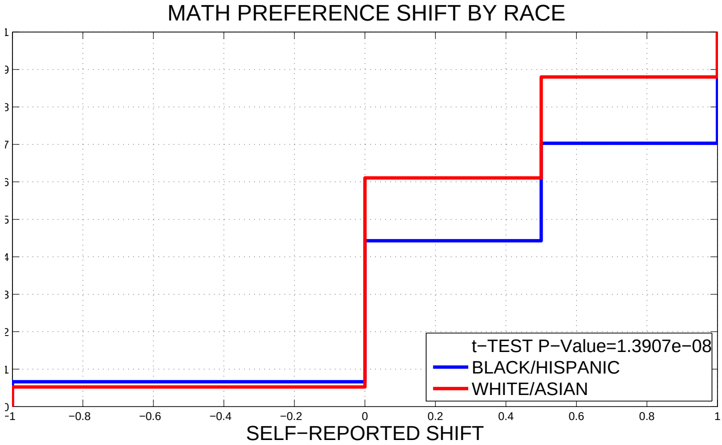
COVARIATES by Race



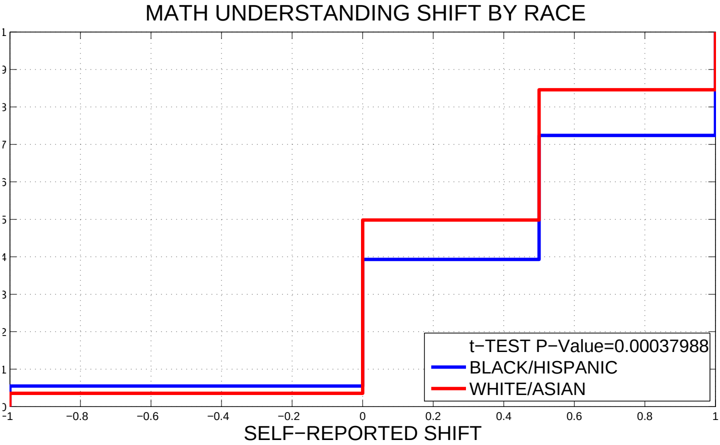
COVARIATES by Race



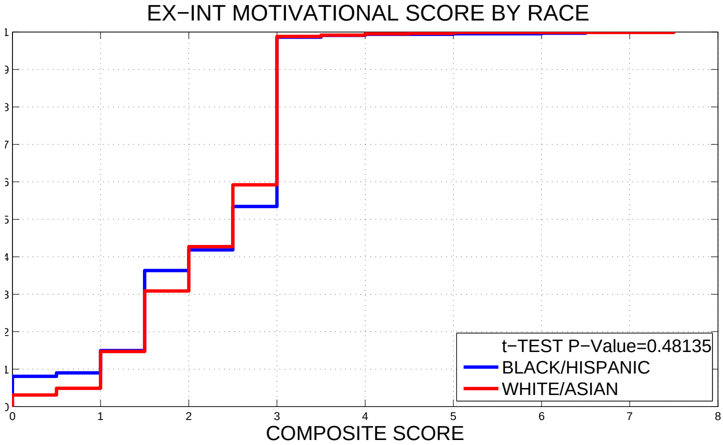
COVARIATES by Race



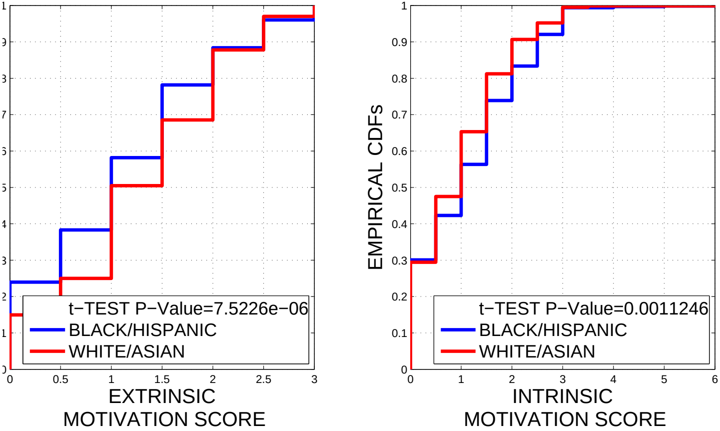
COVARIATES by Race



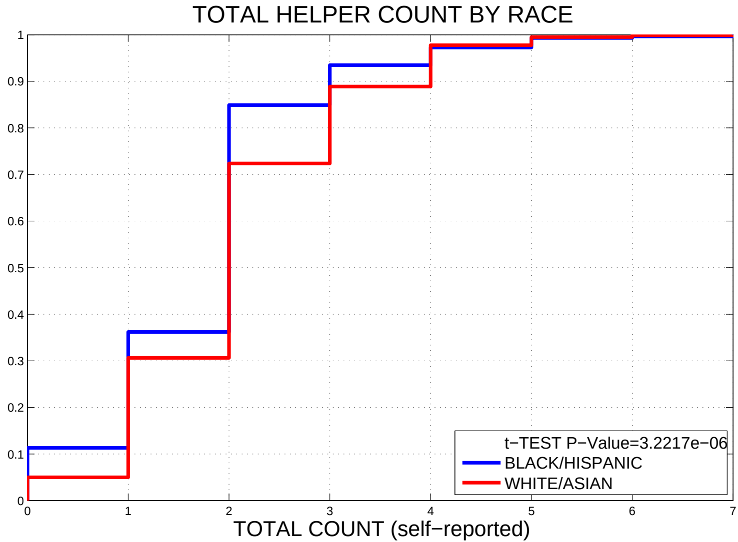
COVARIATES by Race



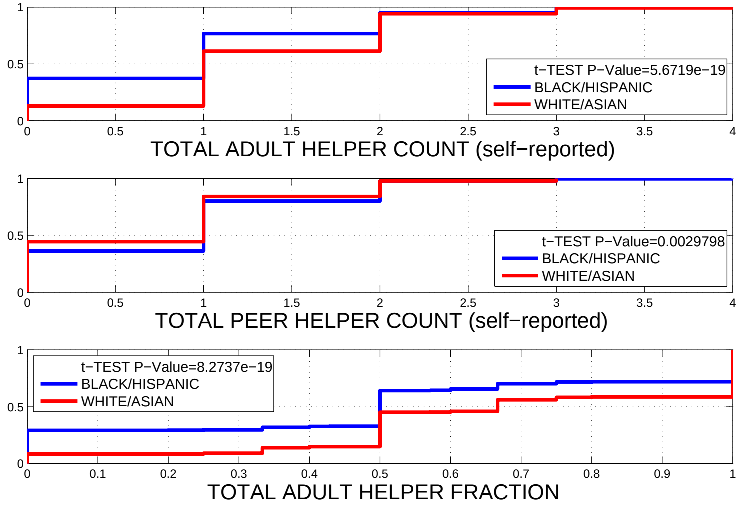
COVARIATES by Race



COVARIATES by Race



COVARIATES by Race



THE END

Time Accounting Issues: Data Truncation

PROBLEM: sometimes students begin a web session and then walk away in the middle, so how to tell spurious time observations apart and what to do about it?

Time Accounting Issues: Data Truncation

PROBLEM: sometimes students begin a web session and then walk away in the middle, so how to tell spurious time observations apart and what to do about it?

DIAGNOSING: Look for “holes" in the support of the distribution of web page times, t.

1. Bin time data by subject category, activity type

Time Accounting Issues: Data Truncation

PROBLEM: sometimes students begin a web session and then walk away in the middle, so how to tell spurious time observations apart and what to do about it?

DIAGNOSING: Look for "holes" in the support of the distribution of web page times, t.

1. Bin time data by subject category, activity type

2. Find first place where a kernel density estimator would equal zero
   - That is, for bandwidth $h = c\sigma N^{-1/5}$,
   - first order time observations from least to greatest: $t_1 < t_2 < \cdots < t_N$ and
   - find $t^* \equiv \min\{t_n \mid t_{n+1} - t_n > 2h\}$.

Time Accounting Issues: Data Truncation

PROBLEM: sometimes students begin a web session and then walk away in the middle, so how to tell spurious time observations apart and what to do about it?

DIAGNOSING: Look for "holes" in the support of the distribution of web page times, t.

1. Bin time data by subject category, activity type
2. Find first place where a kernel density estimator would equal zero
   - That is, for bandwidth $h = c\sigma N^{-1/5}$,
   - first order time observations from least to greatest: $t_1 < t_2 < \cdots < t_N$ and
   - find $t^* \equiv \min \{t_n \mid t_{n+1} - t_n > 2h\}$.
3. If t^* exists, truncate all observations $t > t^*$

Time Accounting Issues: Data Truncation

PROBLEM: sometimes students begin a web session and then walk away in the middle, so how to tell spurious time observations apart and what to do about it?

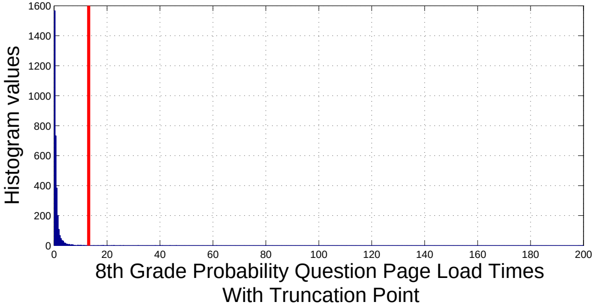
DIAGNOSING: Look for "holes" in the support of the distribution of web page times, t.

1. Bin time data by subject category, activity type
2. Find first place where a kernel density estimator would equal zero
   - That is, for bandwidth $h = c\sigma N^{-1/5}$,
   - first order time observations from least to greatest: $t_1 < t_2 < \cdots < t_N$ and
   - find $t^* \equiv \min\{t_n \mid t_{n+1} - t_n > 2h\}$.
3. If t^* exists, truncate all observations $t > t^*$
4. Repeat from step 2 until $t^* = \emptyset$

Time Accounting Issues: Data Truncation

EXAMPLE: 8th Grade Probability, Non-Practice Page Load (question) Times

- Truncation point: 13.1 minutes
- Percent of truncated observations: 0.63%



RESOLUTION: replace truncated observations by within-bin, within-student censored mean

Time Accounting Issues: Output Time (τ) Aggregation

DEFINITION: "Paid time" is time spent attempting a quiz that was eventually passed, prior to first success on that quiz.

DEFINITION: Students spend some time on quiz attempts which never lead to a paid success, and occasionally re-attempt a quiz after passing it the first time. Call this **"unpaid time"**.

Time Accounting Issues: Output Time (τ) Aggregation

DEFINITION: "Paid time" is time spent attempting a quiz that was eventually passed, prior to first success on that quiz.

DEFINITION: Students spend some time on quiz attempts which never lead to a paid success, and occasionally re-attempt a quiz after passing it the first time. Call this **"unpaid time"**.

- **METHOD 1:** Drop all unpaid time

Time Accounting Issues: Output Time (τ) Aggregation

DEFINITION: "Paid time" is time spent attempting a quiz that was eventually passed, prior to first success on that quiz.

DEFINITION: Students spend some time on quiz attempts which never lead to a paid success, and occasionally re-attempt a quiz after passing it the first time. Call this **"unpaid time"**.

- **METHOD 1:** Drop all unpaid time
- **METHOD 2:** Evenly spread unpaid time over all observed successes.

Time Accounting Issues: Output Time (τ) Aggregation

DEFINITION: "Paid time" is time spent attempting a quiz that was eventually passed, prior to first success on that quiz.

DEFINITION: Students spend some time on quiz attempts which never lead to a paid success, and occasionally re-attempt a quiz after passing it the first time. Call this **"unpaid time"**.

- **METHOD 1:** Drop all unpaid time
- **METHOD 2:** Evenly spread unpaid time over all observed successes.
- **METHOD 3:** Apply unpaid time serially to success times in progress.
  - Tommy passes Q1 in 15 min. Then he spends 2 min re-attempting Q1. Then he spends 1 min attempting Q5 without success. Then he spends 3 min attempting Q2 leading to a pass. Finally, he spends 1 more minute on Q5 but without success.
  - Total output: 2 units
  - $\tau_1 = 15min$
  - $\tau_2 = \tau_1 + 2 + 1 + 3 = \tau_1 + 6min$