

THE NON-MARKET BENEFITS OF ABILITIES AND EDUCATION

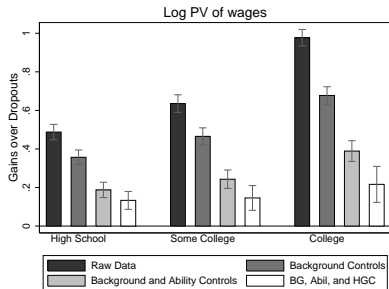
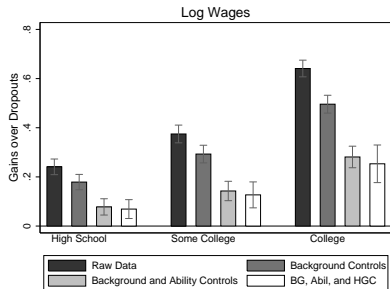
John Eric Humphries with James J. Heckman and Gregory Veramendi

October 1, 2015

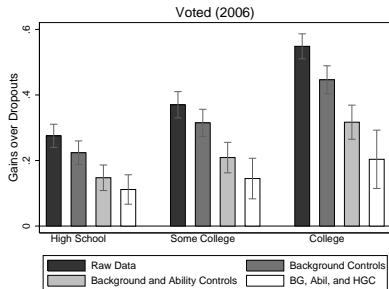
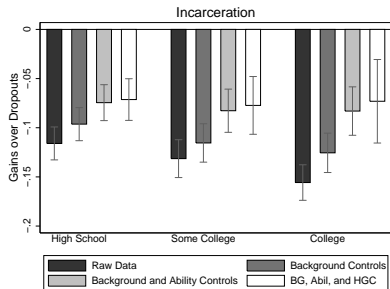
University of Chicago

INTRODUCTION

THE "EFFECT" OF EDUCATION



THE "EFFECT" OF EDUCATION



OUTLINE OF MODEL

Goal: Estimate dynamic model to recover the role of education and the role of skills on non-market outcomes.

Goal: Estimate dynamic model to recover the role of education and the role of skills on non-market outcomes.

- A generalized Roy framework:
 - . Finite vector of unobserved endowments generate dependencies between outcomes and schooling decisions
 - . Approximate agent's decision rule at each stage
 - . Do not impose selection on gains (important for non-market outcomes)

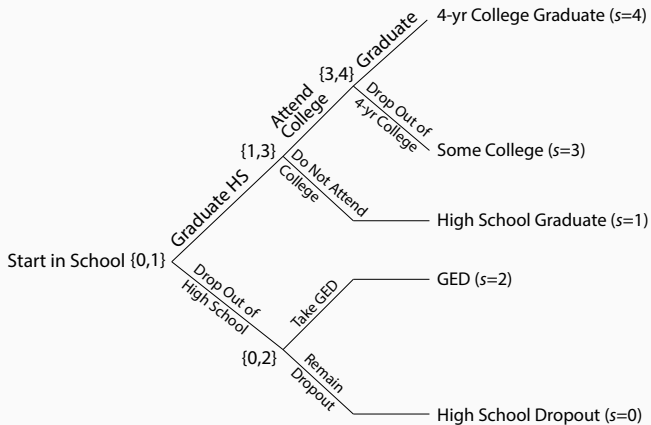
Goal: Estimate dynamic model to recover the role of education and the role of skills on non-market outcomes.

- A generalized Roy framework:
 - . Finite vector of unobserved endowments generate dependencies between outcomes and schooling decisions
 - . Approximate agent's decision rule at each stage
 - . Do not impose selection on gains (important for non-market outcomes)
- Cognitive and socioemotional endowments.
 - . Skill endowments affect educational choices.
 - . Skill endowments affect outcomes conditional on education.
 - . In combination, treatment effects vary by skill endowments.

1. Substantial ability bias.
2. Abilities play an important role in educational decisions and outcomes.
3. Returns to education differ by educational decision and abilities.
4. For many non-market outcomes, low-skill individuals see the largest benefits.

THE MODEL

SEQUENTIAL DECISION MODEL



Decision follows an index threshold-crossing property:

$$D_j = \begin{cases} 0 & \text{if } l_j \geq 0, \quad j \in \mathcal{J} = \{0, \dots, \bar{s} - 1\} \\ 1 & \text{otherwise,} \end{cases}$$

for $Q_j = 1, \quad j \in \{0, \dots, \bar{s} - 1\}$

where:

$$l_j = \underbrace{\phi_j(\mathbf{Z})}_{\text{Observed by analyst}} - \underbrace{\eta_j}_{\text{Unobserved by analyst}}, \quad j \in \{0, \dots, \bar{s} - 1\}$$

Outcomes can be discrete or continuous:

$$Y_s^k = \left\{ \begin{array}{ll} \tilde{Y}_s^k & \text{if } Y_s^k \text{ is continuous,} \\ \mathbf{1}(\tilde{Y}_s^k \geq 0) & \text{if } Y_s^k \text{ is a binary outcome,} \end{array} \right\}$$

$$k \in \mathcal{K}_s, \quad s \in \mathcal{S}.$$

where:

$$\tilde{Y}_s^k = \underbrace{\tau_s^k(\mathbf{X})}_{\text{Observed by analyst}} + \underbrace{U_s^k}_{\text{Unobserved by analyst}}, \quad k \in \mathcal{K}_s, \quad s \in \mathcal{S}.$$

We will use additional measures:

$$\mathbf{T} = \begin{pmatrix} T_1 \\ \vdots \\ T_M \end{pmatrix} = \begin{pmatrix} \Phi_1(\mathbf{X}) + e_1 \\ \vdots \\ \Phi_M(\mathbf{X}) + e_M \end{pmatrix}$$

Assume linear or binary models (though not a required assumption):

- Typically do not have access to individual test items in survey data
- Tend to be using a relatively small number of additional measures.

Assume a factor structure in errors:

$$\begin{aligned}\eta_j &= -(\boldsymbol{\theta}'\alpha_j - \nu_j), \quad j \in \{0, \dots, \bar{s} - 1\} \\ U_s^k &= \boldsymbol{\theta}'\alpha_s^k + \omega_s^k, \quad k \in \mathcal{K}_s, s \in S \\ e_m &= \boldsymbol{\theta}'\alpha_m + \epsilon_m, \quad m \in \{1, \dots, M\}\end{aligned}$$

- θ can be multidimensional.
- Agents know and act on θ .
- Allows for flexible correlations.

- Basic factor model:

$$T^m = \alpha^m \theta + \text{epsilon}^m$$

- Basic factor model:

$$T^m = \alpha^m \theta + \text{epsilon}^m$$

- Accounting for incentives or other observables:

$$T^m = \chi \beta^m + \alpha^m \theta + \text{epsilon}^m$$

- Basic factor model:

$$T^m = \alpha^m \theta + \text{epsilon}^m$$

- Accounting for incentives or other observables:

$$T^m = X\beta^m + \alpha^m \theta + \text{epsilon}^m$$

- Accounting for schooling at the time of the test:

$$T_s^m = X_s \beta_s^m + \alpha_s^m \theta + \text{epsilon}_s^m$$

THE FACTOR MODEL: WHICH MEASURES TO USE?

- Using this framework, we can use:
 - . Tests
 - . Self-reported behaviors
 - . Observed outcomes
- Measures can load on multiple factors.
- Choice of measures, imposed restrictions, and control variables can all affect the interpretation of the factors.
- We find our results are similar across specifications.

ESTIMATION AND DATA

- We allow for correlated endowments.
- We use robust mixture of normal approximations to the underlying endowments' distributions.

$$\begin{bmatrix} \theta_C \\ \theta_S \end{bmatrix} \sim p_1 \Phi(\mu_1, \sigma_1) + p_2 \Phi(\mu_2, \sigma_2)$$

- We allow for correlated endowments.
- We use robust mixture of normal approximations to the underlying endowments' distributions.

$$\begin{bmatrix} \theta_C \\ \theta_S \end{bmatrix} \sim p_1 \Phi(\mu_1, \sigma_1) + p_2 \Phi(\mu_2, \sigma_2)$$

- The sample likelihood is

$$\prod_{i=1}^N \int_{(\theta_C, \theta_S) \in \Theta} f(Y_i, D_i, C_i, S_i | X_i, t_C, t_S) dF_{\theta}(t_C, t_S)$$

- We allow for correlated endowments.
- We use robust mixture of normal approximations to the underlying endowments' distributions.

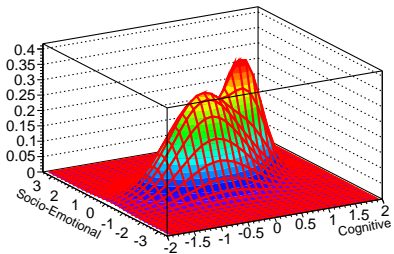
$$\begin{bmatrix} \theta_C \\ \theta_S \end{bmatrix} \sim p_1 \Phi(\mu_1, \sigma_1) + p_2 \Phi(\mu_2, \sigma_2)$$

- The sample likelihood is

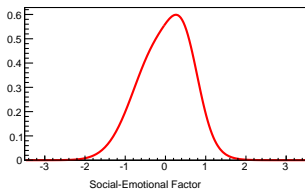
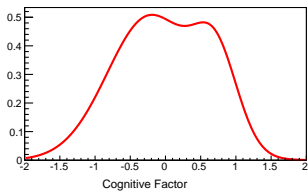
$$\prod_{i=1}^N \int_{(\theta_C, \theta_S) \in \Theta} f(Y_i, D_i, C_i, S_i | X_i, t_C, t_S) dF_{\theta}(t_C, t_S)$$

- Model is estimated in two stages using MLE
- Standard errors are calculated via bootstrap

FACTOR DISTRIBUTION



Distribution of Factors



Measurement System

- Cognitive endowment uses **ASVAB achievement tests**
- Both endowments use a set of grades from core courses (9th grade) and educational choice.

Outcomes

- Wages
- Incarceration
- Welfare Receipt
- Self-Esteem
- Depression
- Civic Participation
- Smoking

THE MEASUREMENT SYSTEM

- ASVAB sub-tests are assumed to measure only cognitive ability:

$$\text{ASVAB}^j = X\beta^j + \alpha^j\theta_c + \varepsilon^j$$

- 9th grade GPA in core subjects assumed to measure both cognitive ability and socio-emotional ability (Duckworth and Seligman 2005; Borghans, Golsteyn, Heckman, and Humphries 2012).

$$\text{GPA}^j = X\beta^j + \alpha_c^j\theta_c + \alpha_{se}^j\theta_{se} + \varepsilon^j$$

- Only need one dedicated measure that loads on only one factor (assuming two correlated factors) (Williams, 2013).

- Early self-reported behaviors also load on both endowments.
- Early behaviors include:
 - . early risky or reckless behavior
 - . early smoking
 - . fighting at a young age.

- Early self-reported behaviors also load on both endowments.
- Early behaviors include:
 - . early risky or reckless behavior
 - . early smoking
 - . fighting at a young age.
- Behaviors clearly depend on environment, but also provide a noisy signal of latent endowments.

- Early self-reported behaviors also load on both endowments.
- Early behaviors include:
 - . early risky or reckless behavior
 - . early smoking
 - . fighting at a young age.
- Behaviors clearly depend on environment, but also provide a noisy signal of latent endowments.
- Concerns of using early behavior to predict later behaviors (smoking)
- Other work shows using factors extracted from behaviors can have same explanatory power as factors extracted from measures of the Big-5 (Humphries and Kosse, 2015).

- Our estimates relatively unchanged when:
 - . Including or excluding risky behaviors.
 - . Restricting risky behaviors to load only on socio-emotional factor.
 - . Assuming risky behaviors measure third unrelated factor.
 - . Assuming ASVAB measures two dimensions of ability.

THE EFFECTS OF ENDOWMENTS

- Endowments impact outcomes two ways:

1. Endowments affect educational decisions:

$$\Pr(D_j = 1 | \boldsymbol{\theta} = \bar{\boldsymbol{\theta}}, \mathbf{X} = \mathbf{x})$$

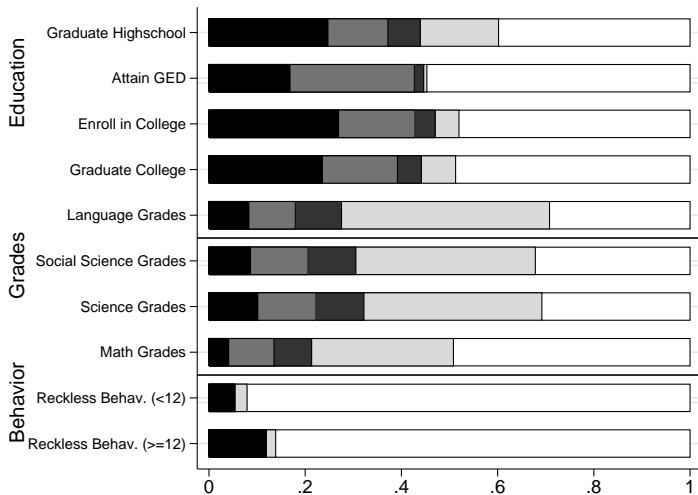
2. Endowments affect outcomes conditional on educational decisions:

$$E[Y_j | \boldsymbol{\theta} = \bar{\boldsymbol{\theta}}, \mathbf{X} = \mathbf{x}]$$

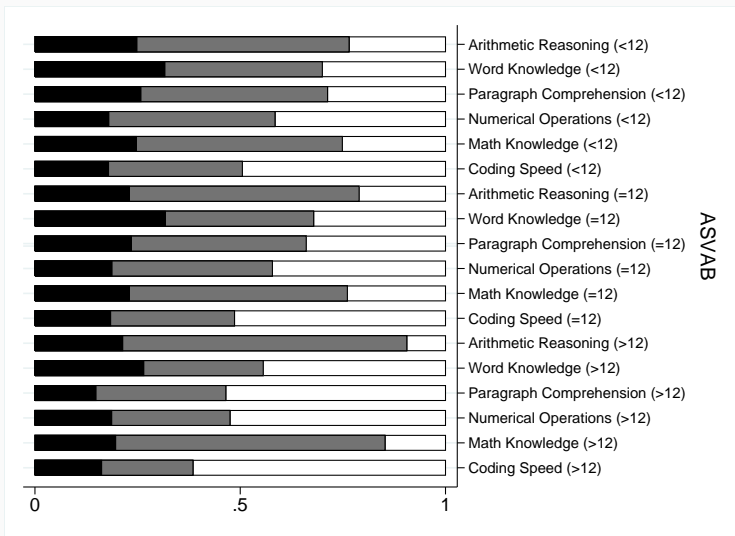
- Our model lets us decompose the role of abilities into the two components.

EXPLAINED VARIANCE

VARIANCE DECOMPOSITION

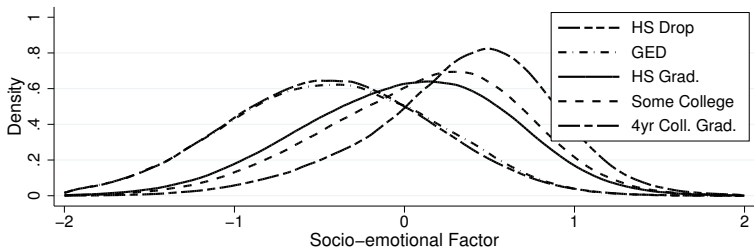
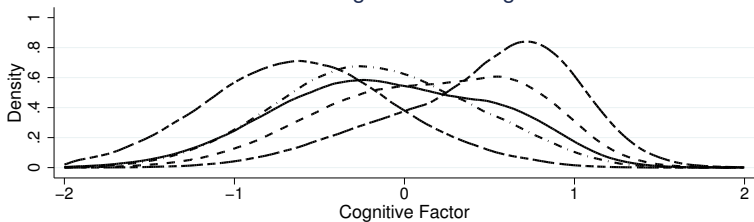


VARIANCE DECOMPOSITION

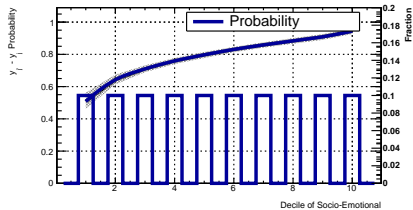
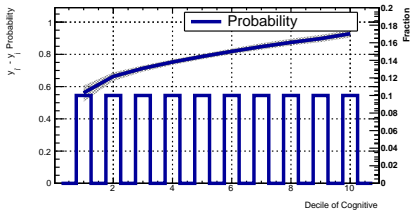
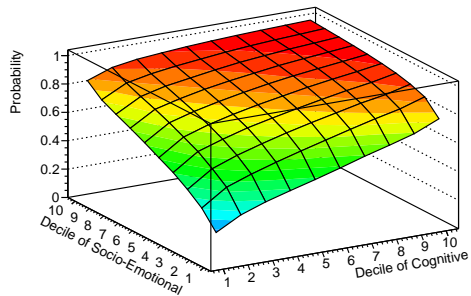


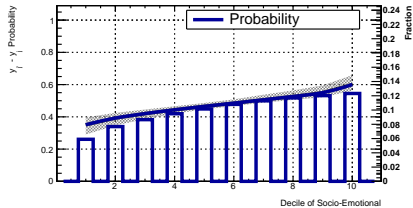
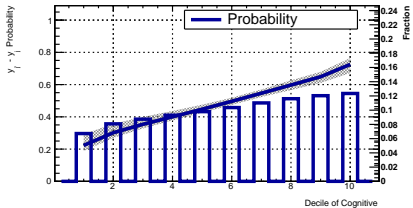
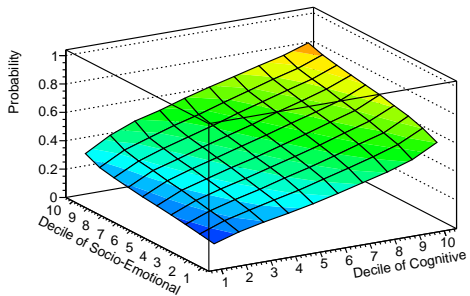
ENDOWMENTS ON EDUCATIONAL DECISIONS

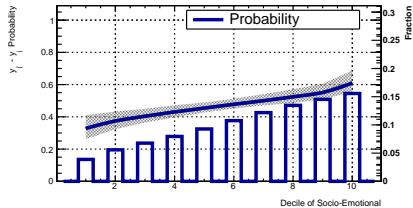
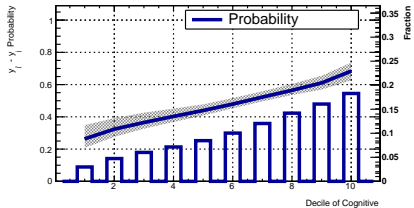
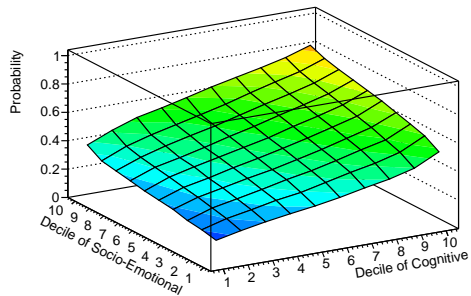
Sorting into Schooling



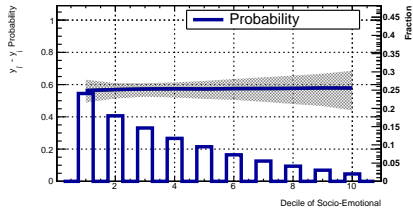
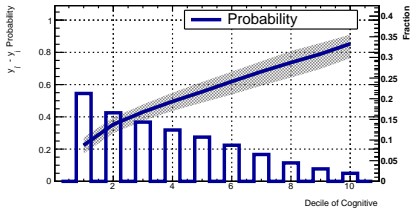
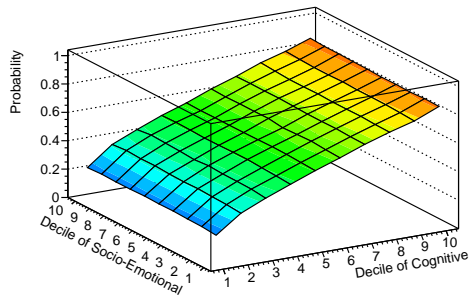
HIGH SCHOOL GRADUATION





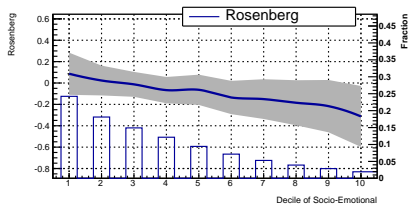
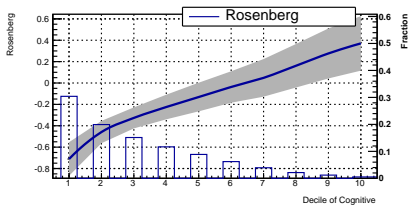
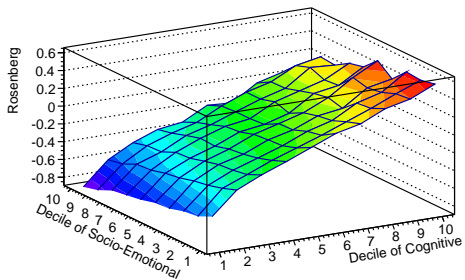


GED CERTIFICATION

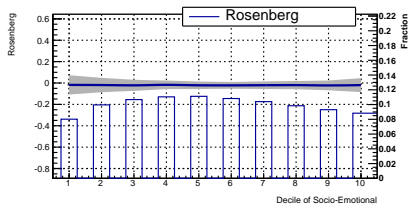
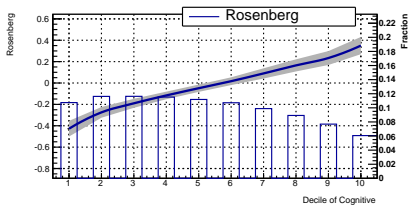
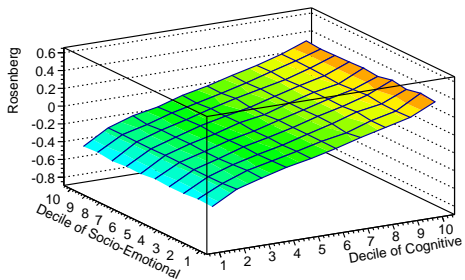


ENDOWMENTS ON CONDITIONAL OUTCOMES

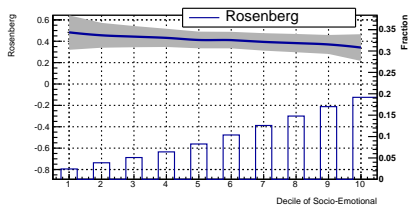
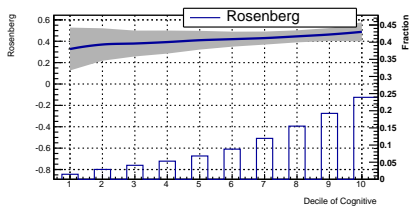
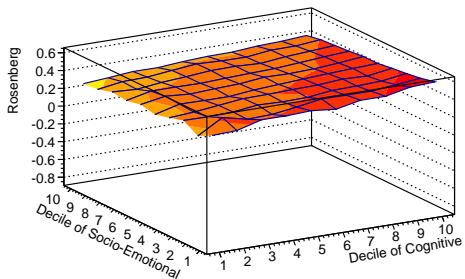
ROLE OF SKILLS ON SELF-ESTEEM (HIGH SCHOOL DROPOUTS)



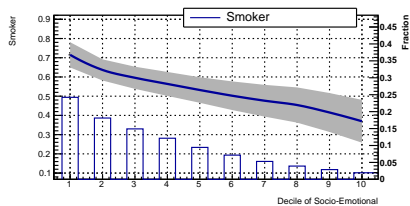
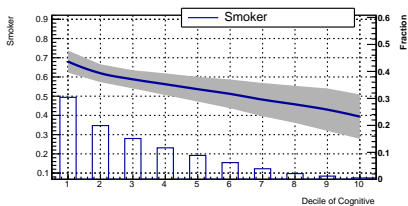
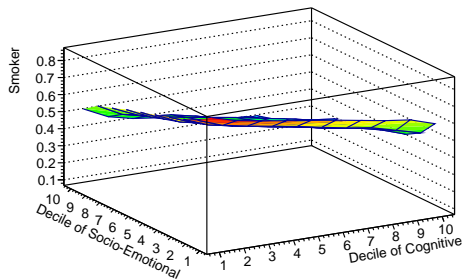
ROLE OF SKILLS ON SELF-ESTEEM (HIGH SCHOOL GRADS)



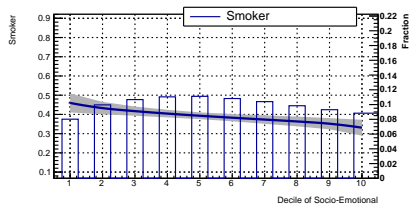
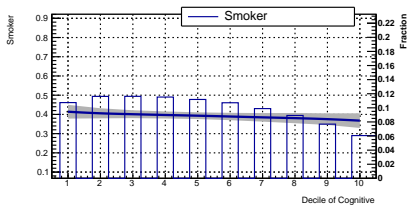
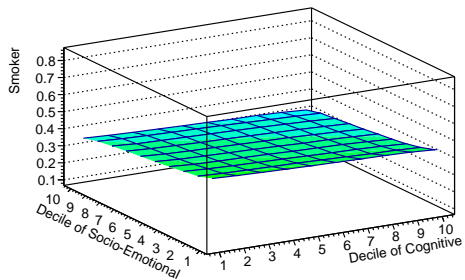
ROLE OF SKILLS ON SELF-ESTEEM (COLLEGE GRADS)



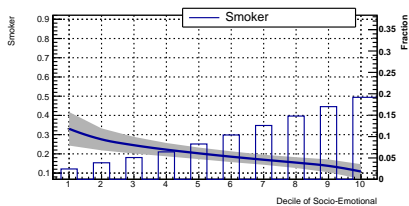
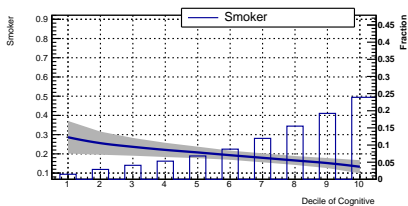
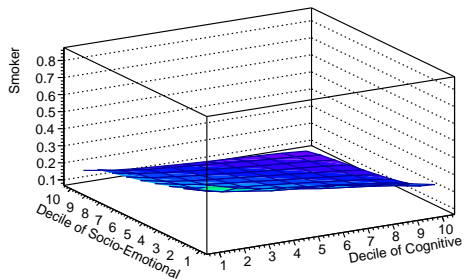
ROLE OF SKILLS ON SMOKING (HIGH SCHOOL DROPOUTS)



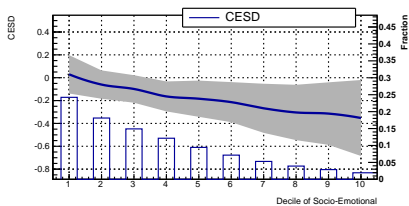
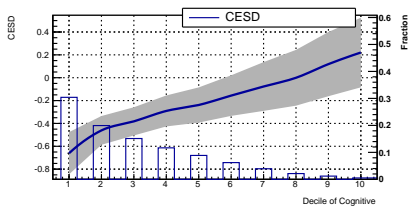
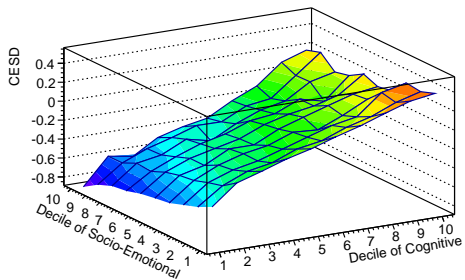
ROLE OF SKILLS ON SMOKING (HIGH SCHOOL GRADS)



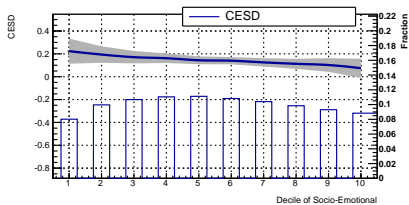
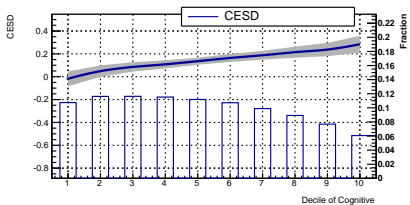
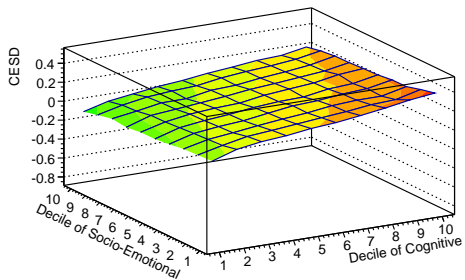
ROLE OF SKILLS ON SMOKING (COLLEGE GRADS)



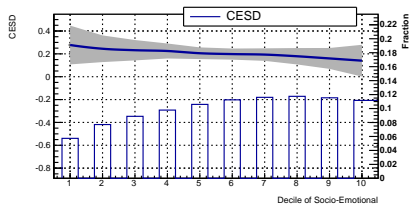
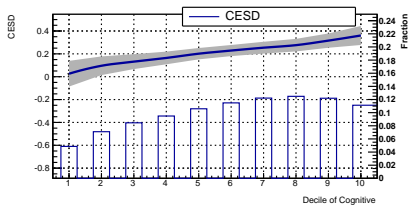
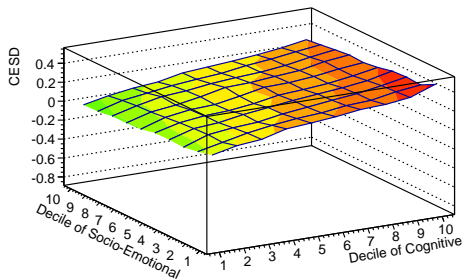
ROLE OF SKILLS ON DEPRESSION (HIGH SCHOOL DROPOUTS)



ROLE OF SKILLS ON DEPRESSION (HIGH SCHOOL GRADS)



ROLE OF SKILLS ON DEPRESSION (COLLEGE GRADS)



TREATMENT EFFECTS

- We can now consider the returns to education
- The effect depends on skills in two ways:
 - . Skills are priced differently by education level
 - . Skills affect the probability the individual goes on to pursue additional education (affects continuation values)

For each individual:

$$T_j^k[Y^k|X = x, Z = z, \theta = \bar{\theta}] := (Y^k|X = x, Z = z, \theta = \bar{\theta}, \text{Fix } D_j = 0, Q_j = 1) \\ - (Y^k|X = x, Z = z, \theta = \bar{\theta}, \text{Fix } D_j = 1, Q_j = 1)$$

Which can be decomposed into a direct effect (DE) and continuation value (CV)

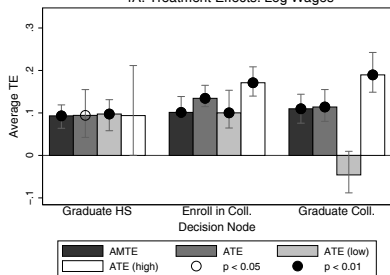
$$T_j^k = DE_j^k + C_{j+1}^k.$$

Where

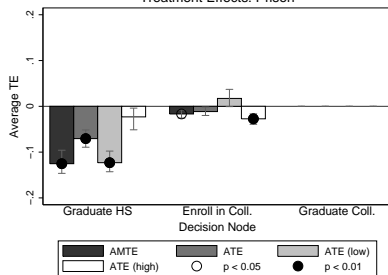
$$DE_j^k = Y_{j+1}^k - Y_j^k \\ C_{j+1}^k = \sum_{r=1}^{\bar{s}-(j+1)} \left[\prod_{l=1}^r D_{j+l} \right] (Y_{j+r+1}^k - Y_{j+r}^k).$$

TREATMENT EFFECTS BY DECISION NODE

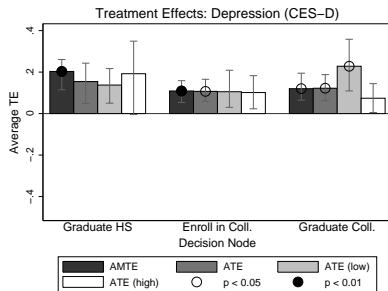
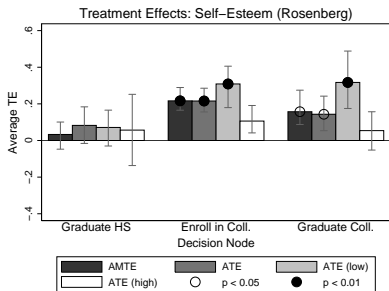
4A. Treatment Effects: Log Wages



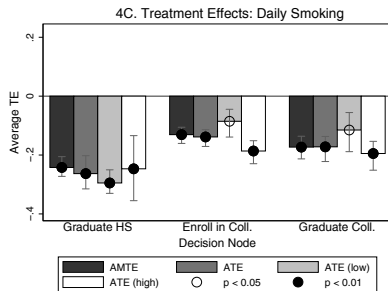
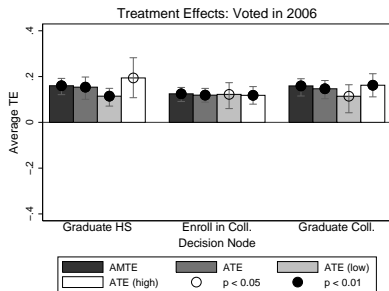
Treatment Effects: Prison



TREATMENT EFFECTS BY DECISION NODE



TREATMENT EFFECTS BY DECISION NODE



CONCLUSIONS

- Behaviors and self-reports can be used to extract measures of underlying ability.
- Cognitive and socio-emotional endowments influence schooling decisions and non-market outcomes.
- Skills influence outcomes most by their impact on educational decisions.
- Gains from education are higher for low-skill individuals for many non-market outcomes.

THANK YOU!

- Consider the case with three measures or test scores:

$$T^1 = X\beta_1 + \alpha_1\theta + \varepsilon_1$$

$$T^2 = X\beta_2 + \alpha_2\theta + \varepsilon_2$$

$$T^3 = X\beta_3 + \alpha_3\theta + \varepsilon_3$$

- Take the covariance:

$$\frac{\text{cov}(T^1, T^2|X)}{\text{cov}(T^2, T^3|X)} = \frac{\alpha_1}{\alpha_3}$$

$$\frac{\text{cov}(T^1, T^2|X)}{\text{cov}(T^1, T^3|X)} = \frac{\alpha_2}{\alpha_3}$$

- All loadings are identified with one normalization
- Factor distributions are non-parametrically identified (Kotlarski,1967)

- Factors can be predicted, but with error. Ignoring sampling error in $\hat{\alpha}^j$ and $\hat{\beta}^j$ and using one test:

$$\begin{aligned}\hat{\theta}_i &= \frac{1}{\hat{\alpha}^j} (T_i^j - X_i \hat{\beta}^j) \\ &= \theta_i + \varepsilon_i^j / \hat{\alpha}^j\end{aligned}$$

- Factors can be predicted, but with error. Ignoring sampling error in $\hat{\alpha}^j$ and $\hat{\beta}^j$ and using one test:

$$\begin{aligned}\hat{\theta}_i &= \frac{1}{\hat{\alpha}^j} \left(T_i^j - X_i \hat{\beta}^j \right) \\ &= \theta_i + \varepsilon_i^j / \hat{\alpha}^j\end{aligned}$$

- Using predicted factors leads to attenuation bias:

$$y = \alpha^y \hat{\theta} + \gamma \text{educ} + \epsilon$$

so

$$\text{plim}(\hat{\alpha}^y) = \alpha^y \left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_{\varepsilon_i^j}^2 / \hat{\alpha}^2} \right)$$

- Use measurement system to model measurement error
 - . Density is a pdf of errors given θ and X
 - . Can take flexible parametric assumptions like mixture of normals.

$$\prod_j f_j(T_i^j | X_i, \theta)$$

$$= \prod_j \left[\frac{1}{\sigma_{\epsilon^j} \sqrt{2\pi}} e^{-\frac{(T_i^j - X_i \beta^j - \alpha^j \theta_i)^2}{2\sigma_{\epsilon^j}^2}} \right]$$

- Estimation of measurement system gives us estimates of $\hat{\alpha}^j$, $\hat{\beta}^j$ and \hat{F}_θ
- Correct for attenuation bias using MLE
 - . Model the measurement error using the measurement system
 - . Integrate over the factor distribution

$$\mathcal{L} = \prod_{i=1}^N \int \left[f_{\text{wage}}(w_i | X_i, \theta) \underbrace{\prod_j f_t(T_i^j | X_i, \theta)}_{\text{Measurement System}} \right] dF_\theta$$

- Gives unbiased estimates of α^y and γ

A POLICY EXPERIMENT: INCREASING SKILLS

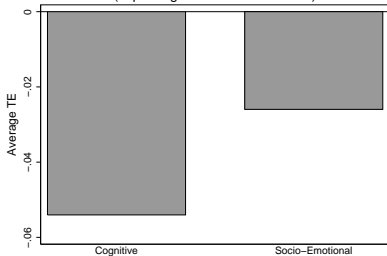
- Consider increasing the bottom decile of skill (cognitive or socio-emotional).
- Increase the bottom decile's skill by the difference between average skill in the 1st and 2nd deciles.
- Impacts schooling decisions and conditional outcomes.

Table: Policy Experiment: The impact of increasing skill in the bottom decile on educational sorting

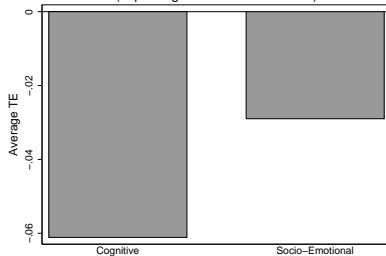
Increased Cognitive Skill						
	Proportion	DO	GED	HS	Enroll Coll	Grad Coll
DO	0.372	0.669	0.134	0.162	0.029	0.006
GED	0.107	0.000	0.735	0.195	0.053	0.017
HS	0.401	0.000	0.000	0.867	0.094	0.039
Enroll in Coll	0.086	0.000	0.000	0.000	0.841	0.159
Grad Coll	0.034	0.000	0.000	0.000	0.000	1.000

SIMULATED POLICY EXPERIMENT

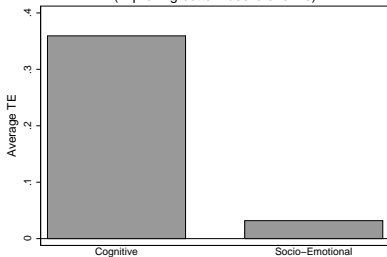
Policy Experiment: Any Welfare
(improving bottom decile of skills)



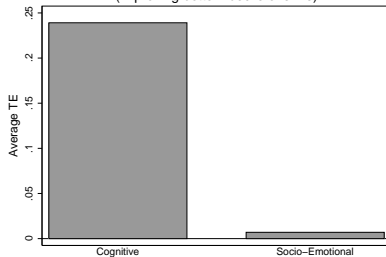
Policy Experiment: Prison
(improving bottom decile of skills)



Policy Experiment: Self-Esteem (Rosenberg)
(improving bottom decile of skills)

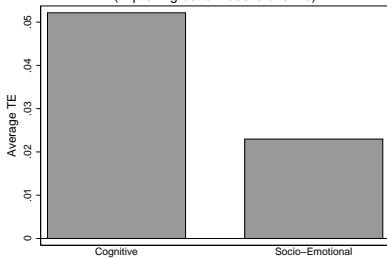


Policy Experiment: Depression (CES-D)
(improving bottom decile of skills)



SIMULATED POLICY EXPERIMENT

Policy Experiment: Voted in 2006
(improving bottom decile of skills)



9D. Policy Experiment: Daily Smoking
(improving bottom decile of skills)

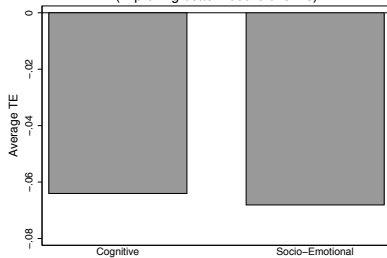


Table: Measurement System of Different Non-cognitive Constructs

Model	Measurement System
NC-LOCUS (NC-L)	Rotter's Locus of Control, Self-esteem.
NC-ENGAGEMENT (NC-E)	Frequency of engagement (volunteering, sport, technical work, reading), number of close friends.
NC-RELATIONS (NC-R)	Relation to parents and friends (bonding, love, argues or fights, problems solving), number of close friends.
NC-BEHAVIORS (NC-B)	Consumption behavior of alcohol and tobacco, eating behavior, argues or fights with family or friends.
Baseline (BASE)	Big-5 (conscientiousness, agreeableness, neuroticism, openness, extraversion), economic preferences (risk and time).

Source: Humphries and Kosse (2015)

Table: Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-Locus	NC-Engagement	NC-Relations	NC-Behaviors
NC-L	1			
NC-E	0.113	1		
NC-R	0.214	0.0968	1	
NC-B	-0.116	-0.0367	0.0844	1
Cons.	0.204	0.0959	0.186	0.134
Agree.	0.134	0.0217	0.218	0.0668
Neuro.	-0.302	-0.0125	-0.0741	0.110
Open.	0.128	0.187	0.182	-0.0546
Extra.	0.173	0.125	0.151	-0.144
Time	0.0954	0.0741	0.123	0.0974
Risk	0.0911	0.0952	-0.0320	-0.186
IQ	0.227	0.116	0.137	-0.131

Source: Humphries and Kosse (2015)

Table: Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-Locus	NC-Engagement	NC-Relations	NC-Behaviors
NC-L	1			
NC-E	0.113	1		
NC-R	0.214	0.0968	1	
NC-B	-0.116	-0.0367	0.0844	1
Cons.	0.204	0.0959	0.186	0.134
Agree.	0.134	0.0217	0.218	0.0668
Neuro.	-0.302	-0.0125	-0.0741	0.110
Open.	0.128	0.187	0.182	-0.0546
Extra.	0.173	0.125	0.151	-0.144
Time	0.0954	0.0741	0.123	0.0974
Risk	0.0911	0.0952	-0.0320	-0.186
IQ	0.227	0.116	0.137	-0.131

Source: Humphries and Kosse (2015)

Table: Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-Locus	NC-Engagement	NC-Relations	NC-Behaviors
NC-L	1			
NC-E	0.113	1		
NC-R	0.214	0.0968	1	
NC-B	-0.116	-0.0367	0.0844	1
Cons.	0.204	0.0959	0.186	0.134
Agree.	0.134	0.0217	0.218	0.0668
Neuro.	-0.302	-0.0125	-0.0741	0.110
Open.	0.128	0.187	0.182	-0.0546
Extra.	0.173	0.125	0.151	-0.144
Time	0.0954	0.0741	0.123	0.0974
Risk	0.0911	0.0952	-0.0320	-0.186
IQ	0.227	0.116	0.137	-0.131

Source: Humphries and Kosse (2015)

Table: Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-Locus	NC-Engagement	NC-Relations	NC-Behaviors
NC-L	1			
NC-E	0.113	1		
NC-R	0.214	0.0968	1	
NC-B	-0.116	-0.0367	0.0844	1
Cons.	0.204	0.0959	0.186	0.134
Agree.	0.134	0.0217	0.218	0.0668
Neuro.	-0.302	-0.0125	-0.0741	0.110
Open.	0.128	0.187	0.182	-0.0546
Extra.	0.173	0.125	0.151	-0.144
Time	0.0954	0.0741	0.123	0.0974
Risk	0.0911	0.0952	-0.0320	-0.186
IQ	0.227	0.116	0.137	-0.131

Source: Humphries and Kosse (2015)

Table: Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-Locus	NC-Engagement	NC-Relations	NC-Behaviors
NC-L	1			
NC-E	0.113	1		
NC-R	0.214	0.0968	1	
NC-B	-0.116	-0.0367	0.0844	1
Cons.	0.204	0.0959	0.186	0.134
Agree.	0.134	0.0217	0.218	0.0668
Neuro.	-0.302	-0.0125	-0.0741	0.110
Open.	0.128	0.187	0.182	-0.0546
Extra.	0.173	0.125	0.151	-0.144
Time	0.0954	0.0741	0.123	0.0974
Risk	0.0911	0.0952	-0.0320	-0.186
IQ	0.227	0.116	0.137	-0.131

Source: Humphries and Kosse (2015)

- All four 2-factor models predict GPA and college enrollment.
- They all are positively correlated with conscientiousness.
- Yet, they are not all positively correlated with each other.
- Loadings on many of the other traits differ.
- Suggests some consideration needed in which measures to include.