Hard Evidence on Soft Skills

James J. Heckman and Tim Kautz
University of Chicago

Development Economics Vice Presidency (DEC) Lecture
World Bank
Washington, DC
December 15, 2011
This draft, December 18, 2011
We live in an era of widespread testing. Achievement tests have assumed a prominent role. They are used to:

- Measure skills of persons (e.g. SAT, ACT, GRE) and certify suitability for admission and qualifications in a variety of domains of life.
- Measure the performance of schools and entire national school systems and nations (e.g. PISA scores, NCLB)

These tests are not well understood.
This lecture is about these tests and the consequences of relying uncritically on them.

- How predictive are they?
- What do they measure?
- How are they validated?
- What do they miss?
The short answer is that they miss—or perhaps better do not accurately capture—soft skills—personality traits, goals, motivations and preferences that are valued in the labor market, in school, and in many other domains.
Multiple skills are useful in predicting and causing economic and social success but with different weights in different tasks. People differ in their endowments of these skills. These differences in endowments and values of endowments across tasks give rise to comparative advantage and sorting in the labor market.

Personality traits—“soft skills”—are important skills that get neglected by the current emphasis on achievement tests in evaluating schools and screening people.

1. Soft skills can be measured.
2. They are predictive of many life outcomes—sometimes as strongly as measures of cognition.
3. Establishing the causal status of personality traits is not an easy task, but there is evidence on the question.
Ignoring personality traits deceives. Doing so distorts statistics and conceals major social problems.

Traits are stable across situations. Yet they can also be changed in a gradual way through investment. Enhancing these traits is an important policy option.

No evidence for extreme situational specificity of the sort advocated by Walter Mischel (1968), which still plays a prominent role in modern behavioral economics. There are “enduring traits” that persist across domains of economic and social life.
Origins of Achievement Testing: A Brief Review

- Mass application of IQ testing fostered the creation of tests of other psychological traits and of the knowledge learned in school.
Many creators of the IQ test saw its limitations.
Binet:

“[Success in school] . . . admits of other things than intelligence; to succeed in his studies, one must have qualities which depend on attention, will, and character; for example a certain docility, a regularity of habits, and especially continuity of effort. A child, even if intelligent, will learn little in class if he never listens, if he spends his time in playing tricks, in giggling, in playing truant.”

-Binet (1916, p. 254)
“What are the chief personality traits which, interacting with g, relate to individual differences in achievement and vocational success? The most universal personality trait is conscientiousness, that is, being responsible, dependable, caring, organized and persistent.”

-Jensen (1998, p. 575)
Figure 1: Modern View of “g”: An Hierarchical Scheme of General Intelligence and Its Components

Source: Recreated from Ackerman and Heggestad (1997), based on Carroll (1993).
Origins of the Modern Achievement Test

- Achievement tests were created in the wake of the success and widespread acceptance of the IQ test as a way to capture the knowledge acquired in schools and in general life.
General Knowledge

- Achievement test invented as a way to measure "general knowledge."—"useful knowledge" valuable in functioning at work and in society, not specific knowledge taught in a course.

- Designed to be "objective"—not depend on teacher assessments as captured by grades. This was perceived to be a way to implement meritocratic notions of education.

- Iowa tests; ACT; GED; No Child Left Behind; NAEP; PISA tests are modern versions.
A model of Academic Achievement at age $t$:

$$A_t = \phi_t(K_t, C_t, P_t, R_{A,t}) \quad t = 1, \ldots, T$$

(1)

$$K_t = \eta_t(Q_t, C_t, P_t, R_{K,t})$$

(2)

$A_t$: Academic Achievement at $t$
$K_t$: Cumulated Investment (Schooling, Parenting)
$C_t$: Cognition (Fluid Intelligence $G_f$ and Crystallized Intelligence $G_c$)
$P_t$: Personality
$R_{A,t}$: Incentives to perform on the test
$K_t$: Quality of Inputs
$R_{K,t}$: Incentives to make the investment
After age 10, $G_f$ becomes quite stable.

$P_t$ continues to evolve as does $G_c$.

Achievement tests reflect both cognitive and noncognitive traits, incentives, and acquired knowledge.

Value added models are based on changes in scores on achievement tests.

- Attempt to estimate (1) but typically assume linearity.
- Arbitrary scale (on $A_t$).
- Any monotonic transformation of a test score is a legitimate test score.
- To avoid (iii) anchor test scores in real world outcomes and do so in a nonparametric fashion (see Cunha et al., 2010).
The creators of these tests understood what they missed.

“We lean heavily on written examinations, on a few types of objective tests, and on the subjective impressions of teachers. Many other appraisal devices could be used, such as records of activities in which pupils participate, questionnaires, check lists, anecdotal records and observational records, interviews, reports made by parents, products made by the pupils, and records made by instruments (motion pictures, eye-movement records, sound recordings, and the like).” [Tyler, 1940]
Validation of Achievement Tests

How are these tests validated?
Usually on grades and other tests.
Ironic because the early pioneers decried the reliance of educational evaluations on “subjective” grades of teachers.
Table 1: Predictive Validities of Standard IQ and Achievement Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Domain over which it is validated</th>
<th>Estimated Validities</th>
<th>Source</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT</td>
<td>First year college GPA</td>
<td>0.35 to 0.53</td>
<td>Validity of the SAT for Predicting First-Year College Grade Point Average</td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>Grades in early years of college</td>
<td>0.42</td>
<td>ACT Technical Manual</td>
<td></td>
</tr>
<tr>
<td>Stanford-Binet</td>
<td>Correlations with other intelligence tests</td>
<td>0.77 to 0.87 with WISC-R</td>
<td>Rothlisburg (1987); Greene, Sapp, Chissom (1990)</td>
<td></td>
</tr>
<tr>
<td>WISC (Wechsler Intelligence Scale for Children)</td>
<td>Correlations with academic achievement</td>
<td>WISC: 0.443 to 0.751 with WRAT tests, 0.482 to 0.788 with 1st grade grades, 0.462 to 0.794 with 2nd grade grades; WISC-R: 0.346 to 0.760 with WRAT tests, 0.358 to 0.537 with 1st grade grades, 0.420 to 0.721 with 2nd grade grades</td>
<td>Hartlage and Steele (1977)</td>
<td>WRAT = Wide Range Achievement Test; Ranges are given because correlations vary by academic subject</td>
</tr>
</tbody>
</table>

Source: Almund et al. (2011).
Table 1: Predictive Validities of Standard IQ and Achievement Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlations with Other Intelligence Tests, Achievement Tests, and Outcomes</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAIS</td>
<td>0.67 (median) with verbal tests, 0.61 (median) with nonverbal tests, 0.69 with education attained, 0.32 with employability of mentally challenged, 0.38 to 0.43 with college grades, 0.62 with high school grades, 0.14 with nursing grades</td>
<td>Feingold (1982)</td>
</tr>
<tr>
<td>Raven’s Standard Progressive Matrices</td>
<td>0.74 to 0.84 with WAIS-R</td>
<td>O’Leary, Rusch, Guastello (1991)</td>
</tr>
<tr>
<td>GATB (General Aptitude Test Battery)</td>
<td>Supervisor rating performance in training programs and in job performance 0.23 to 0.65</td>
<td>Hunter (1986)</td>
</tr>
</tbody>
</table>

Source: Almund et al. (2011)
**Table 1: Predictive Validities of Standard IQ and Achievement Tests**

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
<th>Validity Range</th>
<th>Source</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASVAB (Armed Services</td>
<td>Performance in military training programs and military attrition rates</td>
<td>0.37 to 0.78 for training (mean=0.56); -0.15 for attrition</td>
<td>Schmidt (1988) for performance in training programs; Sticht et al (1982) for attrition rates</td>
<td>Large range in training correlations due to a variety of jobs</td>
</tr>
<tr>
<td>Vocational Aptitude Battery</td>
<td>Test difficulty is normed against graduating HS seniors. Test scores of high school seniors and grades of high school seniors</td>
<td>0.33 to 0.49 for HS Senior GPA</td>
<td>Technical Manual: 2002 Series GED Tests</td>
<td></td>
</tr>
<tr>
<td>GED (General Educational Development)</td>
<td>Correlations with academic achievement</td>
<td>0.13 to 0.62 for college GPA</td>
<td>Omizo (1980)</td>
<td>Large range is due to varying validity of eight subtests of DAT</td>
</tr>
<tr>
<td>DAT (Differential Aptitude Tests)</td>
<td>Correlation with other achievement tests; teacher ratings of student achievement</td>
<td>0.80 with grade 4 CAT/2, 0.69 with grade 5 CAT/2, 0.83 with grade 6 CAT/2, 0.67 with teacher ratings</td>
<td>Michalko and Saklofske (1999)</td>
<td>CAT=California Achievement Test</td>
</tr>
</tbody>
</table>

**Source:** Almund et al. (2011)
Psychometric validity of the GED

In two analyses, Means and Laurence (1984) found correlations of .75 and .79 between mean GED subtest scores and the military’s Armed Forces Qualification Test (AFQT) scores.

Correlations of .88 with its progenitor, the Iowa Test of Educational Development; .80 with the American College Test (ACT); .81 with the Adult Performance Level (APL) Survey; .77 with New York’s Degrees of Reading Power (DRP) Test; .66-.68 with the Test of Adult Basic Education (TABE); and .61-.67 on the General Aptitude Test Battery (GATB).

Baldwin et al. (1995): correlation of .78 between a general GED factor and a general National Adult Literacy Survey (NALS) factor.

“Correlations such as these provide evidence that the GED and other tests are measuring a common core of cognitive skills.”
-Boesel, Alsalam, and Smith (1998, p. 37)
Validities in Outcomes That Matter

- Grades and achievement tests more predictive than IQ
- None predict much of the variance in outcomes.
**Table 2: Validities in Labor Market Outcomes from the National Longitudinal Survey of Youth, 1979**

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IQ</td>
<td>GPA (10th grade)</td>
</tr>
<tr>
<td>Hourly Wage Age 35</td>
<td>0.03</td>
<td>0.05***</td>
</tr>
<tr>
<td>Hours Worked Age 35</td>
<td>0.10***</td>
<td>0.12***</td>
</tr>
<tr>
<td>Any Welfare Age 35</td>
<td>-0.09***</td>
<td>-0.11***</td>
</tr>
</tbody>
</table>

Source: Borghans et al. (2011)
Personality: The Missing Ingredient

- Big Five: OCEAN
  - Openness
  - Conscientiousness
  - Extraversion
  - Agreeableness
  - Neuroticism
Table 3: Validities for Personality Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Domain of Validation</th>
<th>Estimated Validities</th>
<th>Source</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hogan Personality Inventory</td>
<td>Correlations with delinquency criterion; Factor correlations with outcomes</td>
<td>0.00 to 0.67 with School Success, 0.68 to 0.73 with Avoids Trouble, 0.22 to 0.33 with Non-experience Seeking, -0.44 to 0.01 with Enjoys Crowds, -0.42 to 0.09 with Exhibitionist, 0.25 to 0.43 with Easy to Live With, 0.36 to 0.44 with Good Sense of Attachment, 0.10 to 0.43 with Not Depressed, 0.26 to 0.54 with No Guilt. Delinquency factor correlates: 0.91 with chargeable accidents, 0.80 with warning letters, 0.44 with suspensions; Absenteeism factor correlates: 0.62 with grievances, 0.61 with absences, 0.55 with medical absences, 0.44 with workers compensation claims; Negative Sanctions factor correlates: 0.68 with suspension letters, 0.67 with discharges; No Fault factor correlates: 0.71 with nonchargeable accidents; Supervisor’s Ratings factor: 0.60 with supervisor’s ratings, -0.38 with health history</td>
<td>Hogan &amp; Hogan (1989)</td>
<td></td>
</tr>
<tr>
<td>Myers-Briggs Type Indicator</td>
<td>Correlations with other personality tests; agreement between reported personality type and best-fit personality type</td>
<td>Correlation with Big Five based on Adjective Check List: -0.70 (E-I to Extraversion), 0.44 (S-Thompson (2009) N to Openness), 0.47 (T-F to Agreeableness), 0.54 (J-P to Conscientiousness); 72.9% report same four preferences as best-fit type, 18.2% report same three out of four preferences as best-fit type</td>
<td>Schaubhut, Herk,</td>
<td></td>
</tr>
</tbody>
</table>

Source: Almlund et al. 2011
### Table 3: Validities for Personality Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlations with other personality tests</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NEO PI-R</strong> (Revised NEO Personality Inventory)</td>
<td>Correlation with Positive Presentation Management Scale: -0.60 (N), 0.48 (E), 0.04 (O), 0.25 (A), 0.41 (C); correlations with Negative Presentation Management Scale: 0.39 (N), -0.46 (E), -0.31 (O), -0.38 (A), -0.54 (C); correlations with Big Five Index: 0.76 (E), 0.66 (A), 0.70 (C), -0.66 (N), 0.68 (O); correlations with Ten Item Personality Inventory: 0.65 (E), 0.59 (A), 0.68 (C), -0.66 (N), 0.56 (O)</td>
<td>Yang, Bagby, Ryder (2000); Gosling, Rentfrow, Swann (2003) N=neuroticism, E=extraversion, O=openness, A=agreeableness, C=conscientiousness</td>
</tr>
<tr>
<td><strong>NEO-FFI (NEO Five Factor Inventory)</strong></td>
<td>0.73 overall with BFI (Big-Five Index)</td>
<td>Gosling, Rentfrow, Swann (2003) Note: This is a shorter version of the NEO PI-R</td>
</tr>
<tr>
<td>Rotter Locus of Control</td>
<td>Correlation with high school GPA</td>
<td>Stipek &amp; Weisz (1981)</td>
</tr>
<tr>
<td>Rosenberg Self-Esteem Scale</td>
<td>Correlations with other self-esteem scales</td>
<td>Robins, Hendin, Trzeziewski (2001); Hagborg (1993) Correlations with Harter's done on an item by item basis</td>
</tr>
<tr>
<td>Short GRIT Scale</td>
<td>Item-level correlations with outcomes</td>
<td>Duckworth &amp; Quinn (2009) Large ranges due to variety of items</td>
</tr>
</tbody>
</table>

Source: Almlund et al. 2011
**Figure 2:** Association of the Big Five and intelligence with years of schooling

**Males**

- Emotional Stability
- Agreeableness
- Extraversion
- Conscientiousness
- Openness
- Fluid Intelligence
- Crystalized Intelligence

*Standardized Regression Coefficient*

Unadjusted for Intelligence

Adjusted for Intelligence

*Source: Almlund et al. 2011*
Figure 3: Correlations of The Big Five and Intelligence with High School Course Grades

Source: Almlund et al. 2011
Figure 4: Associations with Job Performance

Note. The values for personality are correlations that were corrected for sampling error, range restriction, and measurement error. Job performance was based on performance ratings, productivity data, and training proficiency. The authors do report the timing of the measurements of personality relative to job performance. The value for IQ is a raw correlation.

Source(s): The values reported for personality traits come from a meta-analysis conducted by Barrick and Mount [1991]. The value for IQ and job performance was reported in Schmidt and Hunter [2004].
Figure 5: Correlations of mortality with personality, IQ, and socioeconomic status (SES)

Source: Almlund et al. 2011
### Table 4: The Relative Predictive Power of Conscientiousness and SAT Scores for College GPA

<table>
<thead>
<tr>
<th>Source</th>
<th>Sample</th>
<th>Timing of Measurement and Outcome</th>
<th>Controls</th>
<th>Metric</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conard [2005]</td>
<td>University students in the US (N=186)</td>
<td>College GPA and SAT were both self-reported during college. Personality was measured in college.</td>
<td>Class Attendance</td>
<td>Standardized Regression Coefficient (β)</td>
<td>SAT Total Conscientiousness</td>
</tr>
<tr>
<td>Noftle and Robins [2007]</td>
<td>University students in the US (N=10,497)</td>
<td>College GPA and SAT were both self-reported during college. Personality was measured in college.</td>
<td>Gender, Other Big Five Traits</td>
<td>Standardized Regression Coefficient (β)</td>
<td>SAT Verbal SAT Math Conscientiousness</td>
</tr>
<tr>
<td>Wolfe and Johnson [1995]</td>
<td>University students in the US (N=201)</td>
<td>GPA and SAT were provided by the Colleges’ Record Office. Personality was measured in college.</td>
<td>High School GPA</td>
<td>Standardized Regression Coefficient (β)</td>
<td>SAT Total Conscientiousness</td>
</tr>
</tbody>
</table>
Personality Explains A Lot of the Variation in AFQT and Other Achievement Tests

\[ A_t = \phi_t( K_t, C_t, P_t, R_{A,t} ) \quad t = 1, \ldots, T \]

\[ K_t = \eta_t( Q_t, C_t, P_t, R_{K,t} ) \]
Figure 6: Decomposing Achievement Tests and Grades into IQ and Personality [NLSY79]

Source: Borghans, Golsteyn, Heckman et al. [2011].
Figure 7: Decomposing Achievement Tests and Grades into IQ and Personality [Stella Maris]

Source: Borghans, Golsteyn, Heckman et al. [2011].
Questions

- Are these relationships causal?
- Can we change these traits or are they fixed? Is promoting cognition and personality a useful policy option?
- What do we miss by ignoring soft skills?
Difficulties in Establishing Causality

- Reverse causality (clearly present in studies based on contemporaneous measures)
- Using lagged achievement measures creates an errors in variables problem. Current stocks determine current behavior.
- Multiple traits in conjunction with incentives predict task performance.

\[ Y^j_t = \phi_t(G^c_t, G^f_t, P_t, e^j_t) \] (3)

\[ e^j_t = \chi_t(G^c_t, G^f_t, P_t, R^j_t) \] (4)
• All psychological measurements are calibrated on task performances.

• A fundamental identification problem lies at the heart of any psychological measurement system for any particular trait — need to standardize for incentives and the effects of other traits in performing a task.

• Examples: Incentivizing IQ tests.

• All psychological measurements capture performance on some task or set of tasks.

• But typically analysts do not control for all of the factors that promote success on the tasks.
Causal Evidence on the Power of Soft Skills: Two Separate Sources

- GED testing Program
- Evidence from a social experiment
Case Study 1:
The GED as a case study of the power of soft skills and costs of neglecting them

- Evidence is strongly suggestive of the causal power of personality traits.
- GED is a test that secondary school dropouts can take to be certified as the equivalents of ordinary high school graduates.
- Use multiple data sets on outcomes, backgrounds, and abilities for all major economic and social groups in the U.S.
- 12% of all high school certificates issued in the U.S. are currently given to GEDs.
- GED is a group of 5 achievement tests normed against national samples of high school graduates (70% can pass).
GEDs are as smart as HSGs who do not go on to college.

GEDs who go on to college and succeed are indistinguishable from other college graduates in terms of annual wage income (True for AA and BA students).

*Terminal GEDs perform at a level very close to that of dropouts.*
We examine the essential life skills that GEDs lack.

Comparing GEDs to Dropouts standardizes ability (as measured by achievement tests) and examines the importance of the other life skills and, in particular, soft skills.
Figure 8: Cognitive ability by educational status

Source: Heckman, Humphries, Urzua, and Veramendi (2011)
Figure 9: Cognitive ability by educational status

Source: Heckman, Humphries, Urzua, and Veramendi (2011)
GEDs lack noncognitive traits measured in many ways: behaviors and personality test scores.
Figure 10: Distribution of Non-Cognitive Skills by Education Group

Figure 11: Distribution of Non-Cognitive Skills by Education Group

Gaps in both cognitive and noncognitive skills open up early and persist.
**Figure 12:** PIAT Scores across Ages and Education Groups - (All Races, All Post-Secondary Education)

(a) Males  
(b) Females

**Source:** Children of the National Longitudinal Survey of Youth 1979. **Notes:** The Peabody Individual Achievement Test (PIAT) is a widely childhood achievement test.
**Figure 13:** Behavioral Problems Index (BPI) across Education Groups - (All Races, All Post-Secondary Education)

(a) Males

(b) Females

Sources: Children of the National Longitudinal Survey of Youth 1979. Notes: The Behavioral Problems Index (BPI) is based on a 28 question survey given to parents about their child. The BPI is normalized to have mean 0 and variance 1.
The early childhood environments of dropouts and GEDs are worse than those of H.S. Grads. (Moon, 2011; Cunha et al., 2010)

Early family environments are more like those of dropouts.

Female GEDs have better early childhood environments as measured by parental investment—than males—close to those of H.S. Grads.

Males not so

Consistent pattern: Generally female GEDs have better traits then male GEDs—compared to dropouts.
These traits are highly predictive of who graduates from H.S. and who does not.

Noncognitive traits do not predict GED certification.
Figure 14: Probability of Graduating from High School (by cognitive and non-cognitive skill decile)

Source: Reproduced from Heckman et al. (2011).
Figure 15: Probability of GED Certification (conditional on dropping out, by cognitive and non-cognitive decile)

Source: Reproduced from Heckman et al. (2011).
**Figure 16:** Post-Secondary Educational Attainment Across Education Groups - (All Races, NLSY79)

(a) Males

(b) Females


• If a GED gets a BA or a MA, he/she earns at that level.
• However, there is usually delay.
• The present value of earnings for such people is substantially (20–30%) lower.
• Annual payment per unit of skill is the same for GEDs and dropouts.
Figure 17: Annual Income Differences - By Age - NLSY79 - (Males, All Races)

(a) All Post-Secondary Education  (b) No Post-Secondary Education

Source: National Longitudinal Survey of Youth 1979. Note: Estimates for hours worked and hourly wages exclude non-workers. Controls: “Raw” – age and region or state of residence; “Abil” – AFQT adjusted for schooling at time of test; “BG” – mother’s highest grade completed, urban status at age 14, family income in 1979, broken home status in 1979, south at age 14 and AFQT. Regressions exclude those reporting earning more than $300,000 or working more than 4,000 hours.
Figure 18: Hourly Wage Differences - By Age - NLSY79 - (Males, All Races)

(a) All Post-Secondary Education   (b) No Post-Secondary Education

Source: National Longitudinal Survey of Youth 1979. Note: Estimates for hours worked and hourly wages exclude non-workers. Controls: “Raw” – age and region or state of residence; “Abil” – AFQT adjusted for schooling at time of test; “BG” – mother’s highest grade completed, urban status at age 14, family income in 1979, broken home status in 1979, south at age 14 and AFQT. Regressions exclude those reporting earning more than $300,000 or working more than 4,000 hours.
Figure 19: Employment Differences - By Age - NLSY79 - (Males, All Races)

(a) All Post-Secondary Education  
(b) No Post-Secondary Education

Source: National Longitudinal Survey of Youth 1979. Note: Estimates for hours worked and hourly wages exclude non-workers. Controls: “Raw” – age and region or state of residence; “Abil” – AFQT adjusted for schooling at time of test; “BG” – mother’s highest grade completed, urban status at age 14, family income in 1979, broken home status in 1979, south at age 14 and AFQT. Regressions exclude those reporting earning more than $300,000 or working more than 4,000 hours.
**Figure 20:** Hours Worked Differences - By Age - NLSY79 - (Males, All Races)

(a) All Post-Secondary Education  
(b) No Post-Secondary Education

Source: National Longitudinal Survey of Youth 1979. Note: Estimates for hours worked and hourly wages exclude non-workers. Controls: “Raw” – age and region or state of residence; “Abil” – AFQT adjusted for schooling at time of test; “BG” – mother’s highest grade completed, urban status at age 14, family income in 1979, broken home status in 1979, south at age 14 and AFQT. Regressions exclude those reporting earning more than $300,000 or working more than 4,000 hours.
Avoiding Pretest Bias Or Cherry Picking of Results

- Many different specifications.
- Various specifications reported as “significant” in the literature.
- Mostly concentrated in the work of one author and his colleagues and students.
- No correction for pre-testing and model selection when reporting standard errors.
- One simple procedure to avoid such bias is to display results across all models and account for dependence in the estimates.
Figure 21: Distribution of the Effect of the GED Certificate and High School Graduation on Annual Income Across Models (Males)

(a) GED Recipients vs. High School Dropouts (Males)  
(b) High School Graduates vs. High School Dropouts

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 22: Distribution of p-values from Tests Comparing Annual Income of GED Recipients and High School Graduates to High School Dropouts (Males)

(a) GED Recipients vs. High School Dropouts (Males)  
(b) High School Graduates vs. High School Dropouts (Males)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 22: Distribution of p-values from Tests Comparing Annual Income of GED Recipients and High School Graduates to High School Dropouts (Males)

(c) GED Recipients vs. High School Graduates (Males)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
**Figure 23:** Distribution of p-values from Tests Comparing Hourly Wage of GED Recipients and High School Graduates to High School Dropouts (Males)

(a) GED Recipients vs. High School Dropouts (Males)  
(b) High School Graduates vs. High School Dropouts (Males)

**Sources:** National Longitudinal Survey of Youth 1979. **Notes:** Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated \( p \)-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother's highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 23: Distribution of p-values from Tests Comparing Hourly Wage of GED Recipients and High School Graduates to High School Dropouts (Males)

(c) GED Recipients vs. High School Graduates (Males)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated $p$-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Similar patterns found for hours of work and employment for males.
Using R.D. method produces no GED effects (sample of 100,000 GED test takers with long histories before and after the taking test) (Jepsen et al., 2010).

No signalling value — GEDs at the margin of passing (male and female) earn the same before and after certification

For people motivated enough to take the test, there is no effect of the GED on earnings.
Women

- Much of the previous literature focuses on males.
- Simplifies the econometrics: avoids selection into the labor force as an empirical issue.
- But misses an important empirical phenomenon.
- **Controlling for ability and baseline characteristics, there appear to be GED effects for certain groups of females.**
- But only for employment, hours worked, and hence earnings, but not in hourly wage rates.
**Figure 24:** Annual Income Differences - By Age - NLSY79 - (Females, All Races)

(a) All Post-Secondary Education  
(b) No Post-Secondary Education

Source: National Longitudinal Survey of Youth 1979. **Controls:** “Raw” – age and region or state of residence; “Abil” – AFQT adjusted for schooling at time of test; “BG” – mother’s highest grade completed, urban status at age 14, family income in 1979, broken home status in 1979, south at age 14 and AFQT. Regressions exclude those reporting earning more than $300,000 or working more than 4,000 hours.
Figure 25: Distribution of p-values from Tests Comparing Annual Income of GED Recipients and High School Graduates to High School Dropouts (Females)

(a) GED Recipients vs. High School Dropouts (Females)  
(b) High School Graduates vs. High School Dropouts (Females)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 25: Distribution of p-values from Tests Comparing Annual Income of GED Recipients and High School Graduates to High School Dropouts (Females)

(c) GED Recipients vs. High School Graduates (Females)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated $p$-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 26: Distribution of p-values from Tests Comparing Hourly Wage of GED Recipients and High School Graduates to High School Dropouts (Females)

(a) GED Recipients vs. High School Dropouts (Females)  (b) High School Graduates vs. High School Dropouts (Females)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother's highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 26: Distribution of p-values from Tests Comparing Hourly Wage of GED Recipients and High School Graduates to High School Dropouts (Females)

(c) GED Recipients vs. High School Graduates (Females)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated $p$-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Figure 27: Distribution of p-values from Tests Comparing Employment of GED Recipients and High School Graduates to High School Dropouts (Females)

(a) GED Recipients vs. High School Dropouts (Females)  
(b) High School Graduates vs. High School Dropouts (Females)

Sources: National Longitudinal Survey of Youth 1979. Notes: Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
**Figure 27:** Distribution of p-values from Tests Comparing Employment of GED Recipients and High School Graduates to High School Dropouts (Females)

(c) GED Recipients vs. High School Graduates (Females)

**Sources:** National Longitudinal Survey of Youth 1979. **Notes:** Estimates for hours worked and hourly wages exclude non-workers. This graph plots the estimated p-values of GED coefficients from a series of linear regressions. All models control for region, age, year, and AFQT score. The models differ in other controls and sub-populations of the data. The set of models includes all combinations of mother’s highest grade completed, urban residence at age 14, family income, lives in the south at age 14, smoked at 15, has had sex by 15, has committed a major crime, and 9th grade GPA. The sub-populations are all combinations of race, post-secondary education (everyone, has some post-secondary education, no post-secondary education), and age (measured in 5-year categories from 20 to 39) for males and females.
Why Do Women Benefit?

- *Two groups of relatively bright women benefit.* (relative to dropouts)

- One group is bright girls who were screw-ups in high school, dropped out, GED certified, and went on to attend and complete college.

- A second group of bright girls who drop out early (due to pregnancy) but on some dimension of personality and behavior are better than other dropouts—they GED certify late after their children enroll in school.

- They work hard, do not invest much on the job, have little wage growth and do not complete any further education.
This evidence *may* point to change—in preferences and/or constraints that cause some to turn around and the GED is a lifeline that promotes this.

Like “desistance” in the criminology literature. (e.g., Sampson and Laub)
Characteristics of GEDs Who Drop Out Due to Pregnancy

- Males and females drop out of high school for very different reasons.
- Males tend to drop out because they dislike school, lack ability, or work.
- Few males list marriage or parenthood as a major reason.
- Many women do. 42% of all female dropouts in the NLSY79 are pregnant before or during the same year that they dropout.
- 38% of all female GED recipients were pregnant before they dropped out of high school.
- Pregnant GEDs generally have better traits than other female GEDs
  - Better Grades in High School
  - Better AFQT
  - Better behavioral traits
Figure 28: AFQT Scores across Education Groups - (All Races, All Post-Secondary Education)

![AFQT Scores across Education Groups](image-url)

AFQT Scores across Education Groups - (All Races, All Post-Secondary Education)

- Oth (Other) Drop
- Preg (Pregnancy) Drop
- Oth (Other) GED
- Preg (Pregnancy) GED

Significance Levels:
- $p < 0.05$ (Preg vs. Non-Preg)
- $p < 0.05$ (Drop vs. GED)

Sources:

Notes:
The AFQT test was administered to the NLSY79 in 1980 when individuals were age 15 to 22. NLSY79 AFQT scores are adjusted for years of schooling at the time of test as described in the web appendix.

Heckman, Kautz
Hard Evidence on Soft Skills
Figure 29: Crime Factor across Education Groups - (All Races, All Post-Secondary Education)
High “Crime Factor” Means They Are Better

Sources:

Notes:
The crime factor is based on whether the respondent committed a petty crime (damaged property, shoplifted, petty theft, robbery, fraud, fencing), major crime (auto theft, breaking and entering private property, and grand theft), and physical crimes (fighting at work or school, assault and battery, and aggravated assault).
**Figure 30:** Highest Grade Completed across Education Groups - (All Races, All Post-Secondary Education)
Figure 31: 9th Grade GPA across Education Groups - (All Races, All Post-Secondary Education)
Figure 32: Sexual Intercourse Before 15 across Education Groups - (All Races, All Post-Secondary Education)
**Figure 33:** Drinks Age 15 across Education Groups - (All Races, All Post-Secondary Education)
Figure 34: Smokes by Age 15 across Education Groups - (All Races, All Post-Secondary Education)
Aside from early sexual intercourse they tend to engage in lower levels of risky behavior. Lower than the behavior of other pregnant dropouts who do not GED certify.

On these measures, they are much more similar to high school graduates.

Some of the GED recipients would likely have graduated from high school had they not become pregnant even without a change in traits.
We have no direct measures of personality post-pregnancy, but we have evidence on their behaviors.

However, they are persistent in the workplace after they re-enter.
The average age of GED receipt for the women who drop out due to pregnancy is 24.3—about 1.3 years later than other female GED recipients.

They GED certify on average when their youngest child is 3–4 years old – around the age when most children start preschool.
Female GEDs generally have better noncognitive traits, and this shows up in their sorting into positions in the labor market.
**Figure 35: Average Occupational Factor Scores by Final Education - Males**

![Graph showing average occupational factor scores by final education for males]

**Source:** The American Community Survey 2009 and O-Net. **Notes:** All educational categories are final education at time of interview. Each factor is based on the following O-Net occupational importance scores: **Cognitive** - active learning, analytical thinking, complex problem solving, critical thinking, deductive reasoning, inductive reasoning, interpretation of meaning, math reasoning, mathematics, processing information, reading comprehension, creative thinking, updating knowledge and visualization. **Social** - communicate to outside organizations, concern for others, customer or personal service, establish relationships, leadership, oral expression, persuasion, social perceptiveness, speaking, writing, written expression, active listening, and cooperation. **Physical Traits** - arm and hand steadiness, control and precision, coordination, depth perception, explosive strength, finger dexterity, gross body coordination, gross body equilibrium, manual dexterity, multi-limb coordination, reaction time, spatial orientation, stamina, static strength, stress tolerance, trunk strength, and wrist and finger speed.
Figure 36: Average Occupational Factor Scores by Final Education - Females

Source: The American Community Survey 2009 and O-Net. Notes: All educational categories are final education at time of interview. Each factor is based on the following O-Net occupational importance scores: Cognitive - active learning, analytical thinking, complex problem solving, critical thinking, deductive reasoning, inductive reasoning, interpretation of meaning, math reasoning, mathematics, processing information, reading comprehension, creative thinking, updating knowledge and visualization. Social - communicate to outside organizations, concern for others, customer or personal service, establish relationships, leadership, oral expression, persuasion, social perceptiveness, speaking, writing, written expression, active listening, and cooperation. Physical Traits - arm and hand steadiness, control and precision, coordination, depth perception, explosive strength, finger dexterity, gross body coordination, gross body equilibrium, manual dexterity, multi-limb coordination, reaction time, spatial orientation, stamina, static strength, stress tolerance, trunk strength, and wrist and finger speed.
Cross Section Comparisons: Potentially Dangerous

1. People can GED certify at different ages.
2. In a cross section, the experience of GEDs of the same age can be very different.
Longitudinal analysis generally supports the cross section analysis but provides a more nuanced interpretation of it.

Using a variety of procedures, employment gains for women.

No hourly wage gains.
Stability of Traits

- Are traits stable? Can people change?
- The evidence for women suggests this might be a possibility.
- But as a group, post-GED turnover behavior is quite stable.
- High quit rates for all who start anything: job, military, marriage.
- Laurence (2012): High dropout rates in the military
Figure 37: Cumulative Prob of Hazard Out of Employment (Males, All Races, All Levels of Post-Secondary Education)
Figure 38: Cumulative Prob of Hazard Out of Employment (Females, All Races, All Levels of Post-Secondary Education)
Figure 39: Cumulative Prob of Hazard Out of the Same Job (Males, All Races, All Levels of Post-Secondary Education)
Figure 40: Cumulative Prob of Hazard Out of the Same Job (Females, All Races, All Levels of Post-Secondary Education)

Cumulative Prob of Hazard Out of Same Primary Job – Female – All Races

Includes Ever Jailed – Sample With and Without Post–Sec Exp. – All Ability Levels

Years Since Start of Spell vs. Cumulative Prob of Exit by Given Pd

- Dropouts
- GEDs
- HSGs

Source: National Longitudinal Survey of Youth 1979 (NLSY79), nationally representative cross sectional sample.
**Figure 41:** Cumulative Hazard Rate into Divorce
(Males, All Races, All Levels of Post-Secondary Education)

![Cumulative Hazard Rate into Divorce](image)

Source: National Longitudinal Survey of Youth 1979 (NLSY79), nationally representative cross sectional sample.
Figure 42: Cumulative Hazard Rate into Divorce (Females, All Races, All Levels of Post-Secondary Education)

Cumulative Prob of Hazard Into Divorce – Female – All Race
Includes Ever Jailed – Sample With and Without Post–Sec Exp. – All Ability Levels

Years Since Start of Spell
Cumulative Prob of Exit by Given Pd

Source: National Longitudinal Survey of Youth 1979 (NLSY79), nationally representative cross sectional sample.
Figure 43: Cumulative Hazard Rate to First Time Incarcerated (Males, All Races, All Levels of Post-Secondary Education)
• The stability of traits challenges the situational specificity hypothesis that is current in behavioral economics.
• Claim: People adapt fully to situations.
The modern origins of the **situational specificity** hypothesis is based on the work of psychologist Walter Mischel:

“...with the possible exception of intelligence, highly generalized behavioral consistencies have not been demonstrated, and the concept of personality traits as broad dispositions is thus untenable”

-Mischel (1968, p. 146)
Many behavioral economists hold a similar view and appeal to Mischel as a guiding influence.

“The great contribution to psychology by Walter Mischel [...] is to show that there is no such thing as a stable personality trait.”

- Thaler (2008)

The evidence from the GED and much other evidence speaks strongly against the claims of Mischel and the behavioral economists. (See Almlund et al., 2011.)
Why Do People Get a GED If Returns Are So Low?

- Previous models of dropping out of school do not emphasize permanent shocks (e.g. Eckstein-Wolpin).
- Timing due to relative wage changes?
- Such timing is not an important explanation.
- Effects too small.
- The story for women is consistent with preference shocks and/or relaxation of constraints.
  - Birth of a child (stochastic)
  - Activates altruism (Oxytocin and other biological changes)
  - For cash-strapped women, the GED opens doors to employment.
  - Take low wage immediate payoff jobs — consistent with short-run liquidity constraints
Some stress the evolution of the prefrontal cortex and the race between intellectual ability and psychosocial maturity.

This is used to explain adolescent behavior.

May apply to the decision to take the GED.
Figure 44: Proportion of Individuals in Each Age Group Scoring at or Above the Mean for 26- to 30-Year-Olds on Indices of Intellectual and Psychosocial Maturity.

Source: From Steinberg, Cauffman, Woolard et al. (2009), submitted for publication.
• Story also consistent with time constraints: birth of a child $\Rightarrow$ dropout
• Relaxed time constraints as children age coupled with demand for income (to support child) $\Rightarrow$ GED certification
• This is supported by our longitudinal analysis.
• The women who take the GED are also better in certain noncognitive dimensions, closer to the margin of dropping out and also closer to margin of reentry.
• But a powerful factor explained in our book is that there are strong external incentive systems that foster GED test taking even in high schools.
Schooling boosts cognitive and noncognitive traits.
Figure 45: Causal Effect of Schooling on Measures on Cognition (from ASVAB)

Source: Heckman et al. (2006).
Figure 46: Causal Effect of Schooling on Measures on Cognition (from ASVAB)

Source: Heckman et al. (2006).
Figure 47: Causal Effect of Schooling on Two Measures of Socioemotional Skills

Source: Heckman et al. (2006).
**Figure 48:** Causal Effect of Schooling on Two Measures of Socioemotional Skills

Source: Heckman et al. (2006).
Perry Preschool Study: Family Supplementation Interventions Boost Noncognitive Traits

- Early childhood intervention
- Perry had lasting effects on life outcomes for both boys and girls.
- With a 7–10% annual rate of return for both (Heckman et al., 2010)
- Did not boost IQ
Figure 49: Cognitive Evolution Through Time, Perry Males: Male Cognitive Dynamics
Boosted Achievement Test Scores
**Figure 50:** Perry Age 14 Total CAT Scores, by Treatment Group

**CAT** = California Achievement Test  
Treatment: N = 49; Control: N = 46  
Statistically Significant Effect for Males and Females (p-values 0.009, 0.021 respectively)  
Boosted Measured Noncognitive Traits
Figure 51: Personal Behavior Index, by Treatment Group

Control

Treatment
Figure 52: Socio-Emotional Index by Treatment Group

Control

Treatment
### Figure 53: Decomposition of Treatment Effects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cognitive Factors</th>
<th>Socio-Emotional State</th>
<th>Personal Behavior</th>
<th>Other Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT total*, age 14(+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed, age 19 (+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly Income, age 27 (+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No tobacco use, age 27 (+)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of adult arrests, age 27 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jobless for more than 2 years, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever on welfare (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total charges of viol.crimes with victim costs, age 40, (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total charges of all crimes, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of lifetime arrests, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of adult arrests, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of misdemeanor arrests, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total charges of all crimes with victim costs, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any charges of a crime with victim cost, age 40 (-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

 Costs of Neglecting Soft Skills

- GED induces Dropping Out
- GED produces deceptive statistics
- Understates High School Dropout Rate
- Distorts emphasis in schools under test-based incentive systems
Summary

1. Soft skills get neglected in current emphasis on achievement test scores in the economics of education and in educational policy evaluations.

2. Lots of correlational evidence that soft skills predict lifetime success.

3. Causal evidence harder to come by. Evidence from experimental mediation analyses are consistent — as are causal estimates of the effects of education on skills.

4. GED testing program affords a natural experiment as do social experiments.

5. Compare people as bright as ordinary HSGs who don’t go to college with HSGs who go to college.

6. GEDs have lower levels of noncognitive — personality skills.
Most GEDs do not succeed.

Traits generally stable.

But some people might benefit as their circumstances change.

People who benefit are concentrated among above average in ability and noncognitive traits females who drop out because of pregnancy and then GED certify in mid-20s. Most do not go to college.

Benefits come mostly through enhanced employment prospects.

Whatever benefits arise are small in present value terms, compared to graduating high school.

Offsetting this is costs of dropout and deceptive social statistics.

Evidence from social experiments and studies of the effect of schooling on growth of measured traits show that theses traits can be changed by schooling interventions and enriched family environments.
References


Hard Evidence on Soft Skills

125 / 125


