10-29-2015

When You Say Nothing at All: The Predictive Power of Student Effort on Surveys

Collin Hitt
University of Arkansas, Fayetteville, cehitt@uark.edu

When You Say Nothing at All: The Predictive Power of Student Effort on Surveys

Collin Hitt[1]
Julie Trivitt
Albert Cheng
University of Arkansas

EDRE Working Paper No. 2014-10

Last Updated October 29, 2015

Please Do Not Cite Without Permission

Abstract

Character traits and noncognitive skills are important for human capital development and long-run life outcomes. Research in economics and psychology now shows this clearly. But research into the exact determinants of noncognitive skills have been slowed by a common data limitation: most large-scale datasets do not contain adequate measures of noncognitive skills. This is a particularly acute problem in education policy evaluation. We demonstrate that there are important latent data within any survey dataset that can be used as proxy measures of noncognitive skills. Specifically, we examine the amount of conscientious effort that students exhibit on surveys, as measured by their item response rates. We use six nationally representative, longitudinal surveys of American youth. We find that the percentage of questions left unanswered during the baseline year, when respondents were adolescents, is a significant predictor of later-life outcomes. Respondents with higher item response rates are more likely to attain higher levels of education. The pattern of findings gives compelling reasons to view item response rates as a promising behavioral measure of noncognitive skills for use in future research in education. We posit that response rates are a partial measure of conscientiousness, though additional research from the field of psychology is required to determine what exact noncognitive skills are being captured by item response rates.

Keywords: Noncognitive Skills; Educational Attainment; Employment Income; Human Capital
JEL Classifications: J24, I21

[1] Corresponding Author. Email: cehitt@uark.edu

Section I: Introduction

Noncognitive skills influence educational attainment, labor market outcomes, and other measures of well-being (Almlund et al. 2011; Borghans et al. 2008; Borghans, ter Weel and Weinberg, 2008; Bowles, Ginitis, and Osborne 2001; Carneiro, Crawford, and Goodman 2007; Deke and Haimson 2006; Heckman, 2000; Heckman and Rubinstein 2001; Heckman and Kautz, 2012; Heckman, Stixrud, and Urza 2006; Kaestner and Collison 2011; Lindqvist and Vestman 2011; Lundborg, Nystedt, and Rooth, 2014; Mueller and Plug, 2006). This finding has been a key contribution of human capital research and personality psychology over the past two decades. However, as researchers turn to policy questions regarding noncognitive skills, they encounter a pervasive data challenge: the large national datasets commonly used in economics and public policy research do not contain adequate measures of noncognitive skills.

Some survey and administrative datasets contain no measures at all of noncognitive skills. Other survey datasets do contain a few self-reported scales designed to capture skills such as academic effort and locus of control. But even when self-reported data are collected, scale scores based on self-reports contain poor information about students who are not conscientious enough to complete the survey.

We explore a new noncognitive measure based on the effort that students seem to exhibit on the surveys. Specifically we examine the frequency with which students skip questions or answer “I don’t know.” This variable can be used in datasets that contain no other noncognitive variables. And in datasets that contain at least some traditional measures of noncognitive skills, item response rates can be added to gain a fuller picture of students’ noncognitive skills.
Survey methodology research (e.g. Krosnick and Presser 2010, Smith 1995) has shown that survey response rates — the rate at which respondents actually answer the questions posed to

them — are driven strongly by factors other than cognitive ability. Long, low-stakes surveys require conscientious effort to complete, much like the daily busywork of school and employment. In education and human capital research, little work has been done using item response rates, or other indicators of effort on surveys, as a measure of noncognitive skills.

In our analyses of six large-scale datasets, we conduct a number of exercises to validate item nonresponse as a control variable for noncognitive skills. We show that it is predictive of educational outcomes, after controlling for a broad range of student- and household-demographic characteristics. The specific datasets we examine are the National Longitudinal Survey of Youth 1979 (NLSY:79), the National Longitudinal Survey of Adolescent Health (Add Health), The National Educational Longitudinal Study of 1988 (NELS:88), High School and Beyond (HSB:80), the National Longitudinal Study of Youth 1997 (NLSY:97), and the Educational Longitudinal Study of 2002 (ELS:02). These are important datasets for social science research. All of them follow nationally representative samples of American adolescents into adulthood.

We find evidence that survey item response rates capture important behavioral traits that are largely not captured by cognitive tests. By definition, they appear to capture *non*-cognitive skills. Item response rates consistently predict later educational attainment as standalone variables in sparse models. Before controlling for cognitive ability, item response rates are significantly predictive of later educational attainment in all six datasets. In the four datasets where item nonresponse is a significant predictor of educational attainment (independent of cognitive ability), a one standard deviation increase in item response rates is associated with completing 0.10 to 0.30 additional years of schooling. We also examine the association with employment status. Insofar as the skills captured by item response rate and self-reports influence wages and employment, they appear to do so mostly through their effect on educational attainment.

This study makes three important contributions. First, it shows that most surveys actually contain a behavioral, non-self-reported measure of noncognitive skills. It is important to have an objective measure. What respondents say about their noncognitive skills does not always reflect how they behave; item response rates provide behavioral information about respondents who may not have otherwise provided reliable information about themselves. Second, we identify a measure that can be used in datasets that contain no other valid measures of conscientiousness or academic effort. And third, we demonstrate the importance of thinking more creatively about existing data. Surely other latent measures of noncognitive skills exist in survey data that can provide additional new information about noncognitive skills, which we urge other researchers to explore.

The article proceeds as follows. In Section II, we review the economics literature on noncognitive skills, recent work from psychology highlighting measurement challenges, and survey methodology research on the problem of item nonresponse. In Section III, we describe the national datasets used for our analysis. In Section IV we discuss our empirical models. In Section V, we present the results of our analyses. In the final section, we discuss the results that suggest survey item response rates are a relevant source of missing information on important student noncognitive skills.

Section II: Literature Review

Survey Research in Economics and Psychology

Noncognitive skills are called *non*-cognitive for a simple reason. They are the personality traits, character virtues, emotional dispositions, and social skills that tests of cognitive skills fail to capture. Both noncognitive and cognitive skills influence educational attainment and earnings. Economists have recognized that students with similar cognitive abilities vary widely in educational and labor-market outcomes later in life (Heckman and Rubinstein, 2001). However, the specific noncognitive skills that predict educational attainment and earnings are often unobserved. In such analyses, the effect of noncognitive skills on these outcomes was presumably relegated to the residual, ascribed as measurement error or as a problem of omitted variables. This measurement challenge affects program evaluation and public policy analysis: for example, preschool and school-voucher programs have been shown to improve educational attainment without improving cognitive skills. The implied effect on noncognitive skills went unmeasured in the years immediately following the intervention (Chingos and Peterson, 2012; Duncan and Magnuson, 2013).

The field of personality psychology provides key insights into noncognitive skills. A personality trait that continually reemerges in the literature is conscientiousness. It and related behavioral traits such as grit and locus of control are now understood to be independently linked to academic and labor-market outcomes (Almlund et al. 2011). Conscientiousness is "the degree to which a person is willing to comply with conventional rules, norms, and standards" (Borghans et al. 2008; Hogan and Hogan, 2007). Facets of conscientiousness include orderliness, industriousness, responsibility and self-control (Jackson et al., 2009). With respect to educational outcomes, conscientious students are more likely to complete homework assignments, less likely to skip class, and tend to attain higher levels of education (Credé, Roch and Kieszczynka 2010; MacCann, Duckworth and Roberts 2009; Poropat, 2009; Trautwein et al. 2006; Tsukayama, Duckworth and Kim 2013). Conscientious workers are less likely to engage in counterproductive behaviors at work (Dalal 2005; Roberts et al. 2007); for example physicians in Pakistan rated higher in conscientiousness were less likely to miss work and falsify paperwork (Callen et al., 2014). Thus the question emerges: which policy interventions can increase conscientiousness as well as other important noncognitive skills, especially in children?

Unfortunately the datasets used in personality psychology — often limited samples of convenience — are usually inadequate to evaluate the relationship between noncognitive skills, social institutions, and public policy. Conversely, the massive surveys that many economists and public policy researchers depend upon rarely include measures based on the preferred survey instruments of psychologists. This is why Heckman and Rubenstein (2001) lament, “much of the neglect of noncognitive skills in analyses of earnings, schooling, and other lifetime outcomes is due to the lack of any reliable measure of them” (p. 145).

The preferred survey instruments of psychologists are lengthy questionnaires. For example, the highly-regarded Revised NEO Personality Inventory is a 240-item survey designed to measure what psychologists call the Big Five Personality Traits: Conscientiousness, Agreeableness, Neuroticism, Extraversion and Openness (Costa and McCrae, 2008). Psychologists focusing more narrowly on conscientiousness and self-control have used survey instruments like the 60-item Chernyshenko (2003) Conscientiousness Scales. Such scales are far lengthier than the scales usually included in national longitudinal surveys projects and program evaluations.

Yet economics research relies considerably on these very large national datasets. The research on noncognitive skills and educational attainment, in particular, leans heavily on large longitudinal surveys of children (Coughlin and Castilla 2014; Heckman et al. 2006). Such surveys are typically long but at most contain only short subsections about noncognitive skills. These survey design features limit the information on noncognitive skills that can be captured by the survey instruments. The short scales included in these surveys can be useful, but there are some important limitations for research. We present four examples.

First, the same scales are not used across surveys. This means that the same psychological constructs are not measured in all surveys, making it difficult to compare research on noncognitive skills across studies that use different datasets. This point is illustrated in greater detail in the following data section, where we discuss six major longitudinal datasets that we use in our analysis.

Second, even within the same survey, respondents may not interpret the questions about noncognitive skills in a similar way. This is illustrated by the problem of reference group bias. Self-reports of noncognitive skills are influenced by the reference group to which respondents compare themselves. As West et al. (2014) note:

> When considering whether "I am a hard worker" should be marked "very much like me," a child must conjure up a mental image of "a hard worker" to which she can then compare her own habits. A child with very high standards might consider a hard worker to be someone who does all of her homework well before bedtime and, in addition, organizes and reviews all of her notes from the day's classes. Another child might consider a hard worker to be someone who brings home her assignments and attempts to complete them, even if most of them remain unfinished the next day. (p. 6)

This is a particularly acute problem for program evaluation and public policy analysis. Educational interventions that actually increase noncognitive skills may not be measured as doing so. For example, two recent studies of charter schools have found large positive effects on standardized test scores, student behavior, or student educational attainment; yet the charter school students paradoxically report lower scores on self-reported measures of noncognitive skills (Dobbie and Fryer 2013; West et al. 2014). A likely explanation of these contradictory findings is that treatment-group students altered the standards by which they judged their own skills, reflecting different standards within the charter and comparison schools.

A third problem with current methods of measuring noncognitive skills is that some respondents do not even attempt to provide accurate information on surveys. Some engage in so-called “satisficing.” They provide socially desirable answers, select the first attractive answer option, or simply fill in the blanks without regard to the question asked (Krosnick 1991; Krosnick, Narayan and Smith, 1996). Some respondents simply do not answer questions at all, skipping the question or pleading ignorance.

King and Wand (2007) have proposed an alternative solution to tasks by using anchoring vignettes which statistically correct distortions that arise from reference-group bias. In turn, anchoring vignettes are becoming standard for inclusion in future surveys projects. One prominent example: the Programme for International Student Assessment now includes anchoring vignettes in the surveys deployed to thousands of students worldwide. Unfortunately, anchoring vignettes do not address satisficing or item nonresponse. Vignettes are designed to measure how a student views a normative concept; vignettes do nothing to motivate or capture the attention of disengaged students.

In order to measure motivation or self-control, some researchers also ask respondents to complete a task rather than answer survey questions. For example, Toburen and Meier (2010) use a word scramble activity for a behavioral measure of persistence. In the 1979 and 1997 National Longitudinal Surveys of Youth, which were conducted one-on-one, respondents were asked to complete a coding speed exercise, a sort of clerical task. Examining NLSY:79, Segal (2012) concluded that this was a proxy for noncognitive skills, conscientiousness in particular. While tasks may yield interesting information, there are also practical differences between explicitly-assigned tasks and our variable of interest, item response rates. The nature of assigning a task like coding speed alerts the respondent to the fact that her performance is being judged;

there is no such cue for item response. In our analysis, the survey is the task, and item response is a tacit measure of skills.

Behavioral tasks are also especially difficult to implement for large-scale surveys that use self-administered pen-and-paper formats, which brings us to the fourth problem. Neither anchoring vignettes nor behavioral tasks are common parts of the already-completed surveys that make up the datasets that economists and policy evaluators use.

Survey Effort and Survey Response Rates

We explore a partial solution to these challenges: surveys themselves can be viewed as tasks. In taking a survey, respondents are asked to complete a tedious task on mundane topics, with no external incentives to provide accurate information. For some students, surveys must seem much like homework. In the datasets we examine, many adolescent respondents skip questions or frequently answer “I don’t know,” plausibly signaling a lack of effort or focus. When students fail to answer questions, they leave holes in their survey record. Conventionally, researchers simply treat the items that respondents fail to answer as missing data or measurement errors. Observations with missing data are often dropped from analyses or new values are imputed (King et al. 1998).

We take a different approach. Instead of ignoring instances of item nonresponse, we view these so-called measurement errors as valuable pieces of information. Adolescent respondents may inadvertently show us something about how they approach the monotonous and mundane tasks of schooling and employment by how they approach a survey. Item nonresponse or its inverse, item response rates, can be revealing and used as variables in empirical analyses. We posit that the information captured by this variable contains information specifically about noncognitive skills. Following this literature review, we lay out a simple empirical model to estimate whether survey item response rates are predictive of educational attainment and labor-market outcomes, independent of cognitive test scores. We use this as an indirect test of whether item response rates capture noncognitive skills.

Previous literature contains only suggestive evidence on this question. For example, one can test the correlation between noncognitive scale scores and item response rates using cross-sectional data. Based upon the 2010 wave of the NLSY:97 and the 2009 wave of the German Socio-Economic Panel, Hedengren and Strattman (2012) have shown that the correlation between self-reports of conscientiousness and survey item response rates is positive. However, item response rate may be endogenous in Hedengren and Strattman's work because they examine a contemporaneous relationship. Although noncognitive ability as measured by item response rates may influence income or educational attainment, it is also possible that income or educational attainment influences response rates via the increased opportunity cost of time. This raises the possibility of simultaneity bias. Still, Hedengren and Strattman's work suggests that there are conceptual reasons to believe that item survey effort is related to noncognitive skills.

Other evidence from survey methods research suggests that item nonresponse is correlated with the noncognitive skills of respondents, though research methodologists rarely venture a guess at the precise noncognitive factors involved. It has long been established within the field of survey methodology that item nonresponse on surveys is not random (Krosnick and Presser, 2010). Among adults, income and educational attainment are positively correlated with item nonresponse (Smith, 1982). Question salience and survey format influence item response rates (Smith 1995), as can incentives (Singer and Ye 2013), suggesting strongly that item response rates are driven by individual motivation or habits — traits distinct from individual's cognitive ability to understand the questions asked.

We believe previous research provides credible evidence to consider item response a partial measure of noncognitive skills. However, the hallmark of noncognitive skills research is the ability of noncognitive measures to forecast later outcomes. To our knowledge, no published research has used item response rates to forecast later educational and labor-market outcomes. Insofar as previous research has compared item response rates to adult outcomes such as income and educational attainment levels, it has used cross sectional data or contemporaneous correlations. Any assessments of the association with education are typically done post hoc, since most respondents are adults typically finished with school. Comparisons to income are contemporaneous and thus suffer from problems of simultaneity bias (e.g. Hedengren and Strattman 2012). In survey methods research, educational level and income are typically used to explain the variation in item response rates, not vice versa.

It seems highly plausible to us that causation runs in the other direction. Item response rates (as a proxy for other noncognitive skills) may account for variation in educational attainment and income. Longitudinal data is needed to test this hypothesis, with item response rates measured during childhood. For adolescents, a survey is a routine but mundane task, like homework and financial aid applications. In adolescence one's willingness to complete these basic tasks of schooling has significant influence on educational attainment and employment earnings (Lleras 2008; Segal 2013). It stands to reason that item response rates on surveys may predict later outcomes as well. Our study is the first to use panel data to determine whether item response rates predict later outcomes. The use of panel data also addresses the problem of simultaneity bias when investigating contemporaneous correlations.

Before we proceed to a discussion of our data, it is important to note once more that even in the face of the limitations we have discussed, researchers have made remarkable progress investigating noncognitive skills. Research to date has been possible because many (and probably most) respondents indeed provide accurate and important information about their own noncognitive skills when asked. We are essentially examining the subset of students who do not exhibit strong effort on surveys, students whose self-reported noncognitive skills are unlikely to be accurate. Therefore the aim of our study is not primarily concerned with altering the empirical models used by noncognitive skills researchers. Rather, we investigate a measure of student effort that has unknowingly been omitted from those models.

Section III: Data

Our study uses six major longitudinal datasets that follow American middle and high school students into adulthood. Students participating in these surveys were born between 1957 and 1987. Each survey is designed to capture a nationally representative sample of American youth. In our analyses, we always use sampling weights to account for survey design effects and sample attrition so that all results remain nationally representative. Baseline survey years ranged from 1979 to 2002. The surveys contain rich data on student demographics and household characteristics. All participants were tested at baseline for cognitive ability. Below we briefly discuss facets of each dataset: the samples, survey modes, the types of item nonresponse that arise, and other explicit measures of noncognitive skills used.

≪Table 1 Here≫

≪Figure 1 About Here≫

Key features of each dataset are listed in Table 1. The descriptive statistics for item response rates in each dataset are shown in Table 2. Across datasets, the average item response rate is between 95 and 99 percent and between 14 percent and 54 percent of respondents completed every question on the survey – item response rates provide no information to distinguish between students with perfect response rates. Figure 1 shows distributions of response rates for each dataset. All are negatively skewed with obvious ceiling effects.

≪Table 2 Here≫

There is also an apparent relationship between survey mode and item response rate. The two NLSY surveys were administered one-on-one, in a face-to-face format. The response rates are far higher in the NLSY surveys than in the other surveys, which were self-administered and used pen-and-paper formats.

The National Longitudinal Study of 1979 (NLSY:79)

The NLSY:79 began with 12,686 male and female youths ranging in age from 14 to 22 as of December 31, 1978. Our analysis examines respondents aged 14 to 17 in the baseline year. Initial surveys were conducted in-person by professional interviewers following a pen-and-paper manual. Responses were logged by the interviewer. Item nonresponse (or "missing data") in the NLSY:97 stems from three sources: the refusal to respond to a particular item, an answer of "don't know", or the incorrect skipping of an item. Interviewers were responsible for distinguishing between refusals and answers of "don't know." The distinction between these two kinds of item nonresponse is therefore blurred. Also, the incorrect skipping of an item is primarily due to interviewer error. For the NLSY:79, we therefore define item nonresponse rate as the rate of refusals and answers of "don't know."

Regarding measures of noncognitive skills, respondents in the initial round of the NLSY:79 were asked a series of 23 questions adapted from the Rotter (1966) Locus of Control scale for adults. Higher scores indicate a high feeling of individual control over the events of one's life, while lower scores indicate a high level of external control.

High School and Beyond, 1980 (HSB:80)

High School and Beyond (HSB:80) followed two cohorts of students: the sophomore and senior classes from a nationally representative sample of US high schools in 1980. The analysis of HSB:80 begins with nearly 12,000 members of the senior-class cohort. We limit our analysis to this senior-class cohort; adult outcomes of the sophomore-class cohort are unavailable as they had barely completed undergraduate work at the final wave of data collection. The final year of the survey is five to six years after the end of high school, meaning that a substantial portion of the population has yet to enter the workforce after college. Thus, we include HSB:80 in only our educational attainment models. The survey mode was a self-administered pen-and-paper survey, with a proctor present. Questions were primarily multiple-choice of fill-in-the-blank format. “Don’t know” or “refuse” were answers options for very few questions. The most common instances of item nonresponse are when students skipped questions altogether. Some questions were asked only to a subset of students, conditional on answers to previous questions. For HSB:80, we define item nonresponse rate as the proportion of missing answers to all the questions that students should have answered conditional on answers to previous questions. HSB:80 also included two student-reported measures of noncognitive skills: the Rosenberg (1965) Self-Esteem Scale and the Rotter (1966) Locus of Control scale.

Several other longitudinal studies bear strong resemblance to the HSB:80. Among the datasets in our analysis, the National Education Longitudinal Study of 1998 and the Educational Longitudinal Study of 2002 are part of the same longitudinal study project administered by the U.S. Department of Education. We calculate item response rates similarly across those datasets.

The National Educational Longitudinal Study of 1988 (NELS:88)

NELS:88 interviewed about 12,000 eighth-graders during the spring semester of 1988, immediately before most students matriculated to high school. NELS:88 followed students until

2000, twelve years after their eighth grade year. NELS:88 used a self-administered, pen-and-paper survey instrument, similar to that used in HSB:80. Here again we calculate item nonresponse rates as the percentage of questions skipped by respondents. Similar to HSB:80, NELS:88 contains locus of control scale scores, as well as scores on a self-concept scale.

National Longitudinal Study of Adolescent Health (Add Health)

Add Health is a longitudinal survey of US middle and high school students. We use a publicly available version of the Add Health dataset. The public-use version contains roughly 6,000 student records that were randomly-selected from the full sample. These students completed a 45-minute, in-school pen-and-paper survey. The baseline survey year was 1994-1995. About 4,700 of the students were additionally selected for in-home follow up surveys. For our analysis, we use data from those who participated in the in-home surveys because key information such as educational attainment and labor-market outcomes, which are collected in 2007-2008, are available only for this subsample. Survey response rates, however, are based upon the in-school, pen-and-paper survey since in-home interviews were primarily conducted using a computer adaptive system that largely removed the possibility of skipping survey questions. As with other pen-and-paper surveys in our analyses, the primary source of item nonresponse comes from skipping items that should have been answered. For Add Health, we calculate item nonresponse rates as the percentage of questions that respondents were supposed to answer but skipped altogether. Add Health also contains items from the Rosenberg (1965) self-esteem scale, which we incorporate into our analysis.

The National Longitudinal Survey of Youth 1997 (NLSY:97)

NLSY:97 is a survey of 8,984 American youths aged 12 to 17 in 1997. Surveys were computer-adaptive, administered in home with the assistance of a professional interviewer.

Questions were primarily multiple-choice and “unsure” was a frequent answer option. Refusal to answer was also a response option, though prompts from computer software and the interviewer made outright refusal a less likely response than in the NLSY:79. We calculate item nonresponse as the rate at which interviewees answer “unsure” or refuse to answer items.

The NLSY:97 is rare among longitudinal datasets in that it includes a behavioral task that has been shown to measure noncognitive skills. As part of the Armed Services Vocational Aptitude Battery, participants are asked to match words to a numeric code, according to a key. This is a clerical task. Respondents are scored based on the speed and accuracy of their responses. Hitt and Trivitt (2013) found that coding speed is correlated with both item response rates and noncognitive ability in NLSY:97. As discussed in the literature review above, Segal (2013) found that coding speed was a plausible measure of conscientiousness.

The Educational Longitudinal Study of 2002 (ELS:02)

ELS:02 followed a nationally representative sample of over 15,000 tenth graders from 2002 through 2012. Like HSB:80 and NELS:88, the survey mode for the baseline year was a self-administered pen-and-paper survey. Similar to those surveys, “don’t know” or “unsure” were rarely offered as response options in the multiple choice questions that constitute most of the survey. We calculate a respondent’s item nonresponse rate in ELS:02 as the percentage of questions left unanswered among questions that the respondent should have answered based on responses to previous questions. ELS:02 also contains various self-reported measures of self-regulation. In particular, we use the general effort and persistence scale and the control expectations scale, which were used in the 2000 Program for International Student Assessment. These items were also field tested before use in ELS:02 as well as used in other research (Burns et al. 2003; Pintrich et al., 1993).

Summary

The surveys used in each of the six datasets above have common design features. They are supposed to be easily understandable. The pen-and-paper surveys are designed to be readable, even for students with reading skills well below grade level. The surveys are long, averaging more than 300 items, which to some students is undoubtedly boring and tedious.

We hypothesize that item response rates are driven by student motivation and effort, and not just cognitive abilities. Tables 3a through 3f show the raw order correlations between item response rates and cognitive tests, from each survey's baseline year. Response rates are, at most, only moderately correlated with cognitive ability, ranging from null to 0.21. These figures indicate that item response rates are not simply explained by cognitive ability. This alone does not mean that item response rates capture other abilities. Item response rates may largely not capture any abilities at all; they could simply be noise. Thus, in the following section, we turn to our empirical strategy, which aims to establish whether item response rates – as a measure of effort on the survey – capture information about noncognitive skills. A hallmark of noncognitive skills research has been the fact that noncognitive skills are predictive of later-life outcomes, independent of cognitive ability. We examine whether that is the case for item response rates.

≪Tables 3a through 3f Here≫

Section IV: Empirical Strategy

Empirical Models

Our study is concerned with a previously unexploited control variable for noncognitive skills. Failing to control for noncognitive abilities can be problematic when estimating human capital models. Consider the following model that specifies employment income, (Y) as a function of cognitive ability (*A*), educational attainment and work experience (***E***), and demographic and household characteristics (***H***):

$$Y_i = f(A, \mathbf{E}, \mathbf{H}; \boldsymbol{\beta}) + \nu, \qquad (1)$$

where $\boldsymbol{\beta}$ is a vector of parameters to be estimated and ν is the error term. In these models, noncognitive ability is not specified and is therefore relegated to ν. Insofar as noncognitive skills are correlated with other independent variables, insufficiently controlling for noncognitive skills leads to biased estimates of $\boldsymbol{\beta}$ (Heckman and Kautz, 2012). This is not to mention that the importance of noncognitive skills on employment income cannot be identified based on this theoretical formulation.

In our analysis, we explicitly include noncognitive skills as an independent variable in our human capital models. That is, we specify, for example, employment income as

$$Y_i = g(A^c, A^n, \mathbf{E}, \mathbf{H}, \boldsymbol{\gamma}) + \mu, \qquad (2)$$

where A^c captures cognitive ability, A^n captures noncognitive ability, $\boldsymbol{\gamma}$ is a vector of parameters to be estimated, and μ is the error term. Analogous models where employment status or educational attainment is the dependent variable can be specified as well. As discussed above, the difficulty in estimating (2) is that noncognitive skills are difficult to observe and most datasets do not have adequate measures of such skills.

Empirical Model: Educational Attainment

We use a simple empirical strategy to estimate the effect of noncognitive abilities. We begin with educational attainment as our outcome of interest. We model years of schooling as an individual utility maximization decision where the costs and benefits can vary with cognitive and noncognitive ability. The costs of schooling include tuition and foregone wages, and the opportunity costs of effort. This model also allows marginal productivity of time spent to vary with the cognitive and noncognitive abilities. We assume linearity in the parameters and estimate the following empirical model:

$$S_i = \boldsymbol{\alpha}\mathbf{X_i} + \boldsymbol{\beta}\mathbf{H_i} + \gamma_c A_i^c \; + \; \boldsymbol{\gamma_n}\mathbf{A}_i^n \; + \; \epsilon_i \qquad (3)$$

where S_i is the years of education for individual *i*. $\mathbf{X_i}$ is a vector of control variables to detect regional differences in the costs of acquiring additional education (explicit and opportunity costs). The control variables we include in $\mathbf{X_i}$ are gender, and indicator variables for birth year and census region. $\mathbf{H_i}$ is a vector of individual characteristics that influence previously accumulated human capital, expected increase in the benefits gained in the marriage market, and the benefits of household production. The specific variables included in $\mathbf{H_i}$ are the highest grade completed by household head, race, and an indicator for living in a two-parent household[2]. A^c_i is standardized observed cognitive ability, as measured by math and verbal standardized tests included in each dataset. $\mathbf{A^n_i}$ is the observed noncognitive ability of individual *i* as measured by standardized response rate as well as the scores on a variety of scales designed to measure noncognitive skills (e.g. Rotter [1966] Locus of Control Scale). Finally, ϵ_i is a normally distributed error term. Summary statistics for the number of years of education completed by respondents in each dataset are listed in Table 4.

≪Table 4 Here ≫

All equations are initially estimated using OLS with sampling weights to correct for sampling methods utilized. One assumption of model (3) is that the quantity of schooling is continuous and differentiable in the neighborhood of the optimal schooling level. This assumption may not be valid to the degree that diploma effects exist. To allow for this possibility

---

[2]To the degree that discrimination exists in labor markets or households make different investments in male and female offspring, many of our control variables could arguably be included either in X or H or both. We recognize the coefficients we estimate are reduced form, but are primarily interested in γ_n.

we also consider a model where schooling level is a discrete choice rather than continuous. In the discrete choice model the individual chooses the diploma level with the highest net benefit

$$D_s = \text{argmax}\{V_s\} \qquad (4)$$

where S options are available, V_s is the expected present value of lifetime utility for diploma s, and $s \in \{1,\ldots,S\}$.

The level of educational attainment incorporating diploma effects meets the classic example of an unobserved latent variable model that can be treated as a categorical dependent variable model. When first estimating model (4), we initially used ordered logit; however, post estimation testing showed the parallel regression assumption should be rejected at the 0.01 significance level. Therefore we estimated (4) via multinomial logit, using the same explanatory variables with six diploma levels of education: no degree, GED, high school diploma, some postsecondary education, bachelor's degree, and more than bachelor's degree.[3] The proportion of respondents who attained each level of education is shown in Table 5.

≪Table 5 Here≫

Empirical Model: Income and Employment

Turning now to employment and income as the outcome of interest, our theoretical construct is $Y_i = g(A^c, A^n, \mathbf{E}, \mathbf{H}, \boldsymbol{\gamma}) + \nu$. We estimate our income models in two ways. First we estimate equation 3 where log of employment income is the dependent variable and educational attainment is included as another independent variable. However, as in all wage estimation models, to avoid biased coefficient estimates we need to address sample selection that occurs as

[3] The only exception to this is the HSB:80 dataset which does not have a separate category for GED and more than bachelor's degree as educational attainment outcomes. Respondents to HSB are 12th graders so many of them are already on track to receive a high school diploma, while many high school dropouts and eventual GED earners are out of sample. Furthermore, the last wave of data collection for HSB occurred 6 years after the initial wave of data collection, making it uncommon to observe respondents who have already obtained a graduate degree.

some people opt out of the labor market. Thus, our second model of wages incorporating cognitive and noncognitive abilities is to use FIML to estimate the following selection model

$$\ln Y_{1i} = \boldsymbol{\gamma_{ac1}} \mathbf{A^c}_i + \boldsymbol{\gamma_{an1}} \mathbf{A^n}_i + \boldsymbol{\gamma_{e1}} \mathbf{E_i} + \boldsymbol{\gamma_{H1}} \mathbf{H}_i + \psi_{1i} \tag{5a}$$

$$Y_{2i} = \boldsymbol{\gamma_{ac2}} \mathbf{A^c}_i + \boldsymbol{\gamma_{an2}} \mathbf{A^n}_i + \boldsymbol{\gamma_{e2}} \mathbf{E_i} + \boldsymbol{\gamma_{H2}} \mathbf{H}_i + \psi_{2i} \tag{5b}$$

$$Y_{1i} = Y^*_{1i} \quad \text{if} \quad Y^*_{2i} > 0 \tag{5c}$$

$$Y_{1i} = 0 \quad \text{if} \quad Y^*_{2i} \leq 0$$

In this model equation (5b) is a probit selection equation estimating the propensity to be employed and thus an observable wage. $\mathrm{Ln}Y_{1i}$ is log of employment income for individual i. $\mathbf{A^c_i}$ is a vector of variables representing cognitive ability. $\mathbf{A^n_i}$ is a vector measuring noncognitive ability. $\mathbf{E_i}$ is a vector representing human capital in the form of formal schooling and labor market experience in quadratic form, $\mathbf{H_i}$ is a vector of household characteristics including race and gender, and ψ_{1i} and ψ_{2i} are bivariate normally distributed error terms with $\boldsymbol{\gamma}$ representing all the estimated coefficients. We estimate this model using full information maximum likelihood with the heckman command in Stata which allows us to incorporate sample weights and robust standard errors. In equation (5a) we include standardized response rate at baseline, years of experience and experience squared, years of education, race, gender, and the other controls from the education equation which vary by dataset. In equation (5b) we include all of the variables in (5a) and add current marital status, an interaction between gender and marital status, and the number of children living in the household as exclusion restrictions. In the NLSY datasets we also have a measure of spousal income and include an indicator of a spouse with an income in the upper quintile of reported spouse earnings. The exclusion restrictions we use are a common set of information available across many datasets and are basic household information commonly collected in survey data. The marital status indicates a presumed permanent presence of another adult in the household who may provide income and the and marital status-gender interaction allows for a different effect of marriage on men and women's labor market participation stemming from traditional gender expectation or discrimination in the labor market. The presence of children is expected to influence labor market participation by increasing the reservation wage when outside employment of both parents imposes child care costs on the household. Tables 6a and 6b lists the summary statistics for the respondents' employment income and employment status for each of our datasets, separated by gender. Respondents reported annual earnings.

≪Table 6 Here≫

It is also possible that noncognitive skills influence income exclusively via labor market participation. To explore this potential relationship we estimate a model where the dependent variable is the binary indicator of labor market participation. We utilize a standard probit model:

$$\text{Prob}\,(Y = 1) = \boldsymbol{\Phi}\,(\ \mathbf{b_{ac}A^{c}}_{i} + \mathbf{b_{an}A^{n}}_{i} + \mathbf{b_{e}E_{i}} + \mathbf{b_{H}H}_{i}) \qquad (6)$$

where $\mathbf{A^c}$, $\mathbf{A^n}$, $\mathbf{E}$, and $\mathbf{H}$ are all as defined in the models previously discussed, Y is an indicator of labor market participation, and $\boldsymbol{\Phi}$ is the cumulative standard normal distribution .

Section V: Results

To reiterate, our objective is to document the relationship between survey response rate and three life outcomes: educational attainment, employment, and income. All models control for our full spectrum of the respondent's baseline household and demographic characteristics: age, race, gender, household income, parent's education, single-parent household and census region. Additional controls for alternative measures of noncognitive skills and demographic characteristics such as mother's age at birth, are included when available.[4] Given this set of control variables, our results likely represent conservative estimates for the importance of noncognitive skills. Many of the variables we control for likely influence noncognitive skill formation, educational attainment and adult earnings.

Educational Attainment

Table 7 shows the estimates of our empirical models where the number of years of education is the dependent variable. All samples are restricted to observations present in our full model (column 5) as missing data is prevalent for many of our covariates.[5] As depicted in column 1, response rates are positively correlated with educational attainment across all six datasets, before including cognitive ability and survey responses on noncognitive skills. A one-standard-deviation increase in response rates is associated with attaining 0.11 to 0.33 years of additional education in this basic model, all statistically significant at the 0.05 level. Cognitive ability is a significant predictor of educational attainment in all six datasets, per column 2.

When including both response rate and cognitive ability as explanatory variables to predict educational attainment, response rate remains significant in four of the six datasets. As depicted in column 3, when significant, effect sizes range from 0.10 to 0.30 additional years of education for every one-standard-deviation increase in response rate. By comparison, a one standard deviation increase in cognitive test scores is associated with a 0.10 to 1.44 year increase

[4] In the baseline year, we use log of household income and dummy variables indicating the highest grade level of education attainment completed by the head of the household when available. Some data sets, such as ELS:2002, provided categorical instead of continuous measures of household income. In these cases, dummy variables were used to control for household income. Mother's age at birth was included for the HSB:80, NELS:88, Add Health and ELS:02. In order to give more uniform sample sizes in the NLSY:79 and NLSY:97, mother's age at birth was not used as a control variable.

[5] As it turns out, restricting the sample is a more conservative test of whether item response rates are noise or capturing something systematic. We are essentially excluding a proportion of respondents who had missing data in certain variables. That is, we are setting out to determine whether item response rates are predictive of later life outcomes, even among a sample of respondents who did not have such missing data.

in additional years of education attained. Clearly, the co-variation between item response rate and cognitive test scores influences the relationship between item response rate and educational attainment. We discuss concerns about this cause of attenuation in Section VI.

As mentioned, the specification in Column 3 contains no other noncognitive variables. We have argued that item response rates can serve as a measure of noncognitive skills, particularly in datasets that contain no other measure. The specification in Column 3 includes a set of regressors that resembles the data typically available in education program evaluations: test scores, household information, but no explicit measure of students' noncognitive skills. Notably, the NLSY97 contains no baseline-year, self-reported measure of noncognitive skills (the only baseline measure of noncognitive skills is coding speed, which is behavioral and not self-reported). Item response rate is consistently a significant predictor of educational attainment in that dataset, providing new and relevant information about participants' noncognitive skills.

Column 4 contains the model without nonresponse but with self-reported measures of noncognitive skills (or in NLSY:97, the coding speed task). In every instance, self-reported scales are predictive of educational attainment, independent of cognitive ability. Comparing Column 3 to Column 4, the addition of self-reported noncognitive skills adds relatively little to the overall R-squared, no more than 0.017 in the case of ELS:2002. Nevertheless, the coefficients for self-reported noncognitive skills are largely significant. Similarly, when comparing Column 2 to Column 3 the addition of item response rates does not substantially increase the R-squared. For self-reported and behavioral measures of noncognitive skills, this suggests that part of the effect was previously hidden within demographic control variables.

Column 5 of Table 7 displays estimates of a full model in which we include self-reported measures of noncognitive skills along item response rates. Item response rates in these models remain statistically significant in HSB, Add Health, and NLSY:97. Response rate remains positive but falls short of significance in the remaining datasets. The coefficients on self-reported noncognitive skills rarely change when including item nonresponse, when comparing Columns 4 and 5. This suggests that item nonresponse can provide additional information about noncognitive skills, rather than serving as a substitute for traditional measures. This is also consistent with our assertion that item response rates capture information not captured by self-reports. For adolescents with low item response rates, the answers on self-reported measures may be so unreliable that they constitute random noise.

≪Table 7 Here≫

Table 8 displays the results of our multinomial logistic models where educational attainment is treated as discrete categories (equation 4). Reported estimates in Table 8 are in terms of marginal effects, so each reported coefficient may be interpreted as changes in probability of attaining certain levels of education given a one standard-deviation change in response rate or cognitive ability. The model here is equivalent to that in Column 5 of Table 7, the full model. We present only the coefficients for item response rate and cognitive ability, for ease of reading.

≪Table 8 Here≫

In general, higher response rates are associated with a decreased likelihood of being a high school dropout, earning only a GED, having at most a high school diploma, or obtaining some postsecondary education but not earning a bachelor's degree. In other words, a higher response rate is associated with attaining higher levels of education. Notably, estimates from HSB:80 and NLSY:97 data suggest that the likelihood of obtaining a bachelor's degree or any post-baccalaureate degree increases with response rate. No marginal effects for obtaining a GED or a degree beyond a bachelor's degree can be calculated for HSB:80 because very few students did so in that dataset.

The multinomial models all demonstrate that increases in cognitive ability are associated with higher levels of educational attainment, as previous researchers have widely documented. Interestingly, in the NLSY:79, item response rate was not independently predictive of educational attainment in the OLS estimates, as measured continuously via years of education. However, in multinomial logistic models displayed in Table 8, item response rates are significantly associated with a decreased likelihood of dropping out of high school. Altogether, there is a visible pattern across Tables 8. Item response rate is consistently associated with a decreased likelihood of attaining lower levels of education and an increased likelihood of attaining higher levels of education.

Employment and Income

We now turn to the results for employment and income. We first examine whether respondents reported being employed during the most recent survey year. Table 9 shows probit results. These estimates test whether the association of employment with item nonresponse is independent not only of measures collected during childhood but also of educational attainment, workforce experience and marital status. We have already demonstrated that item response rates are associated with later educational attainment.

≪Table 9 Here≫

Item response rate has no additional association with employment. This is also largely true of cognitive ability. Insofar as cognitive ability and noncognitive ability impact later employment, our results suggest they do so via educational attainment.

We then turn to the question of income from employment, per Table 10. Simply regressing the log of income on the same set of covariates as above, we find again that item response rates have no additional association with employment income, except in NELS:88, where a one standard deviation in item response rates is associated a 3.5 percentage point increase in employment income. The OLS estimates, however, may be biased by nonrandom selection into the workforce.

≪Table 10 Here≫

We then estimate a selection model, as discussed in the section above. Per Table 11, the results in this specification are essentially the same as those laid out thus far. In the selection equation, item response rates are never significantly related to the likelihood of being gainfully employed. In the wage equation, item response rates are not significantly associated with increased income, except again in NELS:88. In both the OLS and selection models, educational attainment is a strong and consistent predictor of income. One additional year of formal education is associated with a 1.0 to 6.0 percentage point increase in employment income.

≪Table 11 Here≫

Section VI: Discussion and Conclusion

The importance of our findings rests upon whether we have made a convincing argument that survey response rates capture noncognitive ability. This study began by considering the perspectives of adolescents participating in a survey, who are asked to answer hundreds of boring questions about everyday life. There is strong presumption in the field of survey methodology that item nonresponse signals disinterest or disengagement in the survey process. Additional research has shown that item response rates are correlated, albeit weakly, with self-reports of conscientiousness. We have argued that, seemingly, survey completion mirrors the routine work of school, which in psychological research has consistently been linked to noncognitive skills.

Noncognitive skills have historically been defined as skills not captured by scores on cognitive tests. The raw correlations in response rates and cognitive test scores range from null to 0.21 in the datasets we examine. Clearly, most of the variation in item response rates is driven by something other than cognitive ability. This alone does not mean that item response rates capture a set of abilities and attitudes that is largely independent of cognitive skills; much of item response rates may be statistical noise. Therefore, we test whether item response rates independently predict outcomes that have a well-established relationship with both cognitive and noncognitive skills.

We find that item response rates are a significant predictor of educational attainment in every dataset, before controlling for cognitive ability. Once including cognitive test scores, the effect of item response rates attenuates, but remains significant in four of six datasets. Our results show that item response rate is not predictive of employment status, but our models include educational attainment as a control variable. Additionally, there is essentially no wage premium attributable to item response rate conditional on selection to employment and net of educational attainment, though at least the same can largely be said for cognitive ability. This is not surprising as previous work suggests that wage premiums attributable to noncognitive skills operate through the effect of noncognitive skills on educational attainment (Cawley, Heckman, and Vytlacil, 2001). According to the simple definition of noncognitive skills as “not cognitive skills,” survey response rates possess the characteristics of noncognitive skills that are related to later life outcomes.

It is worth noting that our estimates show that the effect of noncognitive skills attenuates when cognitive test scores are included.[6] Just like surveys, low-stakes cognitive tests require effort. Students showing low effort on surveys might be showing low effort on the accompanying cognitive test as well, leading to an artificially low estimate of their cognitive abilities. In using test scores, part of what we attribute to cognitive ability is simply effort on the test. From previous literature, we know with confidence that test scores are affected by student motivation and noncognitive skills (e.g. Duckworth et al. 2011, Levitt et al. 2012).

In our datasets, test scores and item response rates are moderately correlated, causing attenuation in some of our results. The implication could be that cognitive ability impacts response rates. But the correlation between cognitive tests and item response rates could just as easily indicate the opposite: item response rates partly capture the motivation (or lack of motivation) of students to complete mundane tasks, including low stakes tests.

Interestingly, survey item response rates could serve as a measure of student motivation on standardized tests. Nascent work on this topic was begun over a decade ago in an unpublished manuscript by Boe, May and Baruch (2002) which examined the relationship between student scores on the Trends International Mathematics and Science Study and item response rates on a corresponding survey. Our findings strongly suggest that such work should be revisited.

Item response rates are admittedly a messy measure of noncognitive skills such as student effort. But even a cleaner measure of student effort, if used as a predictor of later outcomes, would

[6] Another source of attenuation in our estimates is the inclusion of demographic and human capital variables in our regression models. While this attenuation makes it difficult to measure the impact of noncognitive skills on later outcomes, it also illustrates that some of the effect attributed to demographic factors is associated with specific behaviors or noncognitive skills. In future research, we intend to explore whether item response rates help explain achievement, educational attainment and employment differences across different student groups.

suffer from attenuation when cognitive test scores were included – because test scores themselves are affected by student effort.

In both of our educational attainment analyses, the statistical significance of item response rate is rarely influenced by in the inclusion of self-reported noncognitive skills. Conversely, the estimates of self-reported noncognitive skills rarely attenuate substantially upon the inclusion of item response rates. The exception to this pattern is ELS:02. Perhaps item response rate measures a set of conscientious behaviors distinct from those that self-reported scales are designed to measure. Or it could mean that item response rate measures noncognitive skills similar to what the scales were designed to capture, and that item response rate contains information from respondents whose self-reports were essentially just noise, due to a lack of attention to the survey. Ultimately, in future research, survey effort should be compared to performance on other tasks or to third-party skills assessments. [7]

We have posited that the skills required for completing a survey are correlated with the skills needed to advance in school and to get a job, many of which go unobserved. We have made a conceptual case that item response rates capture a minimal competency on these routine tasks. That said, we must also acknowledge that item nonresponse is a noisy measure with ceiling effects.

When no other measure of noncognitive skills is available, item response rate can serve as a stand-alone proxy. Item response rate is undoubtedly a noisy measure. Estimates based

---

[7] This is a topic for future research, where the data and methods of psychologists and experimental economists are of considerable value. Under laboratory conditions, it has been shown that financial incentives and fatiguing exercises have temporarily altered a person's observed self-control or conscientiousness (Hagger et al. 2010; McGee and McGee 2011; Segal 2012). Similar experiments could be conducted on survey effort. Evidence from field experiments would also be instructive. Experimental programs have been shown to improve student study habits and focus in school; it would be instructive to learn whether treatment effects also existed on measures of survey effort. Such research could provide considerable insight into what psychological constructs in particular underlie survey effort.

solely on item response rate will be prone to false negatives. Relying on it as the sole noncognitive measure is not advisable, but sometimes the data give no other choice. The NLSY:97, for example, contains no self-reported noncognitive skills in the baseline year.
As a primary explanatory variable, item response rate is of course limited in value. This is true of any single measure of noncognitive skills, including short, self-reported scales. For this reason, it's common for researchers to build composite indices of noncognitive skills (e.g. Heckman, Stixrud and Urzua 2006). Our results suggest that item response rates could be included in such composite measures.

In future research, we will explore how item response rates can be combined with other measures to form stronger, more precise measures of noncognitive skills, whether these measures of noncognitive skills serve as key independent variables of interest or as control variables. It is possible that the inclusion of item response rate as a control variable, when no other noncognitive skill measures are available, could alter estimates of other variables of interest. It is our hope that other researchers join this effort. The object of this paper is to demonstrate that item response rates, and other measures of survey effort, are worthy of further attention. Given a number of challenges, we cannot yet map a precise quantitative relationship between item nonresponse and the skills and everyday behaviors that ultimately shape later outcomes, though such a mapping would be informative and useful for research.

One reason why it is difficult, across datasets, to identify a constant quantitative relationship between item response and other variables is that different surveys use different instruments. We view surveys as tasks. But different surveys, to a degree, represent tasks of different difficulty. All of the surveys we have evaluated are long and tedious, but some are longer than others. The surveys vary in mode, and they vary in subject matter. The age of the respondent varies as well; it is possible that a long survey at age 14 may pose a very different challenge than a short survey at age 17. So, while survey item response requires conscientious effort, different surveys to different adolescents likely require different levels of conscientious effort. These challenges regarding comparability across datasets are not unfamiliar to standard measures of cognitive and noncognitive skills. For example, in our educational attainment and employment estimates, the magnitude of the relationship between cognitive ability and later outcomes varies substantially. The same is true for self-reported noncognitive skills. This is likely due to a number of factors, including the fact that the cognitive and noncognitive measurement instruments are not identical across surveys.

Another reason the relationship between survey effort and conscientious behavior is difficult to map is that conscientious behaviors are not thoroughly documented in our large scale datasets. That is why we cannot yet say that, given a student's level of survey effort, she has a specific probability of completing her homework or finishing a job application – we don't observe students in everyday situations. Understanding the exact relationship between survey effort and true noncognitive abilities is a task for future research. The great value of psychological research is the rich assortment of variables collected, many of which come from third-party observations. So psychologists are perhaps best suited to delve further into the question of what exactly item nonresponse and other measures of survey effort are truly capturing. Absent such research, we have resorted to making a conceptual case. Other than basic reading skills, a respondent needs to pay attention, respect the instructions, and put a little effort into recalling facts about their everyday lives. These skills, by definition, resemble the behaviors associated with conscientiousness.

Our paper makes three important contributions. Primarily, it establishes that response rates capture noncognitive skills that are important to future educational attainment, which ultimately affects other longer-run outcomes, such as labor-market outcomes (Cawley, Heckman, & Vytlcil, 2001). While self-reported measures of noncognitive skills may show what attitudes and character traits are associated with those outcomes, our measure is behavioral. Self-reported noncognitive measures tell us that people who say that they have higher noncognitive skills on balance do better in life. Our findings provide further clues into how people with higher educational attainment behave: they complete mundane tasks given to them by relative strangers in positions of authority, even if the immediate incentive to complete that task is unclear.

Second, the noncognitive variable that we validate can be used in hundreds of existing datasets that do not contain better measures of noncognitive skills. The information captured by item response rates can be used to evaluate the impact of certain policies on those skills. Moreover, even in datasets with explicit measures of noncognitive skills, item response rates do not suffer from the problems of reference group bias and satisficing that plague those measures. That said, as with other measures of noncognitive skills, it should also be noted that this measure has limited viability as a way to evaluate noncognitive skills in data collected in high stakes evaluations, especially in cases where participants would be aware that item response rate is a performance measure. It is also worth noting that recent digital survey designs that force respondents to answer all questions before they can proceed to the next section are eliminating this latent noncognitive skill measure in many datasets — which may incidentally introduce measurement error by generating forced, careless answers.

Third, and perhaps most importantly, our findings show the benefit of thinking more creatively about the data used in economics and education research. In our case, we examine long surveys completed by adolescents. Item response rates are a latent source of data that has been available for decades, but missing answers have been treated simply as measurement errors – even though it has long been understood that item nonresponse is not random. If simple item nonresponse can be shown to be a measurement of other noncognitive skills, then social scientists and psychometricians should begin to explore other latent measures of noncognitive skills that are perhaps more difficult to measure.

The field of economics has made crucial contributions to the understanding of noncognitive skills' importance to education, employment and well-being. The single greatest challenge faced by this research program is the omission of noncognitive measures from key datasets. Discovering and exploiting new and latent measures of noncognitive skills will only enhance future noncognitive skills research. This is what we have set out to do.

Two decades ago, noncognitive skills were "dark matter," relegated to the residual in economic models (Heckman & Rubenstein, 2001, p. 149). Bit by bit, research has rescued noncognitive skills from the error term. In another incremental step, our research brings the role of noncognitive skills into clearer view.

References

Almlund, Mathilde, Angela Lee Duckworth, James J. Heckman, and Tim D. Kautz. 2011. *Personality Psychology and Economics*. National Bureau of Economic Research. Retrieved from: http://www.nber.org/papers/w16822.

Boe, Erling E., Henry May, and Robert F. Boruch. 2002. "Student Task Persistence in the Third International Mathematics and Science Study: A Major Source of Acheievement Differences at the National, Classroom, and Student Levels." *Center for Research and Evaluation in Social Policy, CRESP-RR-2002-TIMSS1*. Retrieved from: http://eric.ed.gov/?id=ED478493.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas Ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43 (4): 972–1059.

Borghans, Lex, Bas Ter Weel, and Bruce A. Weinberg. 2008. "Interpersonal Styles and Labor Market Outcomes." *Journal of Human Resources* 43 (4): 815–58.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*, 1137–76.

Burns, L. J., R. Heuer, S. J. Ingels, J. Pollack, D. J. Pratt, D. Rock, and others. 2003. "Education Longitudinal Study of 2002 Base Year Field Test Report (No. NCES 2003-03)." *Washington, DC: National Center for Education Statistics*.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee. 2015. "Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan." Accessed May 7. http://www.cepr.org/sites/default/files/Callen%20-%20mm_personalities_2sep2014.pdf.

Carneiro, Pedro, Claire Crawford, and Alissa Goodman. 2007. *The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes*. Centre for the Economics of Education, LSE. Retrieved from: http://ideas.repec.org/p/cep/ceedps/0092.html.

Cawley, John, James Heckman, and Edward Vytlacil. (2001). Three observations on wages and measured cognitive ability. *Labour Economics* 8 (4): 419-442.

Chernyshenko, Oleksandr Sergeevich. 2003. "Applications of Ideal Point Approaches to Scale Construction and Scoring in Personality Measurement: The Development of a Six-Faceted Measure of Conscientiousness." Retrieved from: http://psycnet.apa.org/psycinfo/2003-95010-007.

Chingos, Matthew M., and Paul E. Peterson. 2012. "The Effects of School Vouchers on College Enrollment: Experimental Evidence from New York City." *Brookings Institution*. Retrieved from: http://eric.ed.gov/?id=ED534618.

Costa, Paul T., and Robert R. McCrae. 2008. "The Revised Neo Personality Inventory (neo-Pi-R)." *The SAGE Handbook of Personality Theory and Assessment* 2: 179–98.

Coughlin, Conor, and Carolina Castilla. 2014. "The Effect of Private High School Education on the College Trajectory." *Economics Letters*. http://www.sciencedirect.com/science/article/pii/S0165176514003383.

Credé, Marcus, Sylvia G. Roch, and Urszula M. Kieszczynka. 2010. "Class Attendance in College A Meta-Analytic Review of the Relationship of Class Attendance With Grades and Student Characteristics." *Review of Educational Research* 80 (2): 272–95.

Cunha, Flavio, and James J. Heckman. 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Journal of Human*

Resources 43 (4): 738–82.

Dalal, Reeshad S. 2005. “A Meta-Analysis of the Relationship between Organizational Citizenship Behavior and Counterproductive Work Behavior.” *Journal of Applied Psychology* 90 (6): 1241.

Deke, John, and Joshua Haimson. 2006. “Valuing Student Competencies: Which Ones Predict Postsecondary Educational Attainment and Earnings, and for Whom? Final Report.” *Mathematica Policy Research, Inc.* http://eric.ed.gov/?id=ED496390.

Dobbie, Will, and Jr Roland G. Fryer. 2013. *The Medium-Term Impacts of High-Achieving Charter Schools on Non-Test Score Outcomes*. Working Paper 19581. National Bureau of Economic Research. Retrieved from: http://www.nber.org/papers/w19581.

Duckworth, Angela Lee, Patrick D. Quinn, Donald R. Lynam, Rolf Loeber, and Magda Stouthamer-Loeber. 2011. “Role of Test Motivation in Intelligence Testing.” *Proceedings of the National Academy of Sciences* 108 (19): 7716–20.

Duncan, Greg J., and Katherine Magnuson. 2013. “Investing in Preschool Programs.” *The Journal of Economic Perspectives* 27 (2): 109–32. doi:10.1257/jep.27.2.109.

Hagger, Martin S., Chantelle Wood, Chris Stiff, and Nikos LD Chatzisarantis. "Ego depletion and the strength model of self-control: a meta-analysis." *Psychological bulletin* 136, no. 4 (2010): 495.

Heckman, James J. 2000. “Policies to Foster Human Capital.” *Research in Economics* 54 (1): 3–56.

Heckman, James J., and Tim Kautz. 2012. “Hard Evidence on Soft Skills.” *Labour Economics* 19 (4): 451–64.

Heckman, James J., and Yona Rubinstein. 2001. “The Importance of Noncognitive Skills:

Lessons from the GED Testing Program." *American Economic Review*, 145–49.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–82.

Hedengren, David and Thomas Strattman (2012). "The Dog that Didn't Bark: What Item Nonresponse Shows about Cognitive and Noncognitive Ability." Unpublished Manuscript. Retrieved from: http://ssrn.com/abstract=2194373

Hitt, Collin, and Julie Trivitt. "Don't Know? Or Don't Care? Predicting Educational Attainment Using Survey Response Rates and Coding Speed Tests as Measures of Conscientiousness." *EDRE Working Paper 2013-05. University of Arkansas*.

Hogan, Robert, and Hogan, Joyce. *Hogan Personality Inventory Manual, Third Edition*. Tulsa, OK: Hogan Assessment Systems.

Jackson, Joshua J., Dustin Wood, Tim Bogg, Kate E. Walton, Peter D. Harms, and Brent W. Roberts. 2010. "What Do Conscientious People Do? Development and Validation of the Behavioral Indicators of Conscientiousness (BIC)." *Journal of Research in Personality* 44 (4): 501–11.

Kaestner, Robert, and Kevin Callison. 2011. "Adolescent Cognitive and Noncognitive Correlates of Adult Health." *Journal of Human Capital* 5 (1): 29–69.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 1998. "List-Wise Deletion Is Evil: What to Do about Missing Data in Political Science." In *Annual Meeting of the American Political Science Association, Boston*.

King, Gary, and Jonathan Wand. 2007. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes." *Political Analysis* 15 (1): 46–66.

Krosnick, Jon A. 1991. “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology* 5 (3): 213–36.

Krosnick, Jon A., Sowmya Narayan, and Wendy R. Smith. 1996. “Satisficing in Surveys: Initial Evidence.” *New Directions for Evaluation* 1996 (70): 29–44.

Krosnick, Jon A., and Stanley Presser. 2010. “Question and Questionnaire Design.” *Handbook of Survey Research* 2: 263–314.

Levitt, Steven, John List, Susanne Neckerman and Sally Sadoff. 2012. “The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance.” *NBER Working Paper* 18165.

Lindqvist, Erik, and Roine Vestman. 2011. “The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment.” *American Economic Journal: Applied Economics*, 101–28.

Lleras, Christy. 2008. “Do Skills and Behaviors in High School Matter? The Contribution of Noncognitive Factors in Explaining Differences in Educational Attainment and Earnings.” *Social Science Research* 37 (3): 888–902.

Lubotsky, Darren, and Robert Kaestner. 2014. “Do ‘Skills Beget Skills’? New Evidence on Dynamic Complementarities in Early Grades.” Working Paper.

Lundborg, Petter, Paul Nystedt, and Dan-Olof Rooth. 2014. “Height and Earnings: The Role of Cognitive and Noncognitive Skills.” *Journal of Human Resources* 49 (1): 141–66.

MacCann, Carolyn, Angela Lee Duckworth, and Richard D. Roberts. 2009. “Empirical Identification of the Major Facets of Conscientiousness.” *Learning and Individual Differences* 19 (4): 451–58.

McGee, Andrew, and Peter McGee. 2011. *Search, Effort, and Locus of Control*. Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: http://www.econstor.eu/handle/10419/55119.

Morris, Rosenberg. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Mueller, Gerrit, and Erik Plug. 2006. “Estimating the Effect of Personality on Male and Female Earnings.” *Industrial and Labor Relations Review*, 3–22.

Pintrich, Paul R., David AF Smith, Teresa García, and Wilbert J. McKeachie. 1993. “Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ).” *Educational and Psychological Measurement* 53 (3): 801–13.

Poropat, Arthur E. 2009. “A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance.” *Psychological Bulletin* 135 (2): 322.

Puhani, Patrick. 2000. “The Heckman Correction for Sample Selection and it Critique.” *Journal of Economic Surveys*, 14(1), 53-68.

Roberts, Brent W., Peter D. Harms, Avshalom Caspi, and Terri E. Moffitt. 2007. “Predicting the Counterproductive Employee in a Child-to-Adult Prospective Study.” *Journal of Applied Psychology* 92 (5): 1427–36.

Rosenberg, Morris. 1965. “Rosenberg Self-Esteem Scale (RSE).” *Acceptance and Commitment Therapy. Measures Package* 61.

Rotter, Julian B. 1966. “Generalized Expectancies for Internal versus External Control of Reinforcement.” *Psychological Monographs: General and Applied* 80 (1): 1.

Segal, Carmit. 2012. “Working When No One Is Watching: Motivation, Test Scores, and Economic Success.” *Management Science* 58 (8): 1438–57.

Segal, Carmit. 2013. “Misbehavior, Education, and Labor Market Outcomes.” *Journal of the*

European Economic Association 11 (4): 743–79.

Singer, Eleanor, and Cong Ye. 2013. "The Use and Effects of Incentives in Surveys." *The ANNALS of the American Academy of Political and Social Science* 645 (1): 112–41.

Smith, Tom W. 1982. "Educated Don't Knows: An Analysis of the Relationship between Education and Item Nonresponse." *Political Methodology*, 47–57.

Smith, Tom W. 1995. "Little Things Matter: A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses." In *Proceedings of the American Statistical Association, Survey Research Methods Section*, 1046–51. American Statistical Association Alexandria VA.

Toburen, Tina, and Brian P. Meier. 2010. "Priming God-Related Concepts Increases Anxiety and Task Persistence." *Journal of Social and Clinical Psychology* 29 (2): 127–43.

Trautwein, Ulrich, Oliver Lüdtke, Inge Schnyder, and Alois Niggli. 2006. "Predicting Homework Effort: Support for a Domain-Specific, Multilevel Homework Model." *Journal of Educational Psychology* 98 (2): 438.

Tsukayama, Eli, Angela Lee Duckworth, and Betty Kim. 2013. "Domain-Specific Impulsivity in School-Age Children." *Developmental Science* 16 (6): 879–93.

West, Martin R., Matthew A. Kraft, Amy S. Finn, Rebecca Martin, Angela L. Duckworth, Christopher FO Gabrieli, and John DE Gabrieli. 2014. "Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling." Working Paper. Retrieved from: http://www.gse.harvard.edu/cepr-resources/files/news-events/cepr-promise-paradox.pdf.

Table 1: Datasets

Dataset	Years of Data Collection	Respondent Age Range at Initial Year of Data Collection	National-Representativeness	Measure of Cognitive Ability
NLSY:79	1979 to 1992	14 to 22	Adolescents who were 14 to 22 as of December 31, 1978	Armed Forces Qualification Test (AFQT) Percentile
HSB:80	1980 to 1986	15 to 21	Twelfth-grade students in public and private schools during the 1979-1980 school year	Scores on standardized tests of math, reading, and vocabulary
NELS:88	1988 to 2000	12 to 15	Eighth-grade students in public and private schools during the 1987-1988 school year	Scores on standardized tests of math and reading
Add Health	1994 to 2008	10 to 19	Seventh- through twelfth-grade students in public and private schools during the 1994-1995 school year	Scores on an abridged version of Peabody Picture Vocabulary Test
NLSY:97	1997 to 2010	12 to 16	Adolescents who were 12 to 16 years old as of December 31, 1996	Armed Services Vocational Aptitude Battery (ASVAB) Math and Verbal Percentile
ELS:02	2002 to 2012	14 to 19	Tenth-grade students in public and private schools during the 2001-2002 school year.	Scores on standardized tests in math and reading

Table 2: Summary Statistics for Item Response Rate

	Observations	Mode of Survey	Item Response Rate					Questions Faced			
			Avg. %	SD	Min.	Max.	"Perfect"	Avg.	SD	Min.	Max.
NLSY:79	8,230	live interview	99.72	0.43	87.32	100	36.68	750.41	64.50	603	1,094
HSB:80	6,073	pen and paper	96.44	5.88	46.09	100	14.89	370.03	9.58	343	375
NELS:88	9,989	pen and paper	97.10	7.21	17.04	100	38.69	320.00	0.00	320	320
Add Health	2,458	pen and paper	94.86	14.51	6.73	100	54.47	97.16	2.88	87	105
NLSY:97	5,158	live interview	99.01	1.98	56.22	100	41.28	227.90	56.51	114	656
ELS:02	7,150	pen and paper	97.05	4.92	28.49	100	14.17	350.42	8.04	309	381

Note: Summary statistics are presented for the sample present in the full educational attainment model. The column marked "Perfect" indicates the percentage of students with item response rates of 100 percent. For NELS:88, some respondents were routed to additional questions based on answers to previous questions. A substantial portion of the optional questions are targeted at students whose parents are foreign-born or speak a language other than English. Item nonresponse to these questions is plausibly impacted by factors other than effort on the survey. We therefore excluded optional items on NELS:88 from our analysis. The number of observations for ELS:02 is rounded to the nearest ten per restricted data-use license agreements.

Table 3a: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in NLSY79.

	Cognitive Ability	Item Response Rates	Locus of Control
Cognitive Ability	-		
Item Response Rates	.1894 (0.000)	-	
Locus of Control	-0.308 (0.000)	-0.091 (0.000)	-

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N = 8,230

Table 3b: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in HSB80.

	Cognitive Ability	Item Response Rates	Locus of Control	Self-Concept
Cognitive Ability	-			
Item Response Rates	-0.021 (0.104)	-		
Locus of Control	0.013 (0.324)	0.103 (0.000)	-	
Self-Esteem	0.012 (0.342)	0.055 (0.000)	0.205 (0.000)	-

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N = 6,073.

Table 3c: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in NELS:88.

	Cognitive Ability	Item Response Rates	Locus of Control	Self-Concept
Cognitive Ability	-			
Item Response Rates	0.205 (0.000)	-		
Locus of Control	0.312 (0.000)	0.087 (0.000)	-	
Self-Concept	0.157 (0.000)	0.033 (0.001)	0.537 (0.000)	-

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N =9,989

Table 3d: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in Add Health.

	Cognitive Ability	Item Response Rates	Self-Esteem
Cognitive Ability	-		
Item Response Rates	0.209 (0.00)	-	
Self-Esteem	-0.037 (0.07)	-0.012 (0.55)	-

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N = 2,458.

Table 3e: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in NLSY:97.

	Cognitive Ability	Item Response Rates	Coding Speed
Cognitive Ability			
Item Response Rates	0.101 (0.000)		
Coding Speed	0.523 (0.000)	0.034 (0.013)	

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N =5,158

Table 3f: Correlations between Cognitive Ability, Item Response Rates, and Self-Reported Noncognitive Ability in ELS:02.

	Cognitive Ability	Item Response Rates	Control Expectations	General Effort/ Persistence
Cognitive Ability	-			
Item Response Rates	0.186 (0.000)	-		
Control Expectations	0.319 (0.000)	0.120 (0.000)	-	
General Effort/ Persistence	0.206 (0.000)	0.100 (0.000)	0.722 (0.000)	-

Note: Pearson product moment correlation shown. P-values are shown in parenthesis. N = 7,150. The number of observations is rounded to the nearest ten per restricted data-use license agreements.

Table 4: Summary Statistics for Years of Education

	Average	Standard Deviation	Minimum	Maximum	Outcome Year
NLSY:79	12.92	2.39	0	20	1992
HSB:80	13.19	1.67	11	18	1986
NELS:88	14.24	1.85	10	20	2000
Add Health	14.60	2.12	8	20	2008
NLSY:97	13.52	2.81	5	20	2010
ELS:02	14.61	1.96	10	20	2012

Note: In NELS:88 and ELS:02, years of education were imputed based on reports of highest degree completed. Dropouts were coded as 10 in NELS:88 and 11 in ELS:02, where baseline students were in the 8th grade and 10th grade, respectively. GED recipients and HS graduates were coded as 12, two-year college graduates as 14, four-year college graduates as 16, masters degree holders as 18, and higher graduate degree holders as 20. The minimum and maximum values for ELS:02 are rounded to the nearest ten per restricted data-use license agreements.

Table 5: Summary Statistics for Educational Attainment Level

	Less than High School	GED	High School Diploma	Some Postsecondary Education	Bachelor's Degree	More than Bachelor's Degree
NLSY:79	14.51	7.95	35.94	22.88	12.06	6.67
HSB:80	0.36	n/a	57.09	20.09	22.46	n/a
NELS:88	6.45	3.35	12.61	44.82	29.12	
Add Health	4.23	3.13	12.53	42.07	26.69	11.35
NLSY:97	9.94	10.52	46.10	7.24	21.50	4.70
ELS:02	1.69	1.11	6.95	47.07	32.82	10.35

Note: All numbers are percentages. Because Respondents in HSB were in 12th grade during the baseline year, almost all are on track to graduate; high school dropouts and those earning a GED are rare. Likewise, the final year of data collection was 6 years after high school graduation, making it rare for respondents to earn a post-baccalaureate degree.

Table 6a: Summary Statistics for Employment Income among Males

Males							
	N	Average ($)	SD ($)	Minimum ($)	Maximum ($)	Percent of Sample Employed (%)	Outcome Year
NLSY:79	3,716	25,364	17,915	20	90,325	82.93	1992
NELS:88	4,497	30,979	21,634	13	500,000	97.28	1999
Add Health	993	46,510	58,601	7	999,995	92.61	2008
NLSY:97	2,783	35,261	25,293	0	130,254	73.92	2010
ELS:02	2,970	34,185	26,424	0	500,000	93.96	2012

Table 6b: Summary Statistics for Employment Income among Females

Females							
	N	Average ($)	SD ($)	Minimum ($)	Maximum ($)	Percent of Sample Employed (%)	Outcome Year
NLSY:79	3,259	17,891	13,268	30	90,325	71.86	1992
NELS:88	4,589	22,897	14,384	200	500,000	94.91	1999
Add Health	1,105	33,465	35,278	4	500,000	86.68	2008
NLSY:97	2475	27,877	19,971	0	130,254	66.64	2010
ELS:02	3,460	27,649	20,929	0	330,000	92.87	2012

Note: Summary statistics restricted to panel participants who were employed. NLSY79 & 97 truncated income to mean of upper 2% for respondents with income at 98^{th} percentile or higher for summary statistics. The number of observations, minimum values, and maximum values for ELS:02 are rounded to the nearest ten per restricted data-use license agreements.

Table 7: OLS Results for Years of Education

	(1)	(2)	(3)	(4)	(5)
NLSY:79 (N=8,230)					
Item Response Rate	0.134***		0.010		0.007
	(0.034)		(0.026)		(0.027)
Cognitive Ability		1.343***	1.342***	1.314***	1.313***
		(0.036)	(0.036)	(0.037)	(0.037)
Locus of Control				0.103***	0.103***
				(0.024)	(0.024)
R^2	0.290	0.482	0.482	0.483	0.483
HSB:80 (N = 6,073)					
Item Response Rate	0.291***		0.292***		0.269***
	(0.040)		(0.040)		(0.040)
Cognitive Ability		0.096**	0.096***	0.091**	0.092**
		(0.037)	(0.037)	(0.036)	(0.036)
Locus of Control				0.106***	0.097***
				(0.030)	(0.030)
Self-Esteem				0.107***	0.102***
				(0.027)	(0.029)
R^2	0.108	0.103	0.110	0.111	0.118
NELS:88 (N=9,989)					
Item Response Rate	0.107***		0.025		0.020
	(0.031)		(0.031)		(0.031)
Cognitive Ability		0.597***	0.594***	0.547***	0.545***
		(0.031)	(0.031)	(0.030)	(0.030)
Locus of Control				0.125***	0.125***
				(0.029)	(0.029)
Self-Concept				0.089***	0.089***
				(0.026)	(0.026)
R^2	0.332	0.402	0.402	0.411	0.411
Add Health (N=2,458)					
Item Response Rate	0.215***		0.144***		0.141***
	(0.042)		(0.043)		(0.043)
Cognitive Ability		0.519***	0.499***	0.528***	0.508***
		(0.059)	(0.059)	(0.059)	(0.059)
Self-esteem				0.126***	0.124***
				(0.045)	(0.044)
R^2	0.251	0.287	0.290	0.291	0.294
NLSY:97 (N=5,158)					
Item Response Rate	0.287***		0.139***		0.134***
	(0.048)		(0.045)		(0.045)
Cognitive Ability		1.444***	1.433***	1.353***	1.344***
		(0.039)	(0.039)	(0.045)	(0.045)
Coding Speed				0.177***	0.173***
				(0.043)	(0.043)
R^2	0.129	0.331	0.332	0.333	0.334

ELS:02 (N=7,150)					
Item Response Rate	0.325***		0.098*		0.033
	(0.057)		(0.053)		(0.054)
Cognitive Ability		0.726***	0.720***	0.642***	0.640***
		(0.029)	(0.029)	(0.030)	(0.030)
Control Expectations				0.116***	0.115***
				(0.036)	(0.036)
General Effort/ Persistence				0.170***	0.169***
				(0.036)	(0.036)
R^2	0.194	0.278	0.278	0.295	0.295

Notes: All independent variables are standardized. All models control for respondent's household and demographic characteristics. In NELS:88, ELS:02, Add Health, and HSB, years of education were imputed based upon highest degree attained. Such imputation may make the data left-censored and warrant Tobit regressions. However, results do not change whether one uses Tobit or OLS, so we report OLS estimates for simplicity. The number of observations for ELS:02 is rounded to the nearest ten per restricted data-use license agreements. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Multinomial Logistic Regression Marginal Effects for Educational Attainment

	High School Dropout	GED	High School Diploma	Some Postsecondary Education	Bachelor's Degree	More than Bachelor's Degree
NLSY:79						
Item Response Rate	-0.004** (0.002)	0.000 (0.004)	0.019* (0.011)	-0.010 (0.009)	-0.004 (0.004)	-0.001 (0.002)
Cognitive Ability	-0.052*** (0.005)	-0.030*** (0.005)	-0.243*** (0.015)	0.121*** (0.013)	0.141*** (0.008)	0.064*** (0.005)
HSB:80						
Item Response Rate	0.000 (0.000)	n/a	-0.084*** (0.050)	-0.029** (0.014)	0.113*** (0.020)	n/a
Cognitive Ability	-0.000 (0.000)	n/a	-0.027** (0.011)	0.008 (0.008)	0.019*** (0.007)	n/a
NELS:88						
Item Response Rate	-0.000 (0.000)	-0.000 (0.001)	-0.005 (0.006)	-0.004 (0.001)	0.007 (0.010)	0.002 (0.002)
Cognitive Ability	-0.008*** (0.001)	-0.005*** (0.001)	-0.051*** (0.001)	-0.090*** (0.011)	0.138*** (0.010)	0.016*** (0.001)
Add Health						
Item Response Rate	-0.000 (0.00)	-0.000*** (0.000)	-0.028*** (0.010)	-0.017 (0.020)	-0.006 (0.019)	0.028 (0.013)
Cognitive Ability	-0.001*** (0.000)	-0.000*** (0.000)	-0.070*** (0.012)	-0.037** (0.019)	0.072*** (0.017)	0.036*** (0.009)
NLSY:97						
Item Response Rate	-0.007*** (0.003)	0.002 (0.005)	-0.032*** (0.012)	0.007 (0.008)	0.025** (0.011)	0.005 (0.004)
Cognitive Ability	-0.058*** (0.004)	-0.055*** (0.005)	-0.095*** (0.011)	0.016*** (0.006)	0.156*** (0.009)	0.036*** (0.003)
ELS:02						
Item Response Rate	-0.001 (0.003)	-0.000 (0.003)	-0.012 (0.007)	0.010 (0.025)	-0.001 (0.024)	0.004 (0.012)
Cognitive Ability	-0.007 (0.024)	0.006 (0.015)	-0.043*** (0.006)	-0.142*** (0.028)	0.150*** (0.012)	0.048*** (0.005)

Notes: All independent variables are standardized. Coefficients are marginal effects holding all other variables at their mean. All models control for respondent's household and demographic characteristics. Based on a Wald Test, we are able to reject the null hypothesis that the coefficients are jointly equal to zero at 0.05 significance level for datasets where we observed a statistically significant coefficient on item response rate.*** p<0.01, ** p<0.05, * p<0.1

Table 9: Probit Results for Employment Status

	(1)	(2)	(3)	(4)	(5)
NLSY:79 (N=5,353)					
Item Response Rate	0.006 (0.038)		0.007 (0.033)		0.006 (0.038)
Cognitive Ability		-0.007 (0.046)	-0.008 (0.047)	-0.007 (0.047)	-0.008 (0.047)
Rotter Locus of Control				-0.007 (0.034)	-0.007 (0.034)
Years of Education	0.031* (0.015)	0.034* (0.017)	0.033* (0.017)	0.033* (0.017)	0.032* (0.017)
R^2					
NELS:88 (N= 9,091)					
Item Response Rate	0.001 (0.002)		0.001 (0.002)		0.001 (0.002)
Cognitive Ability		0.002 (0.002)	0.002 (0.002)	0.002 (0.003)	0.002 (0.003)
Locus of Control				0.000 (0.002)	0.000 (0.002)
Self-Concept				0.001 (0.002)	0.001 (0.002)
Years of Education	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)
R^2					
Add Health (N = 2,395)					
Item Response Rate	0.004 (0.007)		0.003 (0.007)		0.003 (0.007)
Cognitive Ability		0.008 (0.008)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)
Self-Esteem				0.005 (0.006)	0.005 (0.006)
Years of Education	0.010*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.009*** (0.003)
NLSY:97(N=2,625)					
Item Response Rate	-0.017 (0.045)		-0.013 (0.045)		-0.013 (0.045)
Cognitive Ability		-0.110* (0.059)	-0.109* (0.059)	-0.118* (0.063)	-0.117* (0.054)
Coding Speed				0.016 (0.054)	0.016 (0.054)
Years of Education	0.016 (0.017)	0.032* (0.019)	0.033* (0.019)	0.032* (0.019)	0.032* (0.019)

ELS:02 (N= 6,700)					
Item Response Rate	-0.060 (0.074)		-0.070 (0.075)		-0.081 (0.076)
Cognitive Ability		0.032 (0.036)	0.037 (0.037)	0.047 (0.037)	0.052 (0.038)
Control Expectations				-0.124*** (0.041)	-0.123*** (0.041)
General Effort and Persistence				0.144*** (0.040)	0.147*** (0.040)
Years of Education	0.077*** (0.016)	0.072*** (0.017)	0.072*** (0.017)	0.070*** (0.017)	0.070*** (0.017)

Notes: All independent variables are standardized, except years of education, where the unit of measure is a single year of education completed. Coefficients are marginal effects holding all other variables at their mean. Regressions restricted to panel participants who were employed. All models control for respondent's household and demographic characteristics. The number of observations for ELS:02 is rounded to the nearest ten per restricted data-use license agreements *** p<0.01, ** p<0.05, * p<0.10

Table 10: OLS Results for Log of Employment Income

	(1)	(2)	(3)	(4)	(5)
NLSY:79 (N= 4,280)					
Item Response Rate	0.035		0.028		0.028
	(0.023)		(0.024)		(0.024)
Cognitive Ability		0.126***	0.124***	0.123***	0.121***
		(0.025)	(0.025)	(0.025)	(0.025)
Rotter Locus of Control				-0.012	-0.011
				(0.015)	(0.015)
Years of Education	0.125***	0.102***	0.101***	0.101***	0.101***
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
R^2	0.348	0.355	0.356	0.355	0.356
NELS:88 (N=8,496)					
Item Response Rate	0.038***		0.035**		0.035**
	(0.014)		(0.014)		(0.014)
Cognitive Ability		0.023**	0.018*	0.013	0.008
		(0.022)	0.011	(0.012)	(0.011)
Locus of Control				0.031**	0.031***
				(0.012)	(0.012)
Self-Concept				0.020*	0.019*
				(0.011)	(0.011)
Years of Education	0.065***	0.063***	0.062***	0.059***	0.059***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
R^2	0.381	0.380	0.381	0.383	0.384
Add Health (N=2,098)					
Item Response Rate	-0.008		-0.012		-0.013
	(0.020)		(0.020)		(0.020)
Cognitive Ability		0.045	0.046	0.049	0.051
		(0.032)	(0.032)	(0.032)	(0.032)
Self-esteem				0.043**	0.043**
				(0.021)	(0.022)
Years of Education	0.111***	0.106***	0.106***	0.104***	0.104***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
R^2	0.147	0.149	0.149	0.151	0.151
NLSY:97 (N=4,187)					
Item Response Rate	0.017		0.011		0.011
	(0.022)		(0.022)		(0.022)
Cognitive Ability		0.128***	0.128***	0.110***	0.110***
		(0.023)	(0.023)	(0.024)	(0.024)
Coding Speed				0.038*	0.038*
				(0.021)	(0.021)
Years of Education	0.099***	0.080***	0.080***	0.079***	0.079***

	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
R^2	0.159	0.166	0.166	0.167	0.167
ELS:02 (N= 6,420)					
Item Response Rate	-0.018		-0.048		-0.055
	(0.039)		(0.040)		(0.040)
Cognitive Ability		0.107***	0.110***	0.103***	0.106***
		(0.023)	(0.023)	(0.023)	(0.023)
Control Expectations				-0.008	-0.006
				(0.023)	(0.023)
General Effort and Persistence				0.042**	0.042**
				(0.019)	(0.019)
Years of Education	0.083***	0.067***	0.067***	0.064***	0.064***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
R^2	0.119	0.124	0.124	0.125	0.126

Notes: All explanatory variables in the table are standardized, except years of education, where the unit of measure is a single year of education completed. Regressions restricted to panel participants who were employed. All models control for respondent's household and demographic characteristics. In ELS, years of education were imputed based upon highest degree attained. For NLSY79 the untruncated reported income was used. For the NSLY97 only the truncated income variable is available. We ran a tobit model to account for the truncation of the upper tail. Results were the same as those reported here for all practical purposes. The number of observations for ELS:02 is rounded to the nearest ten per restricted data-use license agreements *** p<0.01, ** p<0.05, * p<0.10

Table 11: Results for Log of Employment Income with Endogenous Selection

	(1)	(2)	(3)	(4)	(5)
NLSY:79 (N= 4,984)					
Wage Equation					
Item Response Rate	0.037		0.030		0.030
	(0.023)		(0.024)		(0.024)
Cognitive Ability		0.120***	0.118***	0.118***	0.116***
		(0.025)	(0.025)	(0.025)	(0.025)
Rotter Locus of Control				-0.009	-0.009
				(0.014)	(0.014)
Years of Education	0.122***	0.100***	0.099***	0.100***	0.099***
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Selection Equation					
Item Response Rate	0.015		0.015		0.015
	(0.039)		(0.040)		(0.040)
Cognitive Ability		0.008	0.006	0.004	0.003
		(0.050)	(0.050)	(0.050)	(0.050)
Rotter Locus of Control				-0.014	-0.013
				(0.035)	(0.035)
Years of Education	0.027*	0.026	0.026	0.026	0.025
	(0.016)	(0.019)	(0.019)	(0.019)	(0.019)
NELS:88 (N=9,063)					
Wage Equation					
Item Response Rate	0.033**		0.032**		0.032**
	(0.014)		(0.014)		(0.015)
Cognitive Ability		0.012	0.008	0.004	0.000
		(0.013)	(0.014)	(0.014)	(0.018)
Locus of Control				0.028**	0.029**
				(0.013)	(0.013)
Self-Concept				0.019	0.019
				(0.012)	(0.011)
Years of Education	0.064***	0.063***	0.063***	0.060***	0.060***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Selection Equation					
Item Response Rate	0.033		0.023		0.024
	(0.039)		(0.041)		(0.046)
Cognitive Ability		0.184***	0.184***	0.177***	0.176***
		(0.043)	(0.044)	(0.045)	(0.046)
Locus of Control				0.037	0.039
				(0.036)	(0.037)
Self-Concept				0.016	0.016
				(0.033)	(0.033)
Years of Education	0.020	-0.007	-0.008	-0.010	-0.011
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)

Add Health (N=2,343)					
Wage Equation					
Item Response Rate	-0.011 (0.023)		-0.013 (0.023)		-0.015 (0.023)
Cognitive Ability		0.032 (0.034)	0.034 (0.034)	0.036 (0.034)	0.037 (0.033)
Self-esteem				0.037 (0.024)	0.037 (0.024)
Years of Education	0.096*** (0.014)	0.092*** (0.014)	0.092*** (0.014)	0.090*** (0.014)	0.091*** (0.014)
Selection Equation					
Item Response Rate	0.026 (0.038)		0.021 (0.038)		0.020 (0.038)
Cognitive Ability		0.047 (0.046)	0.045 (0.046)	0.047 (0.045)	0.044 (0.045)
Self-esteem				0.032 (0.038)	0.033 (0.038)
Years of Education	0.037* (0.022)	0.033 (0.022)	0.033 (0.022)	0.032 (0.022)	0.031 (0.022)
NLSY:97(N= 1,996)					
Wage Equation					
Item Response Rate	0.074 (0.051)		0.073 (0.051)		0.074 (0.050)
Cognitive Ability		0.010 (0.046)	0.003 (0.046)	0.005 (0.049)	-0.002 (0.049)
Coding Speed				0.014 (0.043)	0.014 (0.043)
Years of Education	0.061*** (0.016)	0.065*** (0.017)	0.063*** (0.017)	0.065*** (0.018)	0.063*** (0.017)
Selection Equation					
Item Response Rate	-0.045 (0.039)		-0.051 (0.038)		-0.051 (0.038)
Cognitive Ability		0.151*** (0.047)	0.155*** (0.048)	0.131*** (0.049)	0.135*** (0.049)
Coding Speed				0.047 (0.042)	0.048 (0.042)
Years of Education	0.040** (0.016)	0.019 (0.016)	0.020 (0.016)	0.020 (0..016)	0.020 (0.016)
ELS:02 (N= 6,320)					
Wage Equation					
Item Response Rate	0.023 (0.042)		0.001 (0.043)		-0.028 (0.040)
Cognitive Ability		0.082*** (0.025)	0.082*** (0.025)	0.074*** (0.025)	0.070*** (0.019)
Control Expectations				0.013 (0.023)	-0.057* (0.032)
General Effort and Persistence				0.019 (0.023)	0.039* (0.021)

Years of Education	0.056*** (0.010)	0.045*** (0.010)	0.045*** (0.010)	0.042*** (0.011)	0.052*** (0.008)
Selection Equation					
Item Response Rate	-0.051 (0.069)		-0.049 (0.071)		-0.016 (0.060)
Cognitive Ability		-0.023 (0.033)	-0.019 (0.033)	-0.024 (0.034)	-0.031 (0.030)
Control Expectations				-0.047 (0.041)	-0.057* (0.032)
General Effort and Persistence				0.086** (0.039)	0.102*** (0.032)
Years of Education	0.093*** (0.017)	0.096*** (0.017)	0.096*** (0.017)	0.093*** (0.017)	0.078*** (0.014)

Notes: All explanatory variables in the tables are standardized, except years of education, where the unit of measure is a single year of education completed. FIML endogenous selection models were used to correct for nonrandom selection into employment. In the NLSY79 and NLSY97 datasets, marital status, household with kids, married with children, and a spouse with high income are exclusion restriction variables in selection equation. In NELS:88, Add Health and ELS:02, marital status, gender, an interaction between gender and marital status, and the number of children are the exclusion restriction. The number of observations for ELS:02 is rounded to the nearest ten per restricted data-use license agreements *** p<0.01, ** p<0.05, * p<0.10

.