THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 15-32

Chih Ming Tan
University of North Dakota, USA
The Rimini Centre for Economic Analysis, Italy

Dhanushka Thamarapani
Clark University, USA

THE IMPACT OF SUSTAINED ATTENTION ON LABOR MARKET OUTCOMES: THE CASE OF GHANA

The Impact of Sustained Attention on Labor Market Outcomes: The Case of Ghana

Chih Ming Tan* and Dhanushka Thamarapani#

January 14, 2015

Abstract

In this paper, we go beyond traditional measures of cognitive abilities (IQ) in explaining labor market and social outcomes in developing countries. We exploit a rich dataset from Ghana that provides information on demographics, labor market outcomes, and a direct measure of cognitive ability along with other test scores to construct a measure of sustained attention. Our work is therefore related to the broader literature in Psychology on the importance of executive function on individual behavior and outcomes. We find that, at least for the case of Ghana, after controlling for IQ and other covariates, higher levels of sustained attention are associated with higher annual income, higher education and a higher likelihood of being employed in a white collar job.

Keywords: Ghana; Executive Function; Cognitive and Noncognitive Abilities; Earnings; Occupational Choice

JEL Codes: I10, I15, J24, J30, O12, O55

---

*Corresponding author: Department of Economics, College of Business and Public Administration, University of North Dakota, Gamble Hall, 293 Centennial Drive, Grand Forks, ND 58202-8369, USA; email: chihming.tan@business.und.edu.

#Department of Economics, Clark University, 950 Main Street, Worcester, MA 01610, USA; e-mail: DThamarapani@clarku.edu.

1. Introduction

Following the seminal work of, notably, Heckman and coauthors[1], there has been deep interest in the literature on the impact of cognitive abilities on economic outcomes. In terms of cognitive abilities, economists have generally employed standard measures of intelligence in the Psychology literature for their work. These measures are principally based on psychometric tests; e.g., the intelligent quotient (IQ) and general intelligence g-factor. Economists have used these measures of cognitive abilities to predict various economic outcomes; e.g., wages (Cawley et al., 1996), occupational choice, educational attainment, divorce rates, and incarceration rates (Heckman et al., 2006) and economic development (Hanushek and Woessmann, 2008).

There has also been interest in the effects of what economists refer to as noncognitive abilities on economic outcomes.[2] Cawley et al. (1996) note that there are likely factors other than the g-factor that potentially determine wages and occupational choice. Since most people fall within a relatively tight range around average IQ other factors must contribute to predicting economic and social outcomes. A key focus in the literature has been on the Big Five Factors; i.e., openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (or emotional stability); see, Borghans et al. (2008). The Big Five have been employed to analyze the effects of noncognitive abilities on wages and other social outcomes (Heckman and Kautz, 2012), job performance (Barrick and Mount, 1991), relationship quality (Noftle and Shaver, 2006), and conformity (DeYoung, Peterson and Higgins, 2002). Other researchers have also

[1] See, for example, Heckman et al., (2006), Cunha and Heckman (2008), Heckman and Kautz (2012), and Heckman and Mosso (2014).

[2] Borghans et al. (2008) has pointed out that the term "noncognitive ability" might misleadingly imply that such abilities (e.g., attentiveness, motivation, persistence etc.) are devoid of cognition. As a remedial measure the recent literature uses terms like "soft skills" (Heckman and Kautz, 2012) and "personality traits" (Borghans et al., 2008). We use the term noncognitive abilities with caution so as not to imply the abilities that do not require cognition, but to distinguish them from traditional measures of cognitive abilities (IQ test and g-factor).

focused on factors that facilitate goal-directed behavior such as attention, motivation, persistence, self-esteem, and self-discipline; see, Heckman and Rubinstein (2001), Heckman et al., (2006), Carneiro et al. (2007), and Brunello and Schlotter (2011).

One major drawback with the above approaches is that they can sometimes suffer from a lack of strong theoretical underpinning and limited measurement. For example, there is a lack of consensus on concepts like conscientiousness – how to define and measure it given the context specificity and subjectivity of its nature thereby limiting its predictive validity (Paunonen and Ashton, 2001). Furthermore, proxy variables for these noncognitive abilities typically rely on self-reported data. This data could be misleading as respondents may bias their responses by selectively enhancing their positive traits while downplaying negative ones (Hirsh and Peterson, 2008).

We should also note that most of the earlier work analyzing the influence of cognitive (and noncognitive) abilities on economic outcomes has taken place in the context of developed countries. As Hanushek and Woessmann (2008) points out, the lack of data that directly measures cognitive abilities in developing countries have hindered work in that context. However, there is now an emerging body of work examining the effects of cognitive abilities on child (Engle et al., 2007; Grantham-Mcgregor et al., 2007) and economic outcomes (Glewwe, 1996; Ampaabeng and Tan, 2013; Vogl, 2014) with the increasing availability of data.

The contributions of this paper are as follows. We focus on a developing country context; i.e., Ghana, and employ recently available data – the Ghana Living Standards Survey round 2 (GLSS 2) and the Ghana Education Impact Evaluation Survey round 3 (GEIES 3) – to construct a measure of noncognitive abilities; specifically, sustained attention. Moreover, our measure of sustained attention is not based on self-reported data. It is, in fact, based on the respondent's performance on the sequence of Math (and English) tests administered in GEIES 3. The idea behind the construction of our measure for sustained attention is simple, and derived from the fact that respondents were first given a Simple test before taking an Advance test. Only those respondents who scored more than fifty percent in the Simple test were given the Advance test. We exploit the fact that the difficulty levels between the Simple and Advance tests were dramatically different. We provide samples of these tests in the Appendix. The Simple tests are elementary and correspond roughly to proficiency levels that would be attained by 5th graders. The Advance tests correspond more closely to proficiency levels attained by 9th graders. Therefore, given that the respondent has done well on (i.e., passed) the Advance test, there is no reason why he or she should have made any mistakes at all on the Simple test given the rudimentary nature of the questions on the Simple test as compared to those on the Advance test. We attribute any mistakes made on the Simple test in this case to a lack of sustained attention on the part of the respondent.

More generally, our measure for sustained attention (so constructed) is associated with the broader Psychology literature on executive function.[3] Executive function is distinct from IQ and is defined by the National Center for Learning Disabilities as "a set of mental processes that helps connect past experience with present action." Importantly, individuals have a limit on these functions and they are not context-specific unlike the Big Five factors (Morgan and Lilienfeld, 2000). Diminished executive function "makes activities like planning, organizing, strategizing,

[3] The main components of executive function include inhibitory control, working memory, attentional flexibility, planning, error correction and detection, and many other capacities that are implicated in the monitoring and control of thought and action (Sodian and Hulsken, 2005). These executive function skills are well-defined theoretically as directly influencing daily activities (for example (Carlson et al., 2004) and (Pennington and Ozonoff, 1996)). The executive function also possesses quantifiable and reliable measures since an individual's attentiveness and planning or organizing skills can be measured by robust neuropsychological tests.

remembering details, and managing time and space difficult" for the individual. We contend that errors on the Simple test by respondents who have passed the Advance test reveal potential deficiencies in executive function. We interpret the executive function as a measure of sustained attention because the psychological process most likely to underlie both emotion regulation and executive function is attention; see, for example, Rothbart et al. (2006) and Bell and Deater-Deckard (2007).

It is important to note that both GLSS 2 and GEIES 3 administered a standard IQ (Raven's Progressive Matrices) test along with the Math and English tests. By supplementing the GEIES 3 with the GLSS 2, we also have information on a rich set of individual, family, and community controls for each respondent along with information on outcome variables such as annual income, educational attainment and occupational choice. We employ matching and propensity score methods to obtain causal estimates of the effects of higher sustained attention on labor market and social outcomes. To our knowledge this is the first attempt to analyze the impact of noncognitive abilities on such outcomes based on non-self-reported data using propensity score matching techniques in a developing country context[4]. We find that respondents with higher sustained attention earned more annual income compared to their peers, completed more years of schooling and are also more likely to be employed in a white collar job as opposed to a blue collar job.

The rest of the paper is organized as follows. Section 2 discusses our empirical methodology. Section 3 describes the data. We present our results in section 4. We discuss the results of robustness exercises and falsification tests in section 5. Finally, section 6 concludes.

---

[4] Glewwe, Huang and Park (2013) have conducted a study on impact of noncognitive skills on labor market outcomes in rural China. However they employ a panel dataset of self-reported noncognitive skills and do not use matching techniques.

2. Methodology

In terms of our empirical strategy, we follow Thiel and Thomsen (2013). Our benchmark model takes the following form,

$$y_{i,t} = \gamma\, EF_{i,t} + X_{i,t}\,\beta + \varepsilon_{i,t}\,, \tag{1}$$

where the outcome variable, (log of annual income, a discrete variable for occupational choice and years of completed schooling) is denoted by $y_{i,t}$ for individual i at time t. Whether a person has high sustained attention or not is denoted by $EF_{i,t}$ which, in our context is the "treatment" variable. $X_{i,t}$ is a vector of covariates comprising of a respondent's age, cognitive ability (IQ), gender, locality, parents' education, family size, and height. We postpone a detailed discussion of the variables and data to the next section and refer the reader to Table 1A for descriptive statistics of the original (unbalanced) sample.

The coefficient of interest is γ that measures the (causal) impact of higher sustained attention. The regression model above corresponds to the method of regression adjustment (RA) under linear specification in the impact evaluation literature that uses a two-step approach to estimate treatment effects. This method fits regression models for control and treatment groups separately given the same set of covariates. It then calculates the averages of the predicted outcome (e.g. annual income) for those with high sustained attention and low sustained attention. We can then obtain estimates for the average treatment effect on the treated (ATET) by restricting the computation of averages to the subset of individuals with high sustained attention. In the case of RA, our findings are sensitive to the parametric assumptions of the regression model and also potential imbalance in the covariate distribution across treatment and control groups.

We therefore implement matching estimation. The benefits of matching are that it does not impose a parametric specification on the underlying model (Edwards and Magedzo, 2001) and it is able to potentially achieve covariate balance when the data allows it. The basic idea behind this approach is to use the available data to recreate conditions so that the assignment of the treatment variable; sustained attention, in this case, can be thought of as a randomized experiment after controlling for a set of covariates $X_{i,t}$.

If we construct sustained attention to be a binary variable, then, the focus of the treatment effects approach focuses on the ATET:

$$\delta = E(y_1 - y_0 | x, u, EF = 1), \tag{2}$$

where y_1 is the potential outcome for a respondent if she has high sustained attention while y_0 is the potential outcome for that same respondent if she has low sustained attention. The key problem here is that we cannot simultaneously observe both potential outcomes for each respondent; one of the two has to be a counterfactual. The central idea behind matching estimation is to use the data to construct the missing observations that would allow us to evaluate the counterfactual outcome. This is achieved by pairing a respondent in the treatment group with one or more respondents in the control group with similar covariate values. Rosenbaum and Rubin (1983) show that an efficient way of performing this comparison is to employ the propensity score (PS), defined as the probability of participating in the treatment.

In this paper, we consider results from a range of matching strategies. We employ matching with replacement to ensure that the sample size is not constraining our results (without replacement method requires much more data):

1. One nearest neighbor using Mahalanobis distance (NN)
2. One nearest neighbor using the PS (NN-PSM)
3. Caliper matching using the PS with caliper set to 0.1 (PSM)

We also employ a number of other methods based on the PS that are known to have good efficiency and functional form robustness properties (see, Wooldridge (2010)):

1. Inverse Probability Weighting (IPW)
2. IPW with Regression Adjustment (IPWRA)

The IPW estimator addresses the problem arising from the fact that each respondent is observed in only one of the treatment outcomes (that is, a person cannot have both high sustained attention and low sustained attention simultaneously) by comparing her outcomes to respondents in the comparison group weighted according to the probability that those respondents would have been in the original respondent's assignment group. Finally, IPWRA uses weighted regression coefficients to calculate the average of predicted outcomes at the treatment level with weights created by inverse probabilities of the treatment. Thus IPWRA possesses the doubly robust property. We show that matching techniques establish the covariate balance in the resulting matched sample so that most of the covariates are not statistically different from each other for the individuals with high sustained attention versus low sustained attention[5]. We summarize the covariate balance results for the matched sample in Table 1B.

[5] The instances where matching still results in covariate imbalance is with respect to age, IQ, parents' schooling and primary school quality. The treated group is statistically older than the control group. As for IQ, parents' schooling and primary school quality, we conduct robustness checks without them and our results remain robust.

3. Data

Our two cross sectional datasets for Ghana come from the Ghana Living Standards Survey round 2 (GLSS 2) of 1988/89 and the Ghana Education Impact Evaluation Survey round 3 (GEIES 3) of 2003. Both of these nationally representative surveys were fielded by the Ghana Statistical Services (GSS) with technical assistance from the World Bank. The household questionnaire provide detailed information on the educational attainment, demographic characteristics, weights and heights, economic activities, housing conditions, land and livestock ownership, and household expenditure for all the household members. GLSS 2 thereby covered 14,924 individuals in 3,192 households. Furthermore, GLSS 2 tested mathematics, reading, and abstract thinking skills of all household members, and teachers in 85 randomly selected clusters out of 170 clusters surveyed at the time. GEIES 3 was conducted in all the 85 clusters selected by the GLSS 2, thereby surveying a total of 1,740 households corresponding to approximately 8,000 individuals.

In both rounds the household members of age 9 to 55 years were given the Raven's Progressive Matrices (IQ) test. Those who had 3 or more years of schooling received the Simple (English) reading test and Simple Math test. From the Simple test takers only those who scored above 50 percent were asked to take the advanced version of the test. Although the Simple tests and Raven test were given to individuals of 9 years and older, to avoid the cases of child labor in our earnings regressions we only employ the dataset of individuals who are strictly older than 15 years which is the legal working age in Ghana.

We now discuss our treatment variable. To measure sustained attention (or, executive function) we derive a proxy using the Simple and Advance (Math and English) tests. The tests given are exactly the same in terms of the question, the answer choices, and the order in which they appear in both GLSS 2 and GEIES 3 rounds. The Advance tests were given in the form of multiple-choice questions. The respondents selected the correct answer from a pool of 4 choices[6]. Thus there is the possibility that respondents simply "guessed" the answers rather than actually deriving the correct solutions. To address this issue, we calculate the number of correct answers the respondents could have got right had they been blatantly guessing. For example, if a person took both Advance Math (total of 36) and Advance English (total of 29) tests, by random guessing a person should get a combined score of about 16.25 for the Advance tests, i.e. an average Advance test score of 8.125. We then categorize individuals who scored better than random guessing (respondents with an average Advance test score greater than 8.125) as the "smart" pool who passed the Advance tests[7]. Contrary to the Advance tests the Simple test questions were straightforward[8]. Therefore if a person did better than random guessing in the Advance tests then that person is believed to have high cognitive skills. Intuitively such people with high cognitive skills would be expected to get a full score in the Simple tests given the skill level needed. However if the "smart" respondents do not receive full scores for the Simple tests despite their high cognitive skill they are associated with lack of sustained attention. Thus the proxy for sustained attention or executive function $(EF_{i,t})$, takes the value of 1 for those who belonged to the "smart" pool and scored a perfect score in the simple test and 0 otherwise. Our final dataset retains only the observations with non-missing values for sustained attention thus the effective sample size is 1,544 who are strictly older than 15 years.

[6] Refer to appendix 1 for a sample of Advance test questions.

[7] There are cases where some respondents took only one of the Advance Math test or the Advance English test. In such cases, we adjust the random guessing score to be 7.25 for Advance English-only test takers and 9 for the Advance Math-only test takers.

[8] Refer to appendix 1 for a sample of Simple test questions.

The labor market outcomes are measured using annual income and occupational choice. The log of annual income ($lnY_{i,t}$), is originally measured in Cedis and corrected for inflation using the consumer price index (base year 2005) for the incomes generated in GLSS 2 and GEIES 3 rounds. As mentioned above the incomes are only calculated for individuals who are strictly above the legal working age of 15 years. Since farming, livestock rearing, and self-employment are commonplace in Ghana we take into account all forms of payment in which the respondents could have been paid. This includes payments from primary and secondary jobs received in the form of monetary, food supplies, housing, clothing, transportation and other in kind. The annual income is only applied to individuals who are employed for wages, self-employed in agriculture or self-employed in business. Occupational choice ($WC_{i,t}$), is a binary variable which takes the value of 1 if a person is employed in a white collar job and 0 for blue collar jobs. White collar jobs include salaried occupations that do not involve manual labor. In our datasets the workers occupations were categorized according to the International Standard Industrial Classification of All Economic Activities (ISIC). We find that most jobs that fit the white collar criteria are classified under the services sector which include finance, insurance, real estate, business, public administration, medical and health, research, and other such white collar jobs which are more likely to be pursued by individuals with high noncognitive abilities. If a person is employed in these sectors she would get a value of 1 for occupational choice ($WC_{i,t}$) variable. The base group is the blue collar jobs in agriculture, mining and manufacturing, and trade and transport. However there are services like sanitary, repair, domestic, and laundry services that are also categorized under services sector and we impute a value of 0 for such individuals to avoid misclassification to white collar job.

The social outcomes in this study are measured by educational attainment and indebtedness. To measure educational attainment $(Edu_{i,t})$, we calculate the total year of schooling a person has completed. We restrict the sample to respondents who are above 25 years of age since most people in Ghana have completed schooling by that age although some opt to go for doctoral studies before entering the labor market. However in our sample such is not the norm. Our second social outcome variable is indebtedness using the number of loans acquired by the household. GLSS 2 provides information of the number of loans $(Loans_i)$ undertaken at the household level (not at individual level so as to who in fact took the loan). Since the decision whether to get a loan is usually taken by the head of the household we restrict our sample to the heads of the households in the 1989 cohort since the credit and savings module was not re-interviewed in GEIES 2003.

In our exercises, we control for a host of covariates $(X_{i,t})$. We include a direct measure of the respondent's cognitive ability $(IQ_{i,t})$ using scores from the Raven's Progressive Matrices test. We note that the test questions were unaltered in both cross sections hence assuring the consistency of the IQ measurement for individuals across different rounds[9]. We standardize the Raven test scores to have a mean 0 and standard deviation of 1 among all the test takers of 9 to 55 years in the respective rounds.

Ghana underwent rigorous educational reforms in 1988/89 period which changed the resource landscape for many schools in terms of physical resources (textbooks, chalk, blackboards, classrooms), personnel (teacher quality), and community participation in the decision making process which resulted in decentralization of the education system. White (2004) states that physical resources quality improvement has led to quantity (enrollment)

[9] Refer to appendix 1 for a sample of Raven's test questions.

increase in schools by approximately 25 percent. We therefore identify individuals who completed their primary schooling by 1989 (when the education reforms started) through a binary variable for pre education reform cohort ($Cohort_i$) that takes the value of 1 if an individual completed primary school before the reforms and 0 otherwise. All the individuals of GLSS 2 therefore will take a value of 1 in our sample (since they need to be at least 16 years of age to qualify for our sample) and those who are 27 years or more in GEIES 3 (at most 15 years in 1989) will also take the value of 1 signifying the pre-education reform cohort.

Apart from the binary classification for the education reforms we include measures for school quality that prevailed at the time the individual was in school. GLSS 2 and GEIES 3 provide cluster level school quality information regarding Math and English textbooks provision, proportion of unusable classrooms, and teacher IQ as measured by a Raven's test. We match the cluster average school quality to the individuals[10]. Our sample size would be compromised if we include the different types of school quality separately. What we are interested in is to find out which individuals had better educational resources. We therefore use indicator variables to distinguish high quality in school resources by comparing textbooks, proportional of unusable classrooms and teacher quality separately to the regional median. We then construct a composite school quality index ($Index_{i,t,l}$) for individual i at time t for level l of school where the levels are categorized as primary and secondary schooling. For example, if the number of primary textbooks per student in a particular cluster is greater than the regional median we recognize it as high primary textbook quality. Similarly if the cluster average for the proportion of primary unusable classrooms is lower than the regional median it indicates high primary classroom quality and high primary teacher IQ than the regional median indicates high primary teacher

[10] We conduct robustness check without school quality and our results remain robust.

quality. We repeat the process for each element of secondary school quality. The index variable ($Index_{i,t,l}$) will take 1 for primary school quality if all three of the primary textbook quality, classroom quality and teacher quality are categorized as high and 0 otherwise. We allocate primary school quality to all those individuals who had at least 1 year of schooling. We repeat the same process to derive secondary school quality index and allocate it to individuals who completed at least 7 years of schooling (primary education takes 6 years (Oduro, 2000)). In order not to compromise our sample size we include indicator variables for missing primary and secondary school quality.

We also include fixed effects for region and the type of locality (urban or rural), as well as a gender dummy. We use the distance in minutes to the nearest primary school as another measure of remoteness. In order to account for the household characteristics of the individuals growing up, we include the family size and parents' education as suggested by the literature; see, for example, Black et al. (2005). Following Ampaabeng and Tan (2013) we include a measure of parents' education which combines the education status of both parents. We incorporate the standardized height for the individuals as suggested by Vogl (2014). Finally, we include a full set of interaction terms with the locality.

4. Results

In this section we describe our estimation results for labor market and social outcomes. The labor market outcomes include log of annual income and occupational choice. The income results are summarized in Table 2 while the probit estimation results for the PS calculations are summarized in Table 3. Across all methods described in Section 2, ATET results show that there is highly statistically significant positive mean difference between high sustained attention individuals and their low sustained attention counterparts with respect to earning capacity.

According to the RA estimate for ATET, high sustained attention increases an individual's annual income by an average of 33 percent. The full RA results are summarized in Table 4. The full RA results also show that for the set of low sustained attention respondents, higher IQ and higher secondary school quality were associated with higher annual income, whereas for the high sustained attention respondents, only family background (parents' secondary schooling) appeared to be associated with higher annual income. We then checked for both substantial overlap across both treatment and control groups (see, Appendix 2 for the overlap plots), and report similar findings using matching estimation and the other PS methods discussed in Section 2; see, Table 2. The sign and the significance of the estimated ATET are robust to different matching and PS approaches employed while the magnitudes were consistently in the range of 30 to 32 percent, and only slightly smaller than that for RA.

We summarize results for occupation choice in Table 5. As described in the previous section, for the occupational choice exercises, we defined the outcome variable to take the value of 1 if a person is engaged in a white collar job; 0 otherwise. The ATET results reveal that, among high sustained attention individuals, having high sustained attention increases the probability of choosing a white collar job by an average of 13 percent. The ATET findings are also robust in signs, magnitudes and for the most part in significance levels across estimation strategies[11].

We now turn to the findings for the social outcomes variables. In Table 6 we summarize the results for educational attainment. As mentioned in the previous section, for these

[11] The ATET results are not significant for PSM matching.

regressions, we only consider the individuals who are above 25 years of age and thus have completed their education. The estimates for the ATET suggest that high sustained attention results in people getting significantly more schooling. The magnitude of the estimated ATET under RA matching technique suggests an increase of schooling by around 8.6 months. All the other matching techniques yield statistically significant results where the magnitudes range from 7 to 12.8 months. The results for the indebtedness of the household heads are less interesting; see, Table 7. While there is a strong consensus across the estimation approaches that the ATET estimate is negative, the findings are not statistically significant. One reason for the lack of precision for the ATET estimates has to do with the fact that we only consider the household heads of GLSS 2. Our sample size shrinks dramatically as a consequence. Nevertheless, the negative ATET suggest that individuals with higher executive function may be able to assert more control over their financial status. As the above results describe, these high functioning individuals are better educated and are employed at better jobs with higher earned incomes. The (tentative) evidence here thus suggests that these individuals are able to get by with fewer loans when managing their household expenditures.

5. Robustness Checks and Falsification Test

We conduct a series of robustness checks of our benchmark results as shown in Tables 8A, 8B and 8C.

We first check the robustness of our results to variations in the list of covariates. A driving concern in these exercises is to assure that our benchmark findings remain robust after the removal of potentially non-pre-treatment or predetermined (i.e., endogenous) covariates. Table

8A summarizes the coefficients estimated using a model without standardized IQ, pre education reform cohort, and school quality variables. As shown in the table, our benchmark findings remain robust. Under RA, we find that among those who are high sustained attention individuals, having high sustained attention results in a person earning 42 percent more income, 15percent more likely to be a white collar worker, and would have 16 more months (one year and 4 months) of additional schooling. The signs, the magnitudes and the significance of the ATET estimates are broadly similar across the different matching strategies. We next estimate an even more stringent model. Thus we impose more restrictions such that the covariates included are limited to be (undoubtedly) pre-determined variables: age, gender, and locality. The results are summarized in Table 8B and are broadly comparable to those obtained in Table 8A. We observe highly statistically significant ATET estimates for all the outcome variables. In sum, not only do our benchmark results remain robust, the estimated ATET in our robustness exercises for all of the outcome variables are substantially larger than that in the benchmark case.

As a further robustness check, we expand our “control” group to include both low sustained attention individuals and those who were eligible to take the advance test (passed the simple test) but did not. We therefore attempt to address any issues of selection into the test. Answering the advance test was voluntary and there could be any number of reasons as to why a person might have forgone taking the test even when eligible. For example, they may not have been present at the household when the test was given. We impute a value of 0 for these individuals thus expanding the original control group of low sustained attention individuals. As shown in Table 8C, our results for all three outcome variables remain robust.

Finally, we conduct a falsification test using the multiple control groups test. We first construct an alternative (false) treatment group that is equal to the low sustained attention group (the original control group). We also create a false control group by exploiting the fact that some respondents who passed the simple test (and were thus eligible to take the advance test) chose not to take the advance test. We then re-run the labor market outcomes and educational attainment models with the new (false) treatment variable. We summarize the results in Table 9. We expect to see insignificant results (at the 5 percent level) for the false treatment since, by assumption, the false treatment and false control groups should be indistinguishable once we have controlled for the set of observables. We observe unambiguous insignificant results for the occupational choice and educational outcome regressions. However for the annual income regressions, we do observe statistically significant results whenever RA was employed. Nevertheless, the matching and PS approaches that do not employ RA are all robust. The assumption of strong ignorability appears more defensible in those cases where attention has been paid to ensure sufficient overlap and where bias from selection-on-observables is minimized.

6. Conclusions

Using two waves of a unique cross sectional dataset from Ghana we estimate the causal impact of sustained attention (or, executive function) on labor market and social outcomes. We construct a novel proxy for sustained attention using the avoidable mistakes that high cognitive ability respondents made on simple achievement tests. Our proxy is also interpretable as a measure of executive function. Crucially, unlike the existing literature, our measure of sustained attention (noncognitive ability) was not derived from self-reported outcome measures. We investigate the effects of possessing high sustained attention on respondents' annual income, occupational choice and educational attainment using a variety of regression, matching, and propensity score weighting approaches that are now standard in the impact evaluation literature. We also conduct a series of robustness checks and a falsification test to check the validity of our results. Our findings suggest that respondents with high noncognitive abilities earn substantially more annual income compared to their peers, are more likely to be employed in a white collar job as opposed to a blue collar job, and are likely to be more educated. This paper therefore contributes another source of evidence (from a developing, Sub-Saharan African country) to the existing literature pioneered by Heckman and coauthors, that has largely focused on data from the developed world, on the importance of noncognitive skills development to individual socio-economic outcomes.

References

Ampaabeng, K. and Tan, C. M. (2013) "The Long-term Cognitive Consequences of Early Childhood Malnutrition: The case of famine in Ghana," *Journal of Health Economics*, 32, pp. 1013–1027.

Barrick, M. R. and Mount, M. K. (1991) "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis," *Personnel Psychology*, 44(1), pp. 1–26.

Bell, M. A. and Deater-Deckard, K. (2007) "Biological Systems and the Development of Self-regulation: Integrating Behavior, Genetics, and Psychophysiology," *Journal of Developmental & Behavioral Pediatrics*, 28, pp. 409–420.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005) "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 95(1), pp. 437–449.

Borghans, L., Duckworth, A. L., Heckman, J. and ter Weel, B. (2008) "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 43, pp. 972 – 1059.

Brunello, G. and Schlotter, M. (2011) *Non-Cognitive Skills and Personality Traits: Labour Market Relevance and Their Development in Education & Training Systems*. (Discussion Paper Series).

Carlson, S. M., Moses, L. J. and Claxton, L. J. (2004) “Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability.,” *Journal of experimental child psychology*, 87(4), pp. 299–319.

Carneiro, P., Crawford, C. and Goodman, A. (2007) *The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes*. London: London School of Economics.

Cawley, J., Conneely, K., Heckman, J. and Vytlacil, E. (1996) *Measuring the Effects of Cognitive Ability*. (NBER Working Paper Series No. 5645).

Cunha, F. and Heckman, J. (2008) “Formulating, Identifying, and Estimating the Technology of Cognitive and Noncognitive Skill Formation,” *Journal of Human Research*, 43(4), pp. 738–782.

DeYoung, C. G., Peterson, J. B. and Higgins, D. M. (2002) “Higher-order factors of the Big Five predict conformity: Are there neuroses of health?,” *Personality and Individual Differences*, 33(4), pp. 533–552.

Edwards, S. and Magedzo, I. I. (2001) *Dollarization, Inflation and Growth*. (NBER Working Paper No. 8671).

Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L., Martorell, R. and Young, M. E. (2007) “Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world,” *Lancet*, 369(9557), pp. 229–42.

Glewwe, P. (1996) “The relevance of standard estimates of rates of return to schooling for education policy: A critical assessment,” *Journal of Development Economics*, 51(2), pp. 267–290.

Glewwe, P., Huang, Q. and Park, A. (2013) *Cognitive Skills, Noncognitive Skills, and the Employment and Wages of Young Adults in Rural China*.

Grantham-Mcgregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L. and Strupp, B. (2007) “Developmental potential in the fi rst 5 years for children,” *Lancet*, 369(9555), pp. 60–70.

Hanushek, E. and Woessmann, L. (2008) “The role of cognitive skills in economic development,” *Journal of Economic Literature*, 46(3), pp. 607–668.

Heckman, J. and Kautz, T. (2012) “Hard Evidence on Soft Skills,” *Labour Economics*. (NBER Working Paper Series), 19(4), pp. 451–464.

Heckman, J. and Mosso, S. (2014) *The Economics of Human Development and Social Mobility*. (NBER Working Paper No. 19925).

Heckman, J. and Rubinstein, Y. (2001) “The Importance of Noncognitive Skills : Lessons from the GED Testing Program,” *American Economic Review*, 91(2), pp. 145–149.

Heckman, J., Stixrud, J. and Urzua, S. (2006) “The Effects of Cognitive and Noncognitive Abilities on Labor Market,” *Journal of Labor Economics,* 24(3), pp. 411–482.

Hirsh, J. B. and Peterson, J. B. (2008) “Predicting creativity and academic success with a ‘Fake-Proof’ measure of the Big Five,” *Journal of Research in Personality*, 42(5), pp. 1323–1333.

Morgan, A. B. and Lilienfeld, S. O. (2000) “A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function,” *Clinical Psychology Review*, 20(1), pp. 113–136.

Noftle, E. E. and Shaver, P. R. (2006) “Attachment dimensions and the big five personality traits: Associations and comparative ability to predict relationship quality,” *Journal of Research in Personality*, 40(2), pp. 179–208.

Oduro, A. (2000) “Basic education in Ghana in the post-reform period,” *Accra: Centre for Policy Analysis*.

Paunonen, S. V. and Ashton, M. C. (2001) “Big Five Factors and Facets and the Prediction of Behavior,” *Journal of Personality and Social Psychology*, 813(3), pp. 524–539.

Pennington, B. F. and Ozonoff, S. (1996) “Executive Functions and Developmental Psychopathology,” *Journal of Child Psychology and Psychiatry*, 37(1), pp. 51–87.

Rosenbaum, P. and Rubin, D. (1983) “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70(1), pp. 41–55.

Rothbart, M. K., Posner, M. I. and Kieras, J. (2006) “Temperament, Attention, and the Development of Self-Regulation,” in McCartney, K. and Phillips, D. (eds) *Blackwell Handbook of Early Childhood Development*. Malden, MA: Blackwell Publishing, pp. 338–357.

Sodian, B. and Hulsken, C. (2005) “The Developmental Relation of the Theory of Mind and Executive Functions: A Study of Advanced Theory of Mind Abilities in Children With Attention Deficit Hyperactivity Disorder,” in Schneider, W., Schumann-Hengsteler, R., and Sodian, B. (eds) *Young Children’s Cognitive Development*. New Jersey: Lawrence Erlbaum Associates, pp. 175 – 187.

Thiel, H. and Thomsen, S. L. (2013) “Noncognitive Skills in Economics : Models , Measurement , and Empirical Evidence,” *Research in Economics*, 67(2), pp. 189–214.

Vogl, T. S. (2014) “Height, Skills, and Labor Market Outcomes in Mexico,” *Journal of Development Economics*. Elsevier B.V., 107, pp. 84–96.

White, H. (2004) *Books, Buildings, and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana, World Bank*. World Bank Publications.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. MIT Press.

Table 1A – Descriptive Statistics Unmatched Sample

	High Sustained Attention (Treatment)			Low Sustained Attention (Control)			P-value
VARIABLES	N	Mean	SD	N	Mean	SD	
Log Annual Income	194	10.82	1.201	733	10.34	1.298	0.000
Employed in White Collar Job	205	0.410	0.493	830	0.229	0.420	0.000
Years of Schooling	304	11.93	3.199	1,238	10.59	2.857	0.000
Female	304	0.329	0.471	1,240	0.362	0.481	0.273
Rural	304	0.316	0.466	1,240	0.319	0.466	0.926
Age	304	31.13	10.33	1,240	28.42	10.10	0.000
Pre Education Reform Cohort	304	0.753	0.432	1,240	0.735	0.441	0.521
Family Size	304	4.961	2.968	1,240	5.316	3.317	0.068
IQ Standardized	300	1.365	0.879	1,226	25.14	6.647	0.000
Height Standardized	295	0.699	0.540	1,202	0.661	0.576	0.281
Repeated Cross Section	304	0.487	0.501	1,240	0.469	0.499	0.585
Parents Schooling Primary or less	304	0.069	0.254	1,240	0.156	0.363	0.000
Parents Schooling Secondary	304	0.082	0.275	1,240	0.106	0.308	0.195
Parents Schooling Tertiary	304	0.118	0.324	1,240	0.074	0.262	0.027
Distance (minutes) to School	296	10.09	9.472	1,224	10.75	10.74	0.292
Primary School Quality Index	304	0.164	0.371	1,240	0.007	0.084	0.160
Secondary School Quality Index	304	0.128	0.335	1,240	0.127	0.334	0.967

Table 1B – Descriptive Statistics Matched Sample

	High Sustained Attention (Treatment)			Low Sustained Attention (Control)			P-value
VARIABLES	N	Mean	SD	N	Mean	SD	
Log Annual Income	179	10.81	1.229	697	10.31	1.303	0.000
Employed in White Collar Job	179	0.425	0.496	696	0.253	0.435	0.000
Years of Schooling	179	12.36	3.360	697	10.92	2.809	0.000
Female	179	0.324	0.469	697	0.327	0.469	0.937
Rural	179	0.369	0.484	697	0.354	0.479	0.723
Age	179	34.85	8.777	697	32.90	9.173	0.009
Pre Education Reform Cohort	179	0.916	0.278	697	0.897	0.305	0.412
Family Size	179	4.743	2.932	697	4.736	3.116	0.977
IQ Standardized	179	1.372	0.890	697	0.775	0.927	0.000
Height Standardized	179	0.717	0.521	697	0.728	0.538	0.799
Repeated Cross Section	179	0.397	0.491	697	0.400	0.490	0.929
Parents Schooling Primary or less	179	0.0615	0.241	697	0.106	0.308	0.037
Parents Schooling Secondary	179	0.0223	0.148	697	0.042	0.200	0.152
Parents Schooling Tertiary	179	0.0279	0.165	697	0.016	0.125	0.359
Distance (minutes) to School	179	9.866	9.391	697	0.400	0.490	0.291
Primary School Quality Index	179	0.117	0.323	697	0.179	0.384	0.028
Secondary School Quality Index	179	0.0950	0.294	697	0.103	0.305	0.737

Table 2 – Matching Results for Log Annual Income Regressions

	Dependent Variable Log of Annual Income
	ATET
Regression Adjustment (RA)	0.3296***
	(0.099)
Nearest Neighbor (NN)	0.2212*
	(0.120)
Nearest Neighbor PSM (NN - PSM)	0.3101**
	(0.132)
Propensity Score Matching (PSM)	0.3101**
	(0.132)
Inverse Probability Weights (IPW)	0.3086***
	(0.097)
IPW with Regression Adjustment (IPWRA)	0.3011***
	(0.102)
N	876

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.

The sample includes individuals above 15 years of age.

The set of controls include age, gender, standardized IQ, family size, parents' education, locality, pre education reform cohort, standardized height, distance (minutes) to nearest school, school quality, a survey year dummy, and full set of interactions with locality.

ATET refers to average treatment of the treated.

The propensity score model is probit. The RA model assumes linearity.

The reported results for NN and NN-PSM are for one nearest neighbor.

The reported results for PSM use a caliper of 0.3.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3 – Probit Estimation for Propensity Score

VARIABLES	Dependent Variable High Sustained Attention = 1
Age	0.1132***
	(0.034)
Age-squared	-0.0013***
	(0.000)
IQ Standardized	0.4629***
	(0.064)
Female	-0.0928
	(0.116)
Family Size	-0.0100
	(0.018)
Parents Schooling Secondary	0.3094*
	(0.188)
Parents Schooling Tertiary	0.5823***
	(0.190)
Rural	-0.0134
	(0.890)
Pre Education Reform Cohort	-0.5277***
	(0.154)
Height Standardized	-0.2346**
	(0.099)
Distance (minutes) to School	-0.0081
	(0.008)
Primary School Quality	-0.1384
	(0.120)
Secondary School Quality	-0.0837
	(0.145)
Constant	-2.9638***
	(0.556)
N	1,462

The set of controls includes a survey year dummy and a full set of interactions with the locality. None of the coefficients of the interaction terms were statistically significant.
The sample includes individuals above 15 years of age.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4 – Regression Adjustment Results for Log Annual Income Regressions

	Dependent Variable Log of Annual Income		
VARIABLES	ATET	OME0	OME1
Age		0.0880**	0.1197
		(0.044)	(0.082)
Age-squared		-0.0009	-0.0011
		(0.001)	(0.001)
IQ Standardized		0.1499***	0.1649
		(0.048)	(0.115)
Female		0.0249	-0.0052
		(0.130)	(0.238)
Family Size		0.0016	0.0306
		(0.021)	(0.026)
Parents Schooling Secondary		-0.2251	1.4308**
		(0.255)	(0.649)
Parents Schooling Tertiary		-0.3089	0.3452
		(0.367)	(1.091)
Rural		-1.6462	-3.8827
		(1.308)	(3.510)
Pre Education Reform Cohort		-0.3534*	-0.2025
		(0.212)	(0.388)
Height Standardized		0.0719	0.0510
		(0.117)	(0.215)
Distance (minutes) to School		0.0024	0.0147
		(0.007)	(0.013)
Primary School Quality		-0.0636	0.1339
		(0.114)	(0.212)
Secondary School Quality		0.2735**	-0.0157
		(0.136)	(0.246)
High Sustained Attention vs. Low Sustained Attention	0.3296***		
(Executive Function 1 vs 0)	(0.099)		
Constant		8.5692***	7.3837***
		(0.742)	(1.432)
N		876	

ATET refers to average treatment of the treated. Coefficients under OME0 are from the linear equation used to estimate the potential outcome means of low sustained attention individuals. Coefficients under OME1 are from the linear equation used to estimate the potential outcome means of high sustained attention individuals.

The sample includes individuals above 15 years of age.

The set of controls includes a survey year dummy and a full set of interactions with locality. None of the coefficients of the interaction terms were statistically significant.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 - Matching Results for White Collar Regressions

	Dependent Variable Employed in White Collar Job
	ATET
Regression Adjustment (RA)	0.1327***
	(0.040)
Nearest Neighbor (NN)	0.1277**
	(0.053)
Nearest Neighbor PSM (NN - PSM)	0.1064**
	(0.047)
Propensity Score Matching (PSM)	0.0479
	(0.051)
Inverse Probability Weights (IPW)	0.1203***
	(0.040)
IPW with Regression Adjustment (IPWRA)	0.1106***
	(0.042)
N	978

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The sample includes individuals above 15 years of age.
The set of controls includes age, gender, standardized IQ, family size, parents' education, locality, pre education reform cohort, standardized height, distance (minutes) to nearest school, school quality, a survey year dummy, and a full set of interactions with locality.
ATET refers to average treatment of the treated.
Propensity score model is probit. The outcome model for RA was also probit.
The reported results for NN and NN-PSM are for one nearest neighbor.
The reported results for PSM use a caliper of 0.2.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6 - Matching Results for Educational Attainment

Matching Technique	Dependent Variable Years of Schooling ATET
Regression Adjustment (RA)	0.7202***
	(0.265)
Nearest Neighbor (NN)	1.0722***
	(0.321)
Nearest Neighbor PSM (NN - PSM)	0.7778**
	(0.395)
Propensity Score Matching (PSM)	0.7778**
	(0.395)
Inverse Probability Weights (IPW)	0.6645**
	(0.272)
IPW with Regression Adjustment (IPWRA)	0.5961**
	(0.27)
N	803

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The sample includes individuals who are above 25 years of age.
The regressions are controlled for age, gender, standardized IQ, family size, parents' education, locality, pre-education reform cohort, standardized height, distance (minutes) to nearest school, school quality, a survey year dummy, and full set of interactions with locality.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model assumes linearity.
The reported results for NN and NN-PSM are for one nearest neighbor.
The reported results for PSM use a caliper of 0.2.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7 – Matching Results for Number of Loans Taken by the Household Head

	Dependent Variable Number of Loans Taken by Household Head
	ATET
Regression Adjustment (RA)	-0.0884
	(0.062)
Nearest Neighbor (NN)	0.0204
	(0.020)
Nearest Neighbor PSM (NN - PSM)	-0.1939
	(0.132)
Propensity Score Matching (PSM)	-0.3878
	(0.251)
Inverse Probability Weights (IPW)	-0.0826
	(0.062)
IPW with Regression Adjustment (IPWRA)	-0.0879
	(0.064)
N	224

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The sample contains only the head of households in the GLSS 2 (1989).
The set of controls includes age, gender, standardized IQ, family size, locality, and standardized height.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model assumes linearity.
The reported results for NN are for one nearest neighbor. NN-PSM are for two nearest neighbor.
The reported results for PSM use a caliper of 0.1.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8A – Robustness Checks

Matching Technique	Dependent Variable Log of Annual Income ATET	Dependent Variable Employed in White Collar Job ATET	Dependent Variable Years of Schooling ATET
Regression Adjustment (RA)	0.4227***	0.1514***	1.3237***
	(0.093)	(0.038)	(0.272)
Nearest Neighbor (NN)	0.3866***	0.1780***	1.1530***
	(0.110)	(0.045)	(0.386)
Nearest Neighbor PSM (NN - PSM)	0.4759***	0.1466***	1.8142***
	(0.142)	(0.050)	(0.345)
Propensity Score Matching (PSM)	0.4759***	0.1571***	1.8142***
	(0.142)	(0.053)	(0.345)
Inverse Probability Weights (IPW)	0.4309***	0.1535***	1.3147***
	(0.092)	(0.038)	(0.271)
IPW with Regression Adjustment (IPWRA)	0.4260***	0.1530***	1.3093***
	(0.093)	(0.038)	(0.272)
N	883	985	811

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The treatment group is the high sustained attention individuals. The control group is the low sustained attention individuals.
The sample for log annual income and employed in white collar job includes individuals who are above 15 years of age. The sample for years of schooling includes individuals who are above 25 years of age.
The set of controls includes age, gender, family size, parents' education, locality, standardized height, distance (minutes) to nearest school, a survey year dummy, and the full set of interactions with locality.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model assumes linearity for log of annual income and years of schooling results. For the employed in white collar job results RA model is also probit.
The reported results for NN and NN-PSM are for one nearest neighbor. The reported results for PSM use a caliper of 0.1.
Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8B – Robustness Checks

	Dependent Variable Log of Annual Income ATET	Dependent Variable Employed in White Collar Job ATET	Dependent Variable Years of Schooling ATET
Regression Adjustment (RA)	0.4138***	0.1647***	1.1893***
	(0.089)	(0.037)	(0.262)
Nearest Neighbor (NN)	0.4012***	0.1747***	1.2354***
	(0.100)	(0.040)	(0.294)
Nearest Neighbor PSM (NN - PSM)	0.4096***	0.1698***	1.2129***
	(0.104)	(0.041)	(0.290)
Propensity Score Matching (PSM)	0.4096***	0.1719***	1.2129***
	(0.104)	(0.041)	(0.290)
Inverse Probability Weights (IPW)	0.4139***	0.1633***	1.1909***
	(0.088)	(0.037)	(0.262)
Augmented IPW (AIPW)	0.3852***	0.1578***	1.0970***
	(0.091)	(0.037)	(0.251)
IPW with Regression Adjustment (IPWRA)	0.4154***	0.1650***	1.1900***
	(0.089)	(0.037)	(0.262)
N	927	1,035	846

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The treatment group is the high sustained attention individuals. The control group is the low sustained attention individuals.
The sample for log annual income and employed in white collar job includes individuals who are above 15 years of age. The sample for years of schooling includes individuals who are above 25 years of age.
The set of controls includes age, gender, and locality.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model assumes linearity for log of annual income and years of schooling results. For the employed in white collar job results RA model is also probit.
The reported results for NN and NN-PSM are for one nearest neighbor. The reported results for PSM use a caliper of 0.1.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8C – Robustness Checks

	Dependent Variable Log of Annual Income	Dependent Variable Employed in White Collar Job	Dependent Variable Years of Schooling
	ATET	ATET	ATET
Regression Adjustment (RA)	0.5140***	0.1701***	1.8028***
	(0.089)	(0.037)	(0.268)
Nearest Neighbor (NN)	0.3657***	0.1990***	1.8852***
	(0.110)	(0.045)	(0.362)
Nearest Neighbor PSM (NN - PSM)	0.4408***	0.1885***	1.8197***
	(0.132)	(0.045)	(0.356)
Propensity Score Matching (PSM)	0.4408***	0.1832***	1.8197***
	(0.132)	(0.048)	(0.356)
Inverse Probability Weights (IPW)	0.5127***	0.1693***	1.7968***
	(0.089)	(0.037)	(0.266)
IPW with Regression Adjustment (IPWRA)	0.5134***	0.1703***	1.8097***
	(0.089)	(0.037)	(0.266)
N	1,224	1,358	1,137

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The treatment group is the high sustained attention individuals. The control group comprises of the low attention individuals and those who passed the simple test but did not take the advance test.
The sample for log annual income and employed in white collar job includes individuals who are above 15 years of age. The sample for years of schooling includes individuals who are above 25 years of age.
The set of controls includes age, gender, family size, parents' education, locality, standardized height, distance (minutes) to nearest school, a survey year dummy, and the full set of interactions with locality.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model assumes linearity for log of annual income and years of schooling results. For the employed in white collar job results RA model is also probit.
The reported results for NN and NN-PSM are for one nearest neighbor. The reported results for PSM use a caliper of 0.1.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9 – Falsification Tests

MODELS	N	Regression Adjustment	Nearest Neighbor	Nearest Neighbor PSM	PSM	IPW	IPW with RA
Log Annual Income							
ATET	1,031	0.3044***	0.0794	0.0849	0.0849	0.2182*	0.3169***
		(0.112)	(0.119)	(0.124)	(0.124)	(0.132)	(0.090)
White Collar							
ATET	1,156	0.0243	0.0089	-0.0468	-0.0468	-0.0027	0.0045
		(0.033)	(0.037)	(0.046)	(0.046)	(0.038)	(0.032)
Educational Attainment							
ATET	944	0.2635	0.2793	0.3612	0.3612	-0.0501	0.0005
		(0.277)	(0.258)	(0.415)	(0.415)	(0.342)	(0.299)

Coefficients reported are the mean differences between high sustained attention and low sustained attention individuals.
The false treatment group is the low sustained attention individuals. The false control group is those who passed the simple test but did not take the advance test.
The samples for labor market outcomes include individuals above 15 years of age while the educational attainment sample includes individuals who are above 25 years of age.
The set of controls includes age, gender, standardized IQ, family size, parents' education, locality, pre-education reform cohort, standardized height, distance (minutes) to nearest school, school quality, a survey year dummy, and the full set of interactions with locality.
ATET refers to average treatment of the treated.
The propensity score model is probit. The RA model for white collar is also probit while log annual income and educational models assume linearity.
The reported results for NN and NN-PSM are for one nearest neighbor.
The reported results for PSM use a caliper of 0.2.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 1 - Sample Tests

Ravens Test

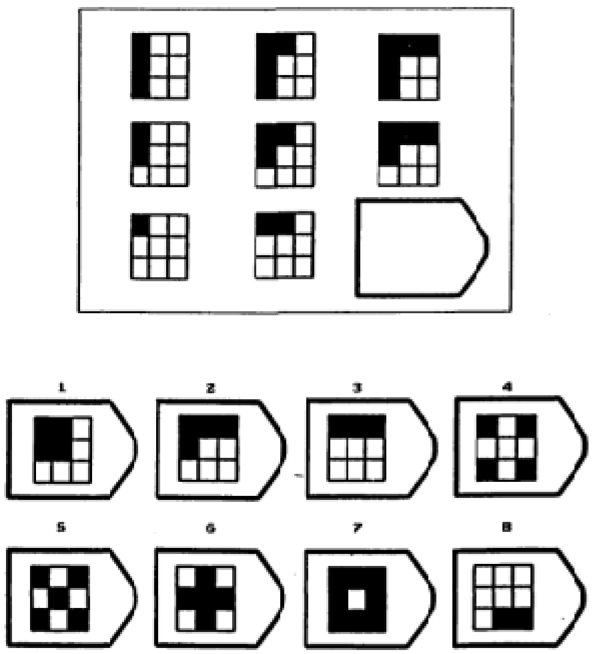


Figure 4a. A problem to illustrate the quantitative pairwise progression rule. The number of black squares in the top of each row increases by one from the first to the second column and from the second to the third column. The number of black squares aong the left remains constant within a row, but changes between rows from three to two to one (The correct answer is #3).

Simple English Reading Test

John is a small boy. He lives in a village with his brothers and sisters. He goes to school every week. In his school there are five teachers. John is learning to read at school. He likes to read very much. His father is a teacher, and his parents want him to become a school teacher too.

1. Who is John?

   (A) An old man
   (B) A small boy
   (C) A school teacher
   (D) A school

2. Where does John live?

   (A) In a village
   (B) In a city
   (C) In a school
   (D) In a forest

3. What does John do every week?

   (A) Works with his father
   (B) Plays with his friends
   (C) Helps his brothers and sisters
   (D) Goes to school

4. How many teachers are there at John's school?

   (A) One
   (B) Three
   (C) Five
   (D) Six

Simple Math Test

1. $1 + 2 =$
2. $5 - 2 =$
3. $2 \times 3 =$
4. $10 \div 5 =$
5. $24 \div 17 =$
6. $33 - 19 =$
7. $17 \times 3 =$
8. $41 \div 7 =$

Advance English Reading Test

Directions: For questions 10-15, read the passage below. Each line of the passage has a number. In each line, there is a box with four possible choices. Pick the choice that best completes the sentence in each numbered line. Mark the letter (A,B,C, or D) of the choice on your answer sheet.

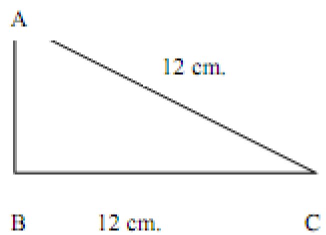
10. Sound is something we [(A) hears. (B) hearing. (C) heard. (D) hear.] It comes to your

11. [(A) Eyes (B) nose (C) ears (D) mouth] in different ways. It might be pleasant,

12. like the voice of a friend, [(A) when (B) as (C) or (D) since] unpleasant, like the yelp of a

13. dog that has been struck by a [(A) horn. (B) car. (C) road. (D) bridge.] Some sounds are loud,

14. and some are soft; some are high, and some are [(A) full. (B) low. (C) quite. (D) big.] Sound is



Note: figure not drawn to scale

21. If the perimeter of the triangle ABC is 30 centimetres, what is the length, in centimetres of side AB?

(A) $2\frac{1}{2}$

(B) 3

(C) 6

(D) 18

---

22. Two cities are 12 kilometres apart. Each day, a bus makes 3 round trips between these cities. How many kilometres does the bus travel each day?

(A) 72

(B) 36

(C) 1

(D) 4

---

23. A meal costs 1500 Cedis. If a 10% service charge is to be added to the bill, what would the total charge be?

(A) 1510 Cedis

(B) 1600 Cedis

(C) 1650 Cedis

(D) 2500 Cedis

---

24. An island has an area of about 300 square miles. The government reports that one third of the island is not suitable for cultivation. About how many square miles of this island are suitable for cultivation?

(A) 50

(B) 100

(C) 150

(D) 200

	Highest	Lowest
Elderet	23.6 °	9.5 °
Magadi	34.9 °	23.1 °
Nakura	26.4 °	10.1 °
Narok	24.4 °	8.3 °

25. The chart above shows the average (mean) high and low temperatures for four cities in a certain year. In which of the cities was there the greatest difference between the average high and the average low?

(A) Eldoret

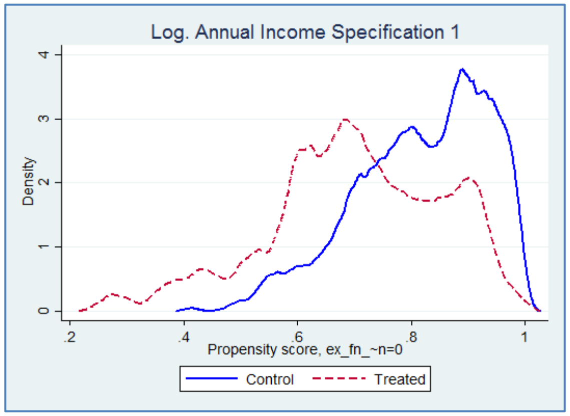
(B) Magadi

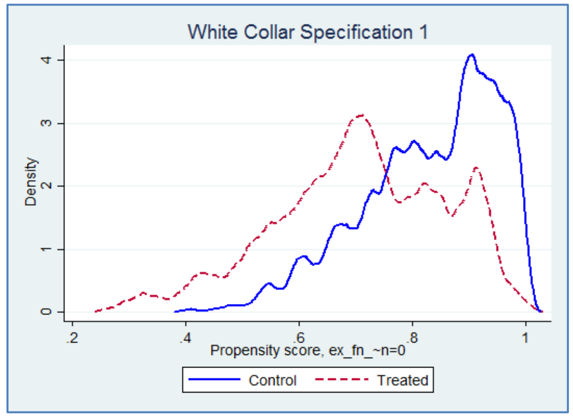
(C) Nakura

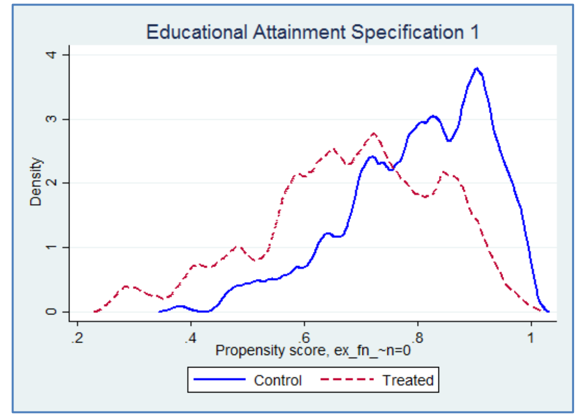
(D) Narok

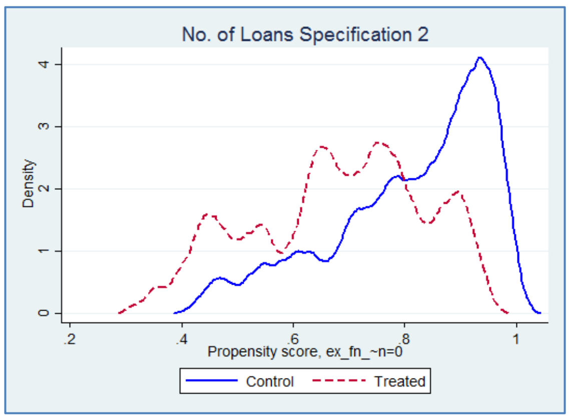
Appendix 2 – Overlap Plots for Matching

Results Regressions [12]





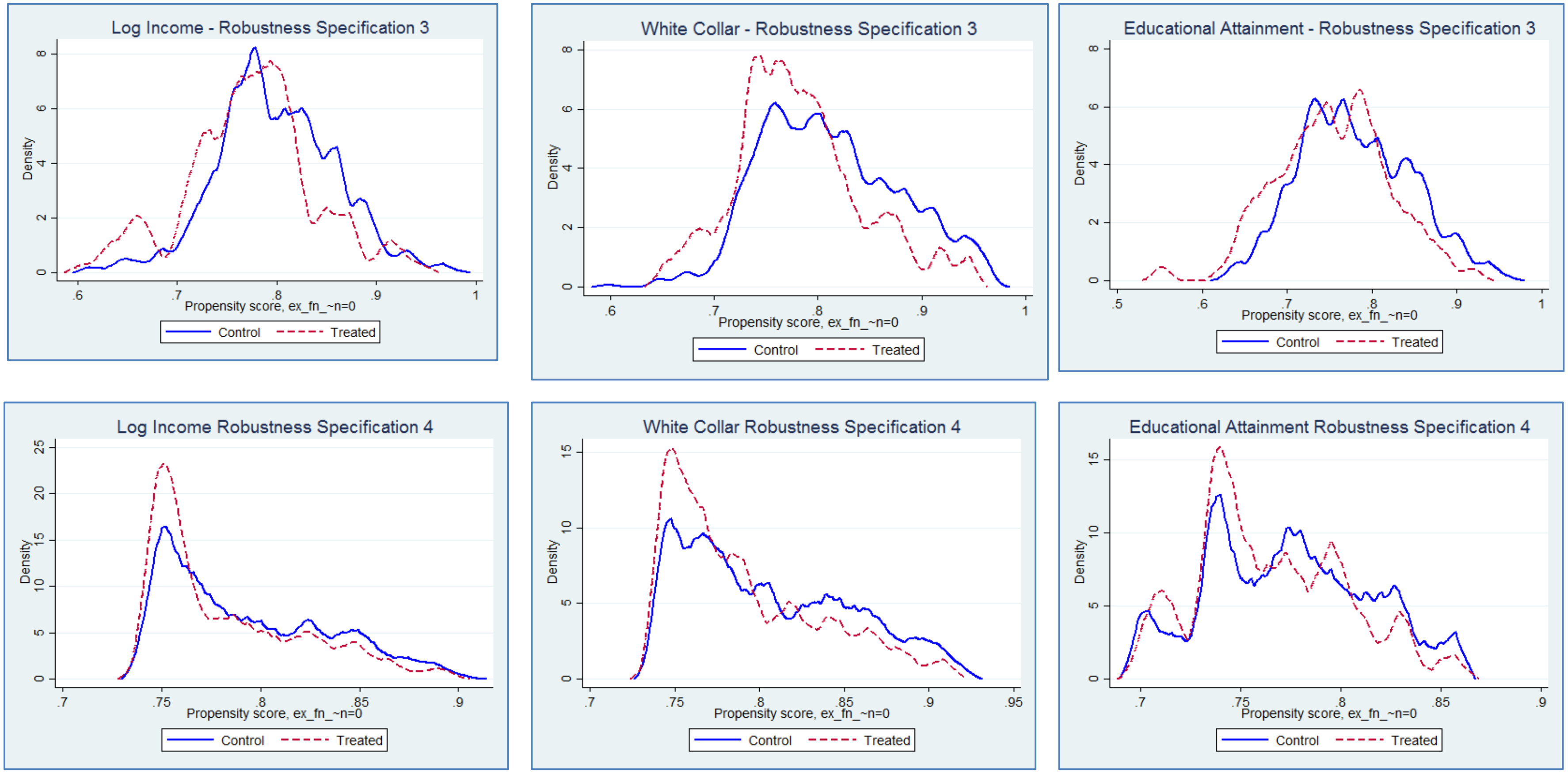




[12] Specification 1 controls for noncognitive ability, age, standardized IQ, gender, family size, parents' education, locality, standardized height, pre education reform cohort, distance (minutes) to nearest school, school quality, a survey year dummy, and a full set of locality interactions.
Specification 2 controls for noncognitive ability, age, gender, family size, standardized IQ, family size, locality, and standardized height.

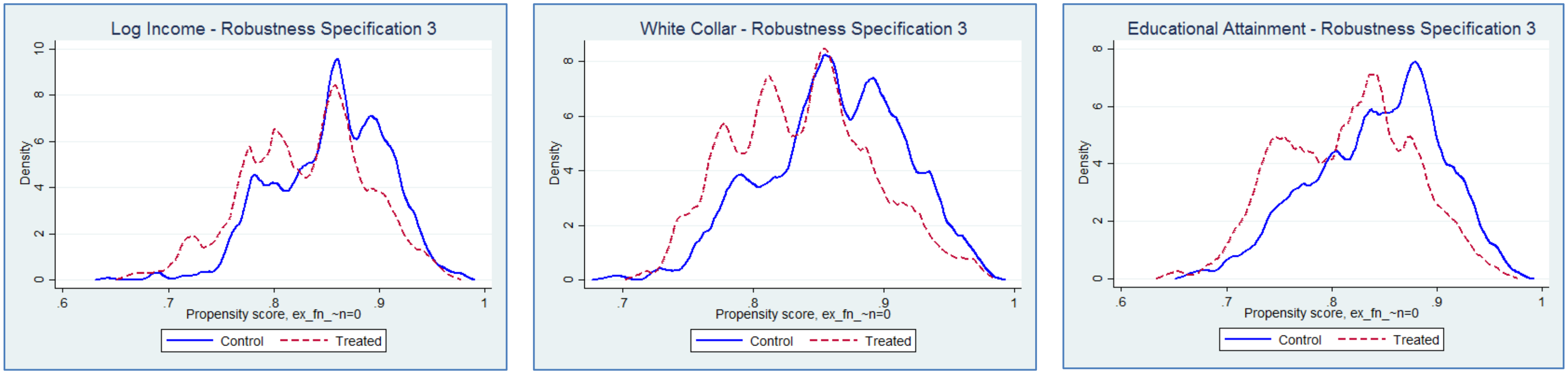
Robustness Checks [13]

With treatment group of high sustained attention individuals and control group of low sustained attention individuals.

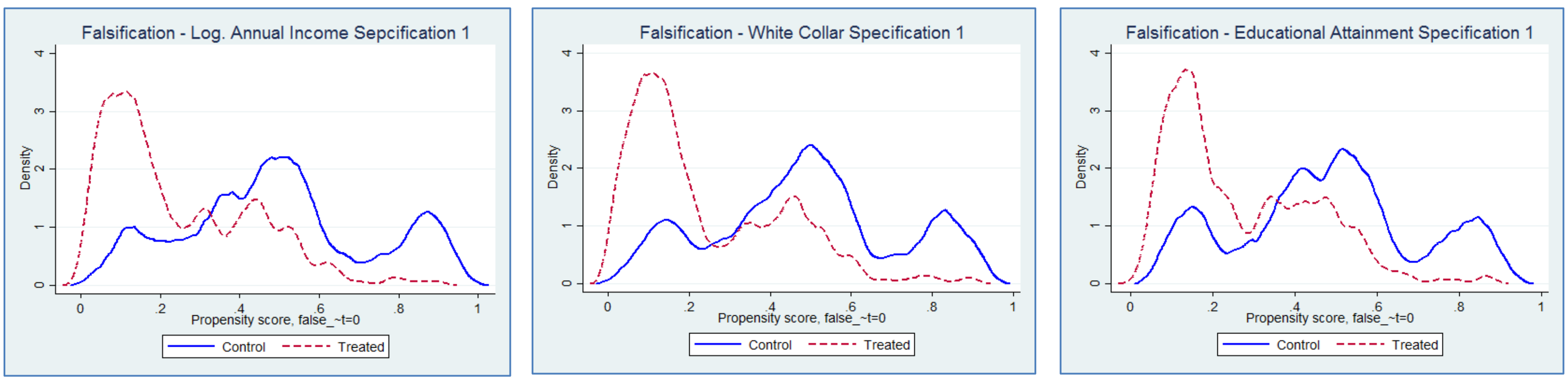


[13] Specification 3 controls for noncognitive ability, age, gender, family size, parents' education, locality, standardized height, distance (minutes to school), and a full set of interactions with locality.
Specification 4 controls for noncognitive ability, age, gender, and locality.

With treatment group of high sustained attention individuals and control group of low sustained attention individuals and those who passed the simple test but did not take the advance test.



Falsification Test[14]



[14] Specification 1 controls noncognitive ability, age, standardized IQ, gender, family size, parents' education, locality, standardized height, pre education reform cohort, distance (minutes) to nearest school, school quality, a survey year dummy, and a full set of locality interactions.