NBER WORKING PAPER SERIES

THE ECONOMIC VALUE OF BREAKING BAD: MISBEHAVIOR, SCHOOLING AND THE LABOR MARKET

Nicholas W. Papageorge
Victor Ronda
Yu Zheng

Working Paper 25602
http://www.nber.org/papers/w25602

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2019

We gratefully acknowledge helpful comments from: Robert Barbera, Pedro Carneiro, Sarah Cattan, Flavio Cunha, Seth Gershenson, Donna Gilleskie, Barton Hamilton, Hans von Kippersluis, Patrick McAlvanah, Robert Moffitt, Albert Park, Richard Spady, Sergio Urzua and Benjamin Williams along with seminar participants at the City University of Hong Kong, Tinbergen Institute, the Brookings Institution, Georgetown University, University of Western Ontario, Aarhus University, DePaul University, European University Institute, the Econometric Society World Congress, SOLE 2016, Barcelona GSE Summer Forum (Structural Microeconometrics). The usual caveats apply. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Economic Value of Breaking Bad: Misbehavior, Schooling and the Labor Market
Nicholas W. Papageorge, Victor Ronda, and Yu Zheng
NBER Working Paper No. 25602
February 2019
JEL No. I20,J10,J20

ABSTRACT

Prevailing research argues that childhood misbehavior in the classroom is bad for schooling and, presumably, bad for labor market outcomes. In contrast, we argue that some childhood misbehavior represents underlying socio-emotional skills that are valuable in the labor market. We follow work from psychology and categorize observed classroom misbehavior into two underlying latent factors. We then estimate a model of educational attainment and earnings outcomes, allowing the impact of each of the two factors to vary by outcome. We find that one of the factors, labeled in the psychological literature as externalizing behavior (and linked, for example, to aggression), reduces educational attainment yet increases earnings. Unlike most models where socio-emotional skills that increase human capital through education also increase labor market skills, our findings illustrate how some socio-emotional skills can be productive in some economic contexts and counter-productive in others. Policies designed to promote human capital accumulation could therefore have mixed effects or even negative economic consequences, especially in the case of policies that target socio-emotional skill formation for children or adolescents which are aimed solely at improving educational outcomes.

Nicholas W. Papageorge
Department of Economics
Johns Hopkins University
3400 N. Charles Street
Baltimore, MD 21218
and IZA
and also NBER
papageorge@jhu.edu

Yu Zheng
School of Economics and Finance
Queen Mary University of London
Mile End, London E1 4NS
Great Britain
and CEPR
yu.zheng@qmul.ac.uk

Victor Ronda
Aarhus University,
Department of Economics and Business Economics
Fuglesangs Alle 4, Office C9
Aarhus, 8210
Denmark
ronda@econ.au.dk

1 Introduction

Economists generally recognize that human capital consists of multiple skills that drive educational and labor market outcomes. An early contribution is Willis and Rosen (1979), who distinguish between academic and manual skill. More recently, a burgeoning literature in economics has extended the concept of human capital to incorporate socio-emotional skills such as perseverance and grit (Heckman and Rubinstein, 2001).[1] It is not controversial that returns to skills can differ across sectors and that some skills are more productive in schooling than in work or in one occupation than in another. For example, to explain career choices, Willis and Rosen (1979) emphasize variation in the returns across occupations to manual versus academic skill.[2]

Despite potential differences in returns, the skills that constitute human capital are all typically seen as enhancing productivity — both in school and on the labor market. This is likely true for cognition and for many socio-emotional skills such as grit, which captures goal setting (Duckworth et al., 2007). However, this view overlooks how some components of human capital could be productive in some economic contexts but could actually be counter-productive in others. If so, then policies designed to promote human capital accumulation could have mixed effects or even negative economic consequences. This is especially the case for policies that target socio-emotional skill formation aimed at children or adolescents, for whom socio-emotional skills have been shown to be relatively malleable (Heckman and Kautz, 2014).

In this paper, we demonstrate that some components of childhood misbehavior predict higher earnings even though they are associated with lower schooling attainment. We examine a widely-studied pair of socio-emotional skills known as *externalizing behavior* and *internalizing behavior*.[3] Externalizing behavior is linked to aggression and hyperactivity while internalizing behavior captures anxiety, depression, shyness, unassertiveness and fearfulness (Ghodsian, 1977; Duncan and Magnuson, 2011; Duncan and Dunifon, 2012). Using a longitudinal dataset from Britain, the National Child Development Survey (NCDS), we estimate an econometric model relating childhood misbehavior to educational attainment and labor market outcomes. We approximate schooling, hours of work and wages using linear-in-parameters equations, and we model correlation across equations as unobserved heterogeneity in the form of three latent factors identified using a measurement system. The first

[1] Excellent summaries of this research are found in Borghans et al. (2008) and Almlund et al. (2011).

[2] This point has its origins in Roy's model of selection into occupations (Roy, 1951).

[3] Regarding the nomenclature: "externalizing behavior" and "internalizing behavior" describe the two socio-emotional skills (sometimes called noncognitive skills) that are measured using teachers' reports of childhood maladjustment or misbehavior.

two latent factors capture the socio-emotional skills described above and are measured using multiple teachers' reports of children's misbehavior or maladjustment among schoolchildren. The third factor captures cognition and is measured using math and reading test scores. We also estimate the model separately for males and females. The key empirical fact we establish is that, for both genders, one of the factors underlying observed classroom misbehavior, externalizing behavior, lowers educational attainment, but is also associated with higher earnings. This result holds across several datasets and using a variety of econometric specifications, including different assumptions on how teacher measurements of misbehavior relate to latent factors. In other words, we provide robust evidence demonstrating that a penchant for *breaking bad* can be good.[4]

Our results have implications for our understanding of the skills comprising human capital. Whereas previous work has recognized variation in skill prices across economic sectors, our findings on externalizing behavior go further, demonstrating that a single skill can be productive in some economic contexts and counter-productive in others. Identifying a skill that raises earnings but lowers educational attainment is a particularly striking illustration since it runs counter to the typical view of ability bias in estimates of the returns to education (Becker, 1967). Often, the presumption is that the unobserved skills leading to success in education also promote earnings.[5] In line with this assumption, among individuals in our sample, we demonstrate that schooling predicts higher earnings; that internalizing behavior predicts lower education attainment and lower earnings; and that cognition predicts higher degrees and higher earnings. In contrast, externalizing behavior has mixed effects. Despite its negative impact on schooling, it is also associated with higher wages for males and females as well as more hours in the labor market for females.[6]

Building on our main findings, we address three questions. The first concerns mechanisms and asks whether externalizing behavior is a direct determinant of earnings or whether it influences some observable third variable or variables which then affect earnings. This would suggest that externalizing is not productive *per se*, but instead leads to higher earnings through intermediate choices and outcomes, such as occupational sorting, marriage or fer-

[4]According to www.urbandictionary.com the definition of the term *breaking bad* is to "challenge conventions" or to "defy authority". *Breaking Bad* is also the title of an American television show in which the protagonist is an unsuccessful chemist who reveals a striking talent for producing illicit drugs. The show offers an extreme example of how certain skills or behaviors may lead to low productivity in one sector and high productivity in another.

[5]There are a number of exceptions. For example Card (2012) shows that IV estimates could lead to larger coefficients on education in wage equations. The argument is based on heterogeneity in treatment effects coupled with the particular group for whom the IV affects attendance.

[6]Several studies have examined the relationship between these two behaviors to better known measures like the "Big 5" personality traits. As we explain below, our key results hold even if we control for personality traits, suggesting that despite correlations, the skills we examine in this study are distinct.

tility. For example, high-externalizing individuals may have relatively strong preferences for lucrative occupations that low-externalizing individuals tend to avoid.[7] A more concerning possibility is that positive returns to externalizing are driven by selection into employment. For example, if high-externalizing individuals work only if they are highly productive due to omitted factors, then estimates of positive labor market returns to externalizing behavior could be an artifact of differential selection into employment. To examine these possibilities, we assess how externalizing behavior relates to labor supply along with occupational choices, work experience, fertility and marriage. While we show evidence that externalizing behavior is strongly related to some of these outcomes, we also demonstrate that these relationships do not drive our main finding that externalizing behavior, despite being unproductive at school, is productive in the labor market.

A second important question is whether the returns to socio-emotional skills vary across socioeconomic groups. This question is partly motivated by the findings of Heckman, Pinto, and Savelyev (2013), who show that an early childhood intervention (the Perry Preschool Program) raised earnings and that about 20% of this rise is attributable to a reduction in externalizing behavior. In contrast, we find that, for a 1958 British cohort, externalizing behavior raises earnings. To explore this difference, we consider a sub-sample of the NCDS British cohort that is selected to mimic the financially disadvantaged group studied in Heckman, Pinto, and Savelyev (2013). We show that among individuals who grew up in poverty, externalizing behavior carries no significant earnings premium. This finding is in line with Lundberg (2013), who demonstrates that the payoff to socio-emotional skills is context-dependent and may vary by socioeconomic status. One possible reason is selection into criminality (Aizer, 2009; Heckman, Pinto, and Savelyev, 2013). However, for our sample, we find that differential sorting into police involvement does not appear to drive differences in returns to externalizing behavior across socioeconomic groups. Therefore, we cannot rule out the possibility that some skills are simply priced differently in the labor market depending on an individual's background. This is particularly troubling since it suggests that individuals who are already disadvantaged due to childhood poverty are excluded from realizing the full returns to skills or attributes that are lucrative for individuals born into wealthier families.

A third question is whether our findings on externalizing are unique to the dataset we use in our main analyses, which follows a single cohort over time and in one country. This would raise the concern that our findings are attributable to factors unique to the individuals that we study, such as labor market shocks occurring in a particular time and place. This would

[7] Becker et al. (2012) examine links between preferences and widely-used measures of socio-emotional skills, in particular, personality traits.

undermine the external validity and limit the policy relevance of our findings. To address this question, we examine the 1970 British Cohort Study (BCS), the National Education Longitudinal Study of 1988 (NELS), the Panel Study of Income Dynamics (PSID), and the National Longitudinal Survey of Youth (NLSY) 1979: Children and Young Adults (CNLSY). We show that our basic results extend to these datasets, which provides compelling evidence that our findings — including the mixed effects of externalizing — are not unique to one group or era, but instead reflect an empirical regularity found across cohorts and countries.

Our findings illustrate the general point that some of the skills, characteristics or attributes that are valuable during childhood may not lead to success in adulthood. In our case, mixed effects of externalizing also suggest a mismatch between the types of skills promoted in school and the skills that are valuable on the labor market. Schooling environments designed to achieve one set of goals, such as obedience, order or conformity, may stifle or penalize skills or attributes that create value in the labor market. Relatedly, our findings also suggest that a productive labor market skill may be easily overlooked or difficult to detect or foster since it is not productive in school. As a result, educational attainment or certification is a potentially flawed signal of a future worker's productivity. This point echoes findings in Heckman and Rubinstein (2001), who show that the GED is a "mixed" signal of productivity since it is taken by high school dropouts with low socio-emotional skill.

More generally, our findings illustrate that broadening our understanding of what constitutes human capital — which the literature on socio-emotional skills has done — also opens up the possibility that some human capital investments can have negative economic returns in some sectors. Despite the positive returns to educational attainment, investments designed to curb or eliminate childhood misbehavior may be ill-conceived or short-sighted since a subset of children who misbehave may be expressing socio-emotional skills that are valuable in the labor market. This is not a hypothetical concern since school districts are increasingly poised to begin using high-stakes tests to evaluate students, teachers and schools based on character or socio-emotional skills (West et al., 2016). Finally, our findings on group differences imply further difficulties in evaluating human capital investments involving children's socio-emotional skills since the returns to skills can differ not only by the economic context in question, but also by socioeconomic status.

The paper is organized as follows. In Section 2, we introduce the NCDS dataset, discuss measurements of misbehavior that identify externalizing and internalizing behavior and conduct a preliminary data analysis. In Section 3, we describe the main "benchmark" econometric model we estimate, including the measurement system used to identify latent skills, along with estimation. In Section 4, we present main results. Section 5 reports results on a host of sensitivity analyses, with particular focus on alternative assumptions to identify the measurement system mapping latent factors to observed measures of misbehavior. Section 6 conducts subgroup analyses to assess whether our main results extend to individuals who grew up in poorer households. Section 7 examines additional datasets to assess whether our main empirical results extend to other countries and cohorts. Section 8 offers brief remarks on potential policy implications and concludes.

2 Data and Preliminary Analysis

In this section, we introduce the NCDS dataset, describe key variables used in our analysis and provide estimates from a preliminary econometric model relating childhood misbehavior with schooling and earnings. We demonstrate that once we treat externalizing and internalizing behaviors separately, externalizing behavior is associated with higher earnings even though it also predicts lower educational attainment.

2.1 The National Child Development Study

The NCDS is an ongoing longitudinal survey that follows the universe of individuals born in the same week in 1958 in Great Britain. It is particularly well-suited for our study since it collects teachers' reports of classroom misbehavior for a large sample of children and then follows these children through adulthood. Therefore, the dataset allows us to relate misbehavior in elementary school to educational attainment along with labor market outcomes. To date, there have been nine surveys to trace all the members of the cohort still living in Great Britain. Surveys occurred when subjects were born and when they were aged 7 (1965), 11, 16, 23, 33, 42, 50 and 55 (2013).

We focus on information gathered at birth and in the first five sweeps, covering ages 7 to 33. The NCDS initially contained information on 18,555 births. In constructing our analytic sample, we keep respondents with valid information on test scores and classroom misbehavior at age 11 and educational attainment and labor outcomes at age 33. We drop individuals with missing information on variables treated in some of our analyses as intermediate outcomes, such as relationship status, fertility, employment status and employment history. We also drop individuals who are reported as employed but have missing information on earnings at age 33. We impute data for individuals missing information on variables used in some specifications as controls, such as parents' education and occupation. The resulting analytic sample has information on 7,241 individuals, of whom 3,573 are males and 3,668 are females.[8]

[8] Most of the drop in observations is due to attrition at the fifth survey. Out of the original 18,555 births, only 11,364 individuals were surveyed in 1991 at age 33. To assess whether sample attrition drives our main

2.2 Key Variables and Summary Statistics

2.2.1 Education and Labor Outcomes

In the UK, schooling is compulsory until age 16. Thereafter, students can leave school without any qualifications (no certificate), study for an exam to obtain a Certificate of Secondary Education (CSE) or study towards obtaining the Ordinary Levels (O-Levels), where the latter are more academically demanding.[9] Individuals aiming to attain a higher degree take another set of examinations, the Advanced Levels (A-Levels). Students who are successful in their A-Levels are able to continue to attain either a higher-education diploma (after two years of study) or a bachelor's degree (after three years of study). At the postgraduate level, students can obtain a higher degree: Master of Philosophy (MPhil) or Doctor of Philosophy (PhD). In summary, individuals in our sample can sort into six mutually exclusive schooling levels: no certificate, CSE, O-Levels, A-Levels, higher education (including diploma and bachelors) or higher degree (including MPhil and PhD).

Summary statistics on education, labor market outcomes and a basic set of controls are found in Table 1. 51% of our sample is female. Females in our sample are less educated compared to males. On average, employed females' wages are 29% lower — and hours are 51% lower — than those reported by employed males. Males are also significantly more likely to be employed and, conditional on employment, to be self-employed. In general, large gender differences in schooling and labor market outcomes suggest that we should allow the parameters of our econometric model to vary by gender.

2.2.2 Socio-Emotional Skills and Cognition

Next, we discuss variables used to construct measures of unobserved skills, including the two socio-emotional skills that are the focus of our analysis, along with cognition. We measure socio-emotional skills using variables describing classroom misbehavior. When a child in the sample was 11 years old, the child's teacher was asked to complete an inventory listing the child's behaviors in the classroom. The teacher was given a list of roughly 250 descriptions of specific behaviors and asked to underline the items which best describe the child. These descriptions include statements such as: "too timid to be naughty," "brags

results, we compare our analytic sample to the sample of all individuals observed at age 11, which we call the "full sample." Compared to the full sample, our analytic sample is slightly more educated, less likely to be self-employed, receives slightly lower wages and works fewer hours. However, none of these differences is statistically significant. Summary statistics for the full sample are reported in Tables S1 and S2 in Appendix A, where we provide additional summary statistics for variables and samples used throughout this study.

[9] CSEs and O-Levels were replaced by the General Certificates of Secondary Education (GCSE) in 1986 after individuals in our sample had finished their schooling.

to other children," "normally honest with school work," "adopts extreme youth fashions," and "has stolen money." Completed inventories were then used to compute scores on a set of ten summary variables known as the Bristol Social Adjustment Guide or *BSAG* maladjustment syndromes.[10] The ten syndromes are: hostility towards adults, hostility towards children, anxiety for acceptance by adults, anxiety for acceptance by children, restlessness, inconsequential behavior, writing off adults and adults standards, depression, withdrawal, and unforthcomingness. The syndromes have been used since their introduction in Stott, Sykes, and Marston (1974) to assess the psychological development of children.

In Table 2, we present averages for each BSAG maladjustment syndrome separately by gender. Values range from 0 to 15, with a higher value indicating a higher prevalence of a particular maladjustment syndrome. The means are usually low due to a clustering around zero and fairly low values in general. Overall, females appear to misbehave less frequently than males. Specifically, males exhibit higher scores for all of the BSAG variables except for "anxiety for acceptance by adults." Gender differences in misbehavior are consistent with earlier findings documented for Great Britain (Duncan and Magnuson, 2011; Duncan and Dunifon, 2012) and the U.S. (Bertrand and Pan, 2013).

Following earlier work (see e.g., Cunha, Heckman, and Schennach (2010)), we measure cognitive skill using a set of math and reading test scores. Test score averages are found in Table 2. These tests are administered when children are 11 years old. According to the table, girls score marginally higher than boys on tests of verbal and non-verbal ability, where non-verbal ability measures identification of shapes and symbols. In contrast, average math scores for boys are marginally higher.

The benchmark econometric model used in our main analysis, described in Section 3, includes a measurement system that uses these observed maladjustment syndromes and test scores as measurements to identify unobserved skills. In contrast, for the preliminary analysis, we use the variables described above to construct crude measures of the unobserved skills. To construct these measures of socio-emotional skill, we follow Ghodsian (1977), who proposed dividing up the BSAG syndromes into two groups based on apparent differences among what behaviors the syndromes capture.[11] Variables assigned to each group are then summed to create two new variables. The first variable, *externalizing behavior*, is constructed from summing over maladjustment syndromes such as "hostility towards adults"

[10]In particular, each item on the inventory was assigned to one of 10 syndromes and the variables are the sum of these items from the teacher inventories. Unfortunately, the original teacher inventory data are not available. If they were, one could use them directly to identify latent skills.

[11]This division proposed in Ghodsian (1977) is also motivated by a principle components factor analysis, which suggests there are two underlying latent factors measured by the BSAG syndromes. We replicate this analysis in Appendix B.

and "restlessness" among others, and expresses anxious, aggressive, and outwardly-expressed behavior. The second variable, or *internalizing behavior*, is constructed by summing over maladjustment syndromes such as "depression" and "withdrawal" among others, and expresses withdrawn and inhibited behavior. Similarly, we obtain a measure of cognitive ability by summing test scores. How we assign measurements to each of the three skills is summarized in Table 3.[12] In addition, we construct a generic measure of misbehavior by simply summing up all ten syndromes. This variable is used to illustrate how findings change once we recognize that misbehavior captures two separate socio-emotional skills. Finally, we normalize these newly constructed crude measures of externalizing behavior, internalizing behavior, cognition, and misbehavior, so that each variable has mean equal to zero and variance equal to one for the full sample. Summary statistics for these measures are reported in Table 2 separately by gender. According to the table, boys exhibit significantly higher externalizing and internalizing behaviors compared to girls. Boys are roughly 0.3 standard deviations higher on average. We also find that average cognition for girls is about 0.06 standard deviations higher than it is for boys.

2.2.3 Additional Control Variables

There are three sets of additional control variables that we use in our subsequent analyses. Table 4 summarizes which additional variables are included in which equations. Conditioning on these variables helps to mitigate concerns related to omitted variables bias, but it is important to state at the outset that their inclusion does not drive our key findings. In both our preliminary analysis using crude measures of skills, as well as our benchmark econometric model that features a formal measurement system, we obtain our main results once we include measures of cognition, externalizing behavior and internalizing behavior.[13]

The first set of additional variables are two basic controls, which are included in all schooling and outcome equations. The first is an indicator for childhood poverty. The variable we construct, "Financial Difficulty," takes the value one if (i) the interviewer reported that the household appeared to be experiencing poverty in 1965 or (ii) a member of the household self-reported having financial difficulties in the 12 months prior to being observed

[12]These measures have been externally validated in the sense that they are positively correlated with a range of other measurements of social maladjustment from teachers, professional observers, parents and peers (Achenbach, McConaughy, and Howell, 1987). Moreover, they have been studied extensively by psychologists researching child development and, of late, by some economists (Blanden, Gregg, and Macmillan, 2007; Aizer, 2009; Agan, 2011; Heckman, Pinto, and Savelyev, 2013). Both Aizer (2009) and Agan (2011) study how externalizing behavior is linked to anti-social and criminal activity. For general surveys of research on externalizing and internalizing behaviors, see Duncan and Magnuson (2011) and Duncan and Dunifon (2012).

[13]Summary statistics for additional control variables are found in Table S3 in Appendix A.

in either 1969 or 1974, and zero otherwise. We include this variable in all equations because it is a variable along which we stratify our sample in one of our subsequent analyses (see Section 6).[14] The second basic control is an indicator variable for living in London. Including this variable is common practice using the NCDS given possible London-specific differences in schooling or labor outcomes.[15] Summary statistics for the financial difficulty and London dummy variables are found in Table 1. 36% of the sample lives in or around London before age 16 versus 30% at age 33. 16% of our sample experienced financial difficulty in their childhood.

We include a second set of control variables in schooling equations, but not in other equations (Heckman, Humphries, and Veramendi, 2018). The reasoning is that externalizing behavior could capture a productive skill on the labor market, but could also relate to family backgrounds that lead to lower schooling, such as an absent father or low parental education. If so, an estimated negative impact on schooling may simply reflect omitted family background variables rather than mixed effects of a socio-emotional skill. To address this concern, we include a set of family background variables, which are excluded from the outcome equations: whether the mother studied beyond the minimum schooling age, whether the father studied beyond the minimum schooling age, whether the father's information is missing, father's occupation, and mother's employment status, all observed when the child is age 11.

A third set of control variables related to school characteristics is included in our measurement system to address possible mis-reporting differences across teachers and schools. We postpone a discussion of this final set of control variables until we introduce the measurement system in Section 3.

2.3 Relating Misbehavior, Schooling and Earnings

Our preliminary analysis relates the crude measures of externalizing behavior, internalizing behavior, and cognition to schooling and labor market outcomes. An advantage of the preliminary analysis is that this approach has been taken in previous studies, which means

[14]Tables S4 and S5 in Appendix A provide summary statistics for samples stratified by the financial difficulties variable.

[15]In the NCDS, the definition of region of residence changed from the first 4 surveys (ages 0, 7, 11 and 16) to the fifth (age 33) survey. Before age 16, we say an individual lives in or around London if he or she lives in East, South East or South England. At age 33, we say an individual lives in or around London if he or she lives in South East England. The reason is that the categories change across surveys. 57%, 85% and 72% of individuals living in East, South East, or South England at age 11 are living in South East England at age 33. Individuals in these regions have higher earnings on average than individuals living in other regions. The results are not sensitive to changes in the classification or whether we include dummies for all the possible regions of residence.

we can directly compare our findings to those in earlier work. In particular, we can show that securing our key results — including the finding that externalizing behavior has mixed effects on schooling and earnings — does not require a more sophisticated measurement system, but emerges once we control for measures of internalizing behavior and cognition as they have been constructed in earlier work. Earlier work includes research using the NCDS dataset studying externalizing and internalizing behaviors (Farmer, 1993, 1995; Jackson, 2006). It also includes research using different samples since the division of misbehavior into these two socio-emotional skills extends to other datasets, including the CNLSY and the PSID (Yeung, Linver, and Brooks-Gunn, 2002; Agan, 2011). Finally, using crude measures facilitates a comparison of empirical patterns across datasets, which we perform in Section 7. The reason is that other datasets often contain summary measures of externalizing and internalizing behaviors, and therefore we cannot always apply the same type of measurement system used in our benchmark econometric model estimated from the NCDS data. As we discuss in Section 3 when introducing the benchmark econometric model, use of these crude measures imposes a number of unattractive assumptions that the formal measurement system allows us to relax.

For the preliminary analysis, we explain educational attainment using an ordered probit model. The outcome variable is one of the six possible schooling levels.[16] Formally, defining s_i^* as a latent variable determining schooling, we estimate regressions of the following form:

$$s_i^* = E_i\psi^{\mathbf{E}} + I_i\psi^{\mathbf{I}} + C_i\psi^{\mathbf{C}} + Z_i'\beta_s + e_i^S \tag{1}$$

where observed schooling $s_i = s$ if $\mu_L^s \leq s_i^* < \mu_H^s$ and μ_L^s and μ_H^s are the particular bounds for schooling level s. E_i and I_i are the crude measures of externalizing and internalizing behaviors and C_i is a crude measure of cognition, constructed according to the description in Section 2.2.2. Recall, we have normalized the measures of unobserved skills. Abusing notation somewhat, Z_i is a vector of control variables, which varies across specifications. Finally, e_i^S is a normally distributed disturbance.

Estimates of equation (1) are reported in Table 5. We start by regressing schooling on the crude measure of generic misbehavior and estimate a negative relationship in Columns [1]-[2]. The magnitude declines when we include cognition, suggesting a negative correlation between the two variables. In Columns [3]-[6], we allow externalizing and internalizing behaviors to have separate effects on schooling. In Column [3], we start by only including externalizing behavior. In the subsequent three columns, we add our measures of cognition

[16]We use an ordered probit in our preliminary analysis to simplify exposition. However, results are robust to using a more flexible specification, such as a multinomial logit or probit model. In the benchmark econometric model used in our main analysis, we estimate a multinomial logit model.

and of internalizing behavior, and then both. Results in Column [6], including all three measures of skills, show that both externalizing and internalizing behaviors independently lower schooling attainment, while a higher level of cognition leads to higher educational attainment. Moreover, the impact of cognition is roughly ten times larger than the impacts of either socio-emotional skill. While models in Columns [1]-[6] include indicators for London, financial difficulties and female, in Column [7] we also include family background variables (the second set of additional control variables described above). Most affect schooling in ways we would expect. For example, higher parental education has a positive impact on the respondent's own education. However, including these variables has very little effect on the size of the coefficients on cognitive and socio-emotional skills.

In Columns [8] and [9], we use the same set of regressors as in Column [7], but stratify the sample by gender. Across genders, externalizing and internalizing behaviors have a negative impact on schooling, while cognition strongly raises educational attainment. Comparing genders, the negative coefficient on externalizing is larger for males, while the coefficient on internalizing is larger for females. Patterns are similar if we compute marginal effects, which are reported in Table S6 in Appendix A.

To explain earnings, we regress log weekly earnings at age 33, conditional on being employed, onto measures of socio-emotional and cognitive skills (Table 6).[17] Defining y_i as log earnings at age 33 for individual i, we estimate OLS regressions of the following form:

$$y_i = E_i\phi^{\mathbf{E}} + I_i\phi^{\mathbf{I}} + C_i\phi^{\mathbf{C}} + X_i'\beta + e_i^Y \tag{2}$$

where X_i includes the basic set of controls (indicators for female, financial difficulties and living in London) and may or may not include schooling outcomes. Columns [1] and [2] contain estimates using the single measure of misbehavior, controlling for cognition or not. In line with previous research (e.g., Segal (2013)), we find this single measure of misbehavior is associated with both lower schooling and lower earnings.

Results change dramatically when we view childhood misbehavior as reflecting two distinct factors and control for cognition. In Columns [3]-[6] of Table 6, we regress log earnings onto E_i and I_i separately. The positive price of externalizing behavior emerges as soon as we control for internalizing behavior and cognition (Column [6]). Comparing Column [4] to [6], we deduce that if externalizing and cognition are negatively correlated while cognition is valuable for earnings, then omitting cognition will downwardly bias the estimated impact of externalizing. Similarly, if externalizing and internalizing are positively correlated while

[17]As with schooling, our choice to use earnings as the outcome variable is for ease of exposition. In our main benchmark model, we allow externalizing to have separate effects on wages and hours.

internalizing is bad for earnings, then omitting internalizing again generates a downward bias on estimated returns to externalizing behavior (Columns [5] and [6]). These results suggest that it is important to account for the correlation patterns across all three skills to produce estimates interpretable as skill prices. In our benchmark econometric model described in Section 3, we thus allow arbitrary correlations across the latent factors capturing underlying skills.

Column [6] presents strong initial evidence that externalizing behavior carries an earnings premium. When we control for schooling outcomes in the earnings equation (Column [7]), the positive coefficient on externalizing behavior becomes even larger. When we further separate the sample by gender (Columns [8] and [9]), we conclude that mixed effects of externalizing behavior hold for both males and females in our sample.

The results from the preliminary analysis presented in Tables 5 and 6 provide initial evidence that a socio-emotional skill that is productive on the labor market is not productive in school.[18] It is also worth highlighting that, according to Table 5, the coefficient on externalizing is positive whether or not we control for schooling. An alternative possibility would be that externalizing behavior predicts higher earnings only after we have controlled for its negative impact on schooling. Such a finding would still support the idea that externalizing is potentially valuable in the labor market. However, it would also suggest that lower levels of externalizing behavior could have a positive net effect on labor market outcomes since the negative effect of externalizing through schooling on earnings would overwhelm the direct positive effect on earnings. In contrast, estimates suggest that externalizing behavior has a positive net effect on earnings despite having a negative impact on schooling.

The positive association between externalizing behavior during childhood and adult earnings has generally not been recognized in previous literature on the economic consequences of childhood misbehavior.[19] There are several reasons for this lack of recognition. First,

[18]This preliminary evidence is robust to a host of alternative specifications, which are explored in Appendix A. One alternative is to measure earnings at age 42 or 50, which yields similar results (Appendix A.5). We continue to use labor market outcomes at age 33 since otherwise we lose a considerable number of observations due to sample attrition as the NCDS cohort ages. We also show explicitly that the positive relationship between externalizing behavior and earnings emerges as soon as we control for internalizing behavior and cognition, and does not required any additional controls (Appendix A.3). We also explore potential non-linear effects and complementarities between factors and find no evidence of either (Appendix A.4). Finally, we report estimates where additional socio-emotional skills are included, in particular, the Big 5 personality traits (Appendix A.6). The externalizing premium decreases by about 20% when we control for the Big 5 personality traits. However, the Big 5 were measured at age 50 after earnings and education were realized, which could introduce bias due to simultaneity. In Section 4.3, we discuss this point in greater detail.

[19]Levine and Rubinstein (2017) recognize an empirical pattern that is similar in spirit. They show that individuals who engage in illicit behaviors as teenagers report high earnings in self employment. One possible extension to our research would be to assess whether the successful entrepreneurs they identify were high-

most of the literature on the long run effects of childhood misbehavior takes for granted that externalizing is broadly unproductive, focusing instead on negative impacts on school-related outcomes (Bertrand and Pan, 2013). This may be a result of data limitations since linking childhood misbehavior to labor market outcomes requires a long panel spanning from childhood well into adulthood. However, even studies using the NCDS dataset have not linked externalizing behavior to earnings (Farmer, 1993, 1995; Jackson, 2006).

Second, many studies use a single aggregated measure of childhood misbehavior or maladjustment. We discuss two such studies which are otherwise similar to ours, highlighting the importance of recognizing that misbehavior reflects distinct socio-emotional skills with potentially different returns in the labor market. Similar to our paper, Fronstin, Greenberg, and Robins (2005) uses the NCDS to study the effect of childhood maladjustment on labor market outcomes. Importantly, to justify the use of a single aggregated measure of misbehavior, the authors refer to earlier work showing that externalizing and internalizing behaviors have a similar effects on mental health in early adulthood, which might suggest similar effects on other outcomes (Chase-Lansdale, Cherlin, and Kiernan, 1995). In contrast, we show that the two factors have opposite effects on earnings.

Another related paper, Segal (2013), uses the National Education Longitudinal Survey (NELS) to relate five different teacher-reported measures of childhood misbehavior to education and labor market outcomes. The author shows that a variable that summarizes five measures of "misbehavior" predicts lower earnings. However, when the five measures are included individually in the same regression, the coefficient for one of the five measures, "disruptiveness," is positively related to earnings. Segal (2013) argues that the positive effect of disruptiveness on earnings is spurious since the association reverses when the other four measures are excluded from the regression (see Footnote 32 on p. 23 of the study). In contrast, we argue that these differences in estimates highlight the importance of including multiple measures of possibly correlated variables capturing misbehavior. We also show that summing multiple measures potentially obfuscates how each skill underlying misbehavior can have different effects on economic outcomes.

3 Model

Summing the BSAG maladjustment syndromes and test scores to create crude measures of underlying skills is simple and straightforward, but also imposes a number of unattractive assumptions. For example, each measurement is assigned to only one underlying skill. For

externalizing children.

example, externalizing behavior is assumed to have no effect on cognitive test scores. Moreover, measurements assigned to each skill are given equal weights. In this section, we develop our benchmark econometric model, which relaxes some of these assumptions. The benchmark model features a formal measurement system, which treats observed maladjustment syndromes and test scores as measures with error of underlying skills. The model produces estimates of the joint distribution of latent skills and the mapping of such skills to observed measurements, which depends in part on the precision of each measure. The measurement system also allows each measure to provide information about more than one factor. For example, a maladjustment syndrome can be a measure of both socio-emotional skill and of cognition. Moreover, externalizing behavior can affect maladjustment syndromes along with cognitive test scores. Using this framework, we are able to secure identification of the impact of underlying skills imposing relatively few assumptions. Results using the benchmark model are discussed in Section 4. As the benchmark model still imposes some somewhat arbitrary assumptions, which we detail below, we assess the sensitivity of the results to alternative assumptions in Section 5.

3.1 Parameterizations of the Schooling Decision Rule and Potential Outcomes

We approximate the schooling decision with a linear-in-parameters multinomial logit model with 6 schooling levels: $s \in \{0, 1, ..., 5\}$. Taking schooling level 0 as the base state, let the log-odds of schooling level s be

$$I_s = \log \frac{Pr(S = s)}{Pr(S = 0)} = \underbrace{Z}_{\substack{\text{observed by}\\ \text{econometrician}}} \cdot \beta_s + \underbrace{\eta_s}_{\substack{\text{unobserved by}\\ \text{econometrician}}}, \quad s = \{1, ..., 5\}, \tag{3}$$

where Z is a vector of variables observed by the econometrician that affect the schooling decision (see Table 4), β_s is a vector of parameters mapping variables in Z to schooling outcomes and η_s is a set of school-level-specific shocks that are unobserved by the econometrician. We impose separability between the observed and unobserved variables in the representation of the schooling decision rule.

We focus on two labor market outcomes in the benchmark model: the hourly wage and the weekly hours worked for individuals who are employed at age 33. More specifically, the log hourly wage, y, and the log weekly working hours, h, are represented by the following two equations:

$$y = \underbrace{X}_{\substack{\text{observed by}\\ \text{econometrician}}} \cdot\beta_Y + \sum_{s=1}^{5} \gamma_{s,Y} \cdot \underbrace{\mathbf{1}[s]}_{\substack{\text{observed by}\\ \text{econometrician}}} + \underbrace{U_Y}_{\substack{\text{unobserved by}\\ \text{econometrician}}} \tag{4}$$

$$h = \underbrace{X}_{\substack{\text{observed by}\\ \text{econometrician}}} \cdot\beta_H + \sum_{s=1}^{5} \gamma_{s,H} \cdot \underbrace{\mathbf{1}[s]}_{\substack{\text{observed by}\\ \text{econometrician}}} + \underbrace{U_H}_{\substack{\text{unobserved by}\\ \text{econometrician}}} . \tag{5}$$

X is the set of basic controls shown in Table 4 and the β's are vectors of associated coefficients. $\mathbf{1}[s]$ is an indicator function indicating the observed schooling level with associated coefficients γ. U_Y and U_H are unobserved determinants of wages and hours worked.

We assume there exists a vector $\mathbf{f}$ of skills that are unobserved by the econometrician and which generate all dependence across the η_s, U_Y, and U_H. More specifically, suppose

$$\eta_s = \mathbf{f}'\alpha_{\mathbf{S}} + \nu_s, \tag{6}$$

$$U_Y = \mathbf{f}'\alpha_{\mathbf{Y}} + \omega_Y, \tag{7}$$

$$U_H = \mathbf{f}'\alpha_{\mathbf{H}} + \omega_H, \tag{8}$$

where the αs are equation-specific vectors of coefficients attached to latent skills $\mathbf{f}$, ν_s is a normal idiosyncratic error term for the schooling choice, and ω_Y and ω_H are normal idiosyncratic error terms for the two labor outcomes, the log hourly wage and the log weekly hours worked.

3.2 Measurement System for Unobserved Skills f

The vector of skills $\mathbf{f}$ is not directly observed, but it can be proxied by a set of observable measurements. We allow for the ten BSAG maladjustment syndromes and four aptitude test scores measured at the age of 11 (Table 3) to be proxies for three latent skills. Specifically, let M be a vector of $K = 14$ measurements of the three latent skills $\mathbf{f} = (f_1, f_2, f_3)$, where f_1 is externalizing behavior, f_2 is internalizing behavior and f_3 is cognition. We propose a linear measurement system:

$$M = \begin{pmatrix} M_1 \\ \vdots \\ M_K \end{pmatrix} = \begin{pmatrix} m_1 + \sum_{j=1}^{3} \lambda_{1j} f_j + W\delta_1 + \varepsilon_1 \\ \vdots \\ m_K + \sum_{j=1}^{3} \lambda_{Kj} f_j + W\delta_K + \varepsilon_K \end{pmatrix}, \tag{9}$$

where m_k is the mean of the measurement k, and λ_{kj} is the factor loading of latent skill j on the kth measurement.[20]

The latent skills follow a joint normal distribution, with mean μ and variance-covariance matrix $\mathbf{\Sigma}$:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \sim N(\mu, \mathbf{\Sigma}) = N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}\right) \tag{10}$$

Referring back to Table 4, notice we include a vector of additional observables denoted W and associated coefficients δ. W includes class size, the percentage of students in the respondent's school taking GCE exams, a dummy for the local educational authority (similar to a school district fixed effect), and the number of full-time teachers in the school. These additional variables are included to address the concern that school attributes simultaneously affect schooling and labor outcomes along with teacher mis-reporting. If we omit these variables, we may misattribute variation in outcomes to variation in skills that is actually due to differences in schooling attributes.[21]

3.3 Identifying Assumptions

The key identifying assumption is that conditional on $\mathbf{f}$, Z, and X, choices and outcomes are statistically independent. Formally, we array the ν_s, $s \in \{1, ..., 5\}$ into a vector $\nu = (\nu_1, \nu_2, \nu_3, \nu_4, \nu_5)$ and array ω_Y and ω_H into a vector $\omega = (\omega_Y, \omega_H)$. We assume that,

$$\nu_s \perp\!\!\!\perp \nu_{s'}, \forall s \neq s', \tag{11}$$

$$\omega_Y \perp\!\!\!\perp \omega_H, \tag{12}$$

$$\omega \perp\!\!\!\perp \nu. \tag{13}$$

Assumptions (11), (12) and (13) maintain independence of the shocks over schooling categories, and across schooling and labor market outcomes. This assumption is testable and in Section 5.2 we provide evidence that $\mathbf{f}$ adequately captures the unobserved covariation of the three outcomes.

In addition, we array the measurement errors, ε_k, $k \in \{1, ..., K\}$ into a vector $\varepsilon =$

[20] The BSAG maladjustment scores range from 0 to 15 but most individuals have a score near 0 (see Table 2). To account for this feature of the data, we use the logarithm of each BSAG score plus one as the relevant measurement in the measurement system.

[21] As with other control variables, results are not affected if these variables are omitted.

$(\varepsilon_1, ..., \varepsilon_K)$ and assume that,

$$\begin{aligned} \varepsilon_k &\perp\!\!\!\perp \varepsilon_{k'}, \forall k \neq k', && (14) \\ (\omega, \nu) &\perp\!\!\!\perp \varepsilon. && (15) \end{aligned}$$

Assumptions (14) and (15) maintain that the measurement errors are independent from each other, and independent from the shocks.[22]

Lastly, we assume that,

$$\begin{aligned} (\nu, \omega, \varepsilon) &\perp\!\!\!\perp (\mathbf{f}, Z, X, W), && (16) \\ \mathbf{f} &\perp\!\!\!\perp (X, Z, W). && (17) \end{aligned}$$

Assumption (16) assumes independence of all the shocks and measurement errors with respect to factors and observables, and Assumption (17) assumes independence of factors with respect to observables.[23] The latter assumption might seem restrictive. In Section 5.6, we discuss alternative sets of models where we change the set of variables in Z and X, including a model where we allow Z and X to be empty vectors.

Identification of the measurement system requires further restrictions. One restriction that secures identification is to choose three "dedicated measures," that is, for each skill we choose one measure that is only affected by that skill (Williams, 2018). We choose "hostility towards children" (M_1) as the dedicated measurement for externalizing behavior (f_1), "depression" (M_2) as the dedicated measurement for internalizing behavior (f_2), and "verbal ability" (M_3) for cognition (f_3). We allow all three skills to load on the remaining 11 measurements. The choice of dedicated measures is somewhat arbitrary, yet is motivated by how we interpret each of the factors. Literature in psychology and medicine posits that externalizing behavior is closely associated with disruptive disorders, which motivates our choice of "hostility towards children" as the dedicated measurement (Duncan and Magnuson, 2011; Kendler and Myers, 2014). Internalizing behavior is commonly associated with depressive disorders, which motivates our choice of "depression" as the dedicated measurement (Regier, Kuhl, and Kupfer, 2013; Kendler and Myers, 2014). In Section 5, we discuss changes to results when we rely on alternative restrictions, including different choices of dedicated measurements. Finally, as factors do not have a natural scale, we normalize the

[22]In a robustness check, we allow for correlation among some of the error terms in our measurement system. We allow the errors for anxiety towards children and anxiety towards adults to be correlated, and for hostility towards children and hostility towards adults to also be correlated. In both cases and for both genders, the estimated correlation is zero. These results are available upon request.

[23]Williams (2018) discusses these assumptions in more detail. In particular,Williams (2018) describes conditions under which Assumption (17) can be relaxed.

coefficients of the dedicated measurements to unity as is commonly done in this literature. These identifying restrictions amount to

$$\begin{aligned} M_1 &= m_1 + 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3 + W\delta_1 + \varepsilon_1 \\ M_2 &= m_2 + 0 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3 + W\delta_2 + \varepsilon_2 \\ M_3 &= m_3 + 0 \cdot f_1 + 0 \cdot f_2 + 1 \cdot f_3 + W\delta_3 + \varepsilon_3 \end{aligned} \tag{18}$$

3.4 Likelihood and Estimation Procedure

We summarize the parameters to be estimated by a vector denoted Φ:

$$\Phi = (\beta, \gamma, \alpha, \Xi) \tag{19}$$

where β denotes the set of coefficients on the vectors of observables absent the schooling level in equations (3)-(5), γ is the set of coefficients governing the returns to schooling, α is the set of coefficients governing the returns to unobserved skills and Ξ are coefficients of the measurement system described in equations (9) and (10).

We estimate the model by simulated maximum likelihood in two stages. In the first stage, we estimate the measurement system for unobserved skills. In the second stage, given the parameter estimates $\widehat{\Xi}$ found in the first step, we estimate the remaining structural parameters, (β, γ, α).[24] We implement the estimation for boys and girls separately; that is, we allow all parameters to differ by gender.

In the first stage, for each suggestion for parameters in the measurement system indexed by g_1 and denoted $\Xi^{(g_1)}$, and for each individual i, we simulate a vector of unobserved factors T times and, for each draw of the factors, compute the probability of observing each measurement.[25] More specifically, given a parameter suggestion, we draw a block matrix of size $T \times I \times J$ from a standard normal distribution, where J is the number of latent factors, i.e. 3, and I is the number of individuals. Then, for each individual i and draw t, we construct a vector of latent factors $(f_{i1t}^{(g_1)}, f_{i2t}^{(g_1)}, f_{i3t}^{(g_1)})$ and compute $\mathbf{f}_{it}^{\mathbf{M},(g_1)}(M_i)$, the probability of observing the classroom misbehavior measurements and test scores, for individual i, draw t and parameter suggestion (g_1).

In the first stage, the simulated log likelihood function is computed as the sum of the log

[24] Along with other robustness checks, in Section 5, we discuss an alternative specification where we estimate the measurement system jointly with outcomes.

[25] For estimation, we set $T = 500$. Results are robust if we use larger or smaller numbers and are available upon request.

of each individual's average likelihood contribution taken over the T draws:

$$\mathcal{L}_1^{(g_1)} = \sum_{i=1}^{I} \log \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{f}_{it}^{\mathbf{M},(g_1)} \left(M_i \right) \right) \tag{20}$$

Using both simplex and gradient methods, we evaluate $\mathcal{L}_1^{(g_1)}$ at different values in the parameter space, indexing these suggestions by (g_1), and continue until a maximum is found.

In the second stage, taking $\widehat{\Xi}$ as given, we follow a similar procedure to compute the density functions corresponding to each outcome: the probability of individual i reaching a schooling level s, $\left(\mathbf{f}_{it}^{\mathbf{S},(g_2)}(s)\right)$, the probability of observing wage y_i, $\left(\mathbf{f}_{it}^{\mathbf{Y},(g_2)}(y_i)\right)$, and hours worked h_i, $\left(\mathbf{f}_{it}^{\mathbf{H},(g_2)}(h_i)\right)$, for individual i, draw t and parameter suggestion (g_2). The simulated log likelihood in the second stage is given by:

$$\mathcal{L}_2^{(g_2)} = \sum_{i=1}^{I} \log \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{f}_{it}^{\mathbf{M},(\widehat{\Xi})} \left(M_i \right) \times \prod_{s=0}^{5} \mathbf{f}_{it}^{\mathbf{S},(g_2)}(s)^{\mathbf{1}[s=s_i]} \right. \\ \left. \times \, \mathbf{f}_{it}^{\mathbf{H},(g_2)} \left(h_i \right)^{\mathbf{1}(e_i=1)} \times \mathbf{f}_{it}^{\mathbf{Y},(g_2)} \left(y_i \right)^{\mathbf{1}(e_i=1)} \right) \tag{21}$$

where s_i represents the observed schooling choice and e_i the observed employment status (with employed taking the value 1) in the data.[26]

4 Empirical Results

Here we present the key empirical findings from our benchmark econometric model described in the previous section. We first discuss estimates of the measurement system mapping unobserved factors to observed BSAG maladjustment syndromes (Section 4.1). Next, we discuss the externalizing schooling penalty (Section 4.2) followed by the externalizing earnings premium (Section 4.3). Thereafter, we assess the role of intermediate choices and outcomes, such as occupation, in explaining our findings (Section 4.4). Finally, we assess whether results are driven by selection into employment (Section 4.5).

[26]Standard errors are computed by constructing the Hessian of the joint likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the joint likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood.

4.1 Mapping Unobserved Skills to Observed Misbehaviors

Starting with the joint distribution of unobserved skills, we find a positive correlation between externalizing and internalizing behavior along with a negative correlation between the two socio-emotional skills and cognition. These patterns hold for both males and females (Table 7). The negative relationship between the two socio-emotional skills and cognition could reflect the distribution of skill endowments at birth. It could also reflect early childhood investments if the same environments that promote externalizing and internalizing behaviors also slow cognitive development (Heckman and Cunha, 2007). An example would be childhood poverty. The positive relationship between externalizing and internalizing behavior is well-documented in the child development literature. Children under stress as a result of poverty or a family disruption tend to develop both aggressive and depressive symptoms (Wolfson, Fields, and Rose, 1987). Accounting for correlation across factors means that we avoid mis-attributing returns to skills. For example, failing to account for the positive association between externalizing and internalizing behavior could lead us to over-estimate the degree to which each socio-emotional skill negatively affects schooling.

In Tables 8 and 9, we report estimates of factor loadings mapping latent skills to BSAG maladjustment syndromes and aptitude test scores. Estimates are reported separately by gender. Consistent with the interpretation of the two socio-emotional skills discussed before, externalizing behavior loads heavily onto disruptive and impulsive syndromes such as hostility towards adults, anxiety towards children or adults, inconsequential behavior and restless behaviors, while internalizing behavior loads heavily onto inhibited syndromes such as withdrawal, unforthcomingness and writing off adults and standards. Cognition loads mostly onto the tests scores. These results are also broadly in line with how we grouped the measurements as reflecting the three skills in the preliminary analysis in Section 2. Across genders, there are some differences in the factor loadings, but they are generally small and insignificant.

Most of the coefficients on the variables related to school characteristics have the expected signs. A higher percentage of students in the school taking GSCE qualification exams is negatively associated with misbehaviors and positively associated with test scores. Being in a public school (LEA) tends to reduce girls' test performances, but not boys'. The number of teachers is an indicator of the size of the school, with bigger schools associated with lower test scores. A larger class size tends to reduce measurements closely related to externalizing behavior and increase test scores, for both boys and girls, which is in contrast with previous research (Fredriksson, Öckert, and Oosterbeek, 2012).[27]

[27]It is possible that class size captures omitted school-level variables which positively affect student out-

4.2 The Externalizing Penalty in School

The marginal effect estimates of the multinomial logit model for educational attainment are reported in Table 10.[28] There is a significant negative relationship between externalizing behavior and educational attainment for boys. A difference from the estimates in the preliminary analysis is that the negative relationship between externalizing and schooling for females is no longer present. The marginal effects are small and the sign of the relationship is unclear. In other words, high-externalizing females are better able to finish school in comparison to high-externalizing males. This finding may reflect how teachers are more likely to punish or refer for special help a male versus a female child for the same level of aggression (Gregory, 1977). On the other hand, we find that internalizing behavior is negatively associated with educational attainment for females, but less strongly so for males. This is also in line with research that finds stronger effects of conduct disorders and weaker effects of anxiety and depressive symptoms for the educational attainment of males in comparison to females (Kessler et al., 1995).

Effect sizes for socio-emotional skills in the schooling model are much smaller than those for cognition, which predicts schooling at similar magnitudes across genders. Also, the effect of family characteristics is consistent with our initial expectations. Having parents with more education and who work in more lucrative occupational categories is related to higher educational attainment for the child. Moreover, individuals living in poverty during their childhood, suggesting relatively few family resources available to invest in children, are less likely to attain higher levels of education.

In general, estimates for the schooling model are broadly consistent with literature that studies the impact of emotional problems in school. One of the key pathways relating behavioral problems to low educational attainment is through early educational failures such as repeating a grade or falling behind in class. If externalizing or internalizing behavior make learning more difficult, this would in part be captured by the strong negative relationship between the two socio-emotional skills and cognition (which is identified from test scores) reported in Table 7. However, the negative impact of the socio-emotional skills on education is not fully explained by these correlations, suggesting additional mechanisms. For example, McLeod and Kaiser (2004) argue that children with internalizing and externalizing behaviors withdraw from social relationships in school, including those with teachers, in order to minimize their exposure to negative interactions. This could make schooling more costly.

comes, such as teacher quality if better teachers are assigned to larger classes. This type of bias would be more concerning if these variables were the focus of our analyses rather than controls to address potential mis-reporting.

[28]Standard errors for the marginal effects are calculated using the delta method.

4.3 The Externalizing Premium on the Labor Market

Literature studying the consequences of externalizing behavior has generally limited attention to educational attainment. In contrast, we assess the relationship between childhood misbehavior and labor market outcomes. Estimates of hours and wage equations conditional on employment are reported in Tables 11 and 12.[29] The benchmark model results are presented in Column [2] of these two tables.

For males, a one-standard-deviation increase in externalizing behavior predicts a statistically significant 6.4% increase in hourly wages, but does not significantly affect weekly hours worked. For females, a one-standard-deviation increase in externalizing behavior predicts a marginally significant 4.7% increase in hours worked per week, but does not significantly affect hourly wages.[30] The evidence points to different ways that externalizing behavior increases earnings for males and females. It tends to raise wages for externalizing males, while it tends to increase labor supply on the intensive margin for externalizing females. The positive effects on hourly wages or weekly hours worked demonstrate that externalizing behavior is productive on the labor market even though it is counter-productive in school, especially for boys. This is a novel finding in the literature on the economic consequences of childhood misbehavior.

One possible explanation for the externalizing premium in the labor market is that externalizing behavior captures unobserved but correlated personality traits. Several studies have examined the relationship between externalizing and internalizing behaviors and better-known measures, such as the "Big 5" personality traits. Evidence suggests that externalizing behavior is negatively associated with conscientiousness, agreeableness, and openness to new experience, while internalizing behavior is mostly related to neuroticism (Ehrler, Evans, and McGhee, 1999; Almlund et al., 2011). Moreover, agreeableness predicts lower earnings (Judge, Livingston, and Hurst, 2012).[31] It is possible that high-externalizing individuals earn more for some of the same reasons that agreeable people earn less, such as a distaste for competition and negotiating. This point relates to earlier work that explores how economic preferences relate to standard measures of socio-emotional skill (Becker et al., 2012). In robustness exercises, we test for this possibility. In Appendix A.6, we show that controlling for

[29] Selection into employment is discussed in the following section.

[30] Using our crude model, we considered an alternative specification where we control for hours worked in the wage equation. For males, the relationship between externalizing and wages increases slightly after we control for hours worked. For females, the relationship becomes negative and is insignificant. These results are available upon request.

[31] To explain why, Barry and Friedman (1998) show that individuals with higher levels of agreeableness are worse negotiators as they are susceptible to being anchored by early offers in the negotiation process. Relatedly, Spurk and Abele (2011) show that less agreeable individuals are more competitive in the workplace and place a higher emphasis on career advancement.

the "Big 5" traits reduces the effect of externalizing behavior on earnings by about 20% and increases the negative effect on education by about 15%. However, our main findings remain after we control for the "Big 5" personality traits, suggesting that, despite correlations, the skills we study are distinct factors with independent impacts on economic outcomes.[32]

Internalizing behavior is negatively related to both productivity in the labor market and hours worked. For males, a one-standard-deviation increase in internalizing behavior predicts a very significant 9.6% decrease in hourly wage and a marginally significant 1.8% decrease in weekly hours worked. For females, the counterpart coefficients in both the wage and hours worked equations are negative, but neither is significant. We also find that cognition significantly increases hourly wages (by 2.5% for males and 4.4% for females), but does not influence the hours decision for either gender. The remaining parameters follow conventional wisdom. For example, higher educational attainment increases worker productivity, but has little effect on the number of hours worked for those already employed. Also, individuals living in or around London earn significantly higher hourly wages, while individuals who experience financial difficulties in childhood receive lower hourly wages.

Note that in the benchmark model (Column [2]), the labor outcome equations condition on the schooling choices. To evaluate whether including endogenous schooling choices affects the estimated effects of the unobserved skills in an intuitive way, in Tables 11 and 12, we also report estimates when we exclude the schooling outcomes from the outcome equations (Column [1]). Excluding schooling variables allows us to estimate the net impact of skills on labor market outcomes. Doing so increases the point estimates of the effect of cognition on hourly wages for both males and females. It also reduces the point estimate of the effect of externalizing on hourly wages for males, though only slightly. Since externalizing reduces schooling for males and schooling improves wages, it is not surprising that excluding schooling would generate a smaller net effect of externalizing on wages. What is notable is that the coefficient is still positive after including schooling, suggesting that more externalizing males earn higher wages despite the negative impact of externalizing on schooling.

Our findings demonstrate a more nuanced relationship between childhood misbehavior

[32] One important caveat to our results on personality using the NCDS is that the "Big 5" personality traits are measured at age 50, after educational and labor market outcomes are realized. Thus, estimates could be biased due to simultaneity, if labor market shocks influence how individuals respond to the personality questions. We therefore address the question of adjusting for additional unobserved skills using the British Cohort Study (BCS). Results are reported in Appendix D. Using the BCS, we construct socio-emotional skills from a larger set of behavioral questions. The larger number of measurements allows us to identify as many as 8 distinct factors, three of them capturing externalizing behavior, internalizing behavior and cognition. We find that the key patterns described in our benchmark model still hold when we identify externalizing behavior using this larger set of measurements, and also when we include additional factors capturing additional socio-emotional skills in schooling and labor outcome equations.

and labor market outcomes than has been recognized in previous literature. They also illustrate how socio-emotional skills can have mixed effects on economic outcomes.

4.4 Externalizing and Other Outcomes

To further examine why externalizing behavior increases earnings, we study its relationship to intermediate outcomes, such as occupational sorting and labor market experience along with household structure, including fertility and marriage. The aim is to assess mechanisms underlying the externalizing premium. One possibility is that high-externalizing individuals are less likely to be in relationships or to have children, which could free up time to work longer hours or to focus on working more productively. Another possibility is that high-externalizing individuals prefer (or experience a lower utility cost from) sorting into lucrative occupations that other individuals avoid.

To explore these mechanisms, we assess how estimated coefficients change when we add endogenous intermediate outcome variables to the wage and hours equations. Results are reported in Table 13 for hourly wage and Table 14 for weekly hours worked. We start from the benchmark model (Column [1]) and add a dummy variable for being married by age 33 (Column [2]) and reported number of children by age 33 (Column [3]).[33]

While having a partner has a strong positive effect on wages for both males and females, having children lowers wages and weekly hours worked for females only. Controlling for partnership and fertility does not change the coefficients on externalizing in any significant way for males, but it roughly doubles the point estimates of the impact of externalizing on wages and hours worked as well as increases their statistical significance for females. To understand the gender difference in how fertility affects the externalizing earnings premium, we estimate a linear regression of the number of children by age 33 on the three skills from the previously estimated measurement system. Estimates are found in Table 15. Externalizing males and females are both more likely to have a larger number of children by age 33, but based on the outcome equations (Tables 13 and 14), having more children is somewhat irrelevant to earnings for males, but is associated with a large drop in both wages and hours for females. Recall from Figure 2 that female earnings are much lower for women with children in comparison to women without children. For males, there is no discernible relationship. Findings relating externalizing behavior, number of children and earnings suggest that the relatively low net impact of externalizing on women is attributable to two countervailing effects, which are (i) higher fertility, which lowers earnings, and (ii) better labor market outcomes.

[33] We keep the measurement system mapping latent skills to observed measurements of misbehavior as in the benchmark model.

Finally, we add months of experience and occupational choices as controls (Columns [4] and [5]). Doing so does not appreciably alter the estimated relationship between externalizing behavior and labor market outcomes. In fact, there is little evidence that externalizing behavior drives individuals into any specific occupation once we have controlled for education. An exception is that high-externalizing females are more likely to self-select into skilled manual occupations but the parameter is only marginally significant. These results are found in Table 16 where we estimate a multinomial logit model of occupational sorting with unskilled occupations as the base group.[34]

To visualize results, we plot earnings against different levels of externalizing separately for men and women in Figure 3. The slope of the curve represents the impact of externalizing behavior on earnings. To generate the figure, we simulate weekly earnings, which is the product of hourly wages and weekly hours worked, as we vary the externalizing behavior from the 5th percentile to the 95th percentile, keeping other latent skills and covariates at the population median. We repeat this exercise conditioning on different sets of intermediate outcome variables. Finally, we produce this figure separately for males and females to illustrate gender differences. For males, conditioning on intermediate outcomes does not change the slope very much (Panel (a)). For females, the slope increases noticeably when we condition on the number of children by age 33 (Panel (b)), reflecting the positive relationship between externalizing and fertility along with the negative relationship for females between number of children and earnings. An interpretation of this result is that there are large labor market returns to high-externalizing women who do not have children.

In summary, though externalizing behavior is related to a host of economic outcomes that also predict earnings, positive labor market returns to externalizing behavior are not driven by differential sorting into these outcomes. This provides further evidence for the idea that externalizing behavior generates higher earnings. High-externalizing males are either more productive or are otherwise able to secure higher payoffs for their labor, and high-externalizing females work more hours.

4.5 Externalizing and Employment

Recall that wage and hours regressions are estimated on individuals who are employed. One possible concern is that the estimated relationship between externalizing behavior and earnings is driven solely by selection into employment. This is slightly different from our analysis of intermediate outcomes, which examines whether externalizing raises wages through sort-

[34]In additional analyses that are available upon request, we also show that the returns to externalizing do not differ significantly across occupations.

ing into other outcomes, such as parenthood or occupation. Here, the concern is that the externalizing premium is an artifact of selection into employment. For example, if high externalizing individuals dislike employment, it is possible that our estimates are driven by high-externalizing individuals who supply labor because they are highly productive due to unobserved factors. This would introduce positive selection bias into our estimates of the impact of externalizing behavior on earnings.

To address this possibility, we first estimate a multinomial logit model of selection into self- and paid employment with the same set of controls as in the outcome equations of the benchmark model, while fixing the measurement system.[35] The results can be found in Table 17 where unemployed individuals are the base group. We find important gender differences in our results. Females with higher levels of externalizing behavior are less likely to be unemployed and are more likely to be self-employed or employed at age 33.[36] For males, externalizing behavior is weakly negatively related to employment. Moreover, men and women with high levels of internalizing behavior are significantly more likely to be unemployed. Cognition predicts higher employment for males and lower employment for females, though the effects are not significant for either gender. The main impact of cognition on employment likely works through schooling, for which we control and which predicts employment for both genders.

The results for externalizing behavior among females are especially concerning since they raise the possibility that high-externalizing women who are relatively productive (or who work more hours when employed) tend to self-select into employment. This could be the case if high-externalizing women face a lower disutility of working and are therefore observed in unemployment only if they are particularly unproductive due to other (omitted) factors. To address this concern, we exploit earnings data for individuals who were not employed at age 33, but reported earnings in a previous employment spell. The idea is that labor market outcomes at other periods would provide some insight into how much unemployed individuals would have earned if they had worked at age 33 (Neal and Johnson, 1996). Using this approach, the proportion of individuals in our sample for whom we obtain a measure of earnings rises from 62% to 92% (90% for males and 93.5% for females).[37] If results are driven by highly productive, high-externalizing individuals entering employment, we would expect the estimated relationship between externalizing behavior and earnings to fall once

[35] In other words, we keep the measurement system mapping latent skills to observed maladjustment syndromes and test scores as in the benchmark model so that changes in the parameters are solely attributable to changes in the control variables and not in the measurement system.

[36] This finding is similar to the one in Levine and Rubinstein (2017). They show that teenagers who engage in risky or illicit activities are more likely to self-select into entrepreneurship.

[37] This percentage is somewhat lower for males because a higher percentage of males are always classified as self-employed.

we include earnings information on unemployed individuals.

We re-estimate the benchmark model outlined in Section 3 using the larger sample that includes individuals with earnings information from other years. Estimates are presented in Table 18. In Columns [1] and [3], we present the estimated parameters using weekly earnings as an outcome in the benchmark model. In Columns [2] and [4], we use the new measure of earnings from an enlarged sample that includes individuals not working at age 33. For males, including earnings for the unemployed tends to reduce the point estimate of the externalizing earnings premium slightly, but does not affect the significance level. For females, including the imputed earnings of the unemployed does not make a difference in the estimated impact of externalizing on earnings. These results provide some evidence against the possibility that selection into employment explains the estimated results for the males in our sample. The bottom line is that the results from our benchmark model continue to hold after we account for the possibility of self-selection into employment.[38]

5 Sensitivity Analyses

This section discusses a host of sensitivity analyses, beginning with changes to assumptions on the measurement system. Most results discussed in this section are reported in Appendix C.

5.1 Alternative Dedicated Measurements

Throughout the paper, we have assumed there are three unobservable skills, externalizing behavior, internalizing behavior and cognition, which are identified from measures of childhood classroom misbehavior and test scores. In the estimation of the measurement system that links the unobserved skills to the maladjustment syndromes and test scores, we designated one particular measure as a sole measurement of each skill. In this section, we discuss the implications of these assumptions. As explained in Section 3.2, we have chosen “hostility towards children” as the dedicated measurement for externalizing behavior, “depression” for internalizing behavior, and “verbal ability” for cognition. To assess sensitivity, we re-estimate the model iterating over all possible candidates for the dedicated measurements of the two

[38]As an additional robustness check, we also experimented with a formal Heckman selection model for hourly wages using partnership and number of children as exclusion restrictions. We do not present these results since they suggest a similar story to the one presented in Table 18 and because the exclusion restrictions are difficult to defend.

socio-emotional skills.[39] We plot the effect on weekly earnings from a one-standard-deviation increase in externalizing behavior for each different choice of dedicated measurements in Figure 4.[40] The dashed bars indicate the results from the benchmark econometric model for males and females.

There are several points to note about the figure. While different dedicated measurement choices imply different magnitudes of the effects on education and earnings, externalizing behavior almost always has a significantly positive earnings premium for females and in a majority of cases has a significantly positive premium for males. Moreover, the benchmark specification is not the one that produces the largest externalizing earnings premium for males or females. Remarkably, in no specification do we find significant evidence against our main result from the benchmark model. The specifications under which the earnings premium becomes insignificant tend to be those in which withdrawal or unforthcomingness is chosen as the dedicated measure for internalizing behavior. As shown in Appendix C.1, in such cases, depression loads heavily on the "externalizing" factor. In this case, we identify a factor that is a mixture of what we typically regard as outwardly expressed externalizing behavior and inwardly expressed internalizing behavior, and the impact of the "externalizing" factor on earnings is muted, which is expected given the negative correlation between depression and productivity. That said, the positive returns to externalizing behavior in the labor market do not require that depression be the dedicated measurement for internalizing behavior. For example, results are similar when the BSAG measure "writing off of adults and adult standards" is chosen as the dedicated measurement for internalizing behavior (see Tables S19 and S20 in the Appendix).

Findings from this exercise illustrate the fundamental identification problem in measuring underlying traits, discussed in Almlund et al. (2011). Creating a summary variable of measurements (as we did in our preliminary analysis) is simple, but implicitly imposes a number of unattractive assumptions. The measurement system in our benchmark model permits the relaxation of some assumptions, though a minimal set of assumptions, including which variable to use as a dedicated measurement, is still required for identification — and the analyst must choose which to use. A benefit of the measurement system is that such assumptions are explicit, and highlight the trade-off between letting the data guide the analysis versus imposing just enough structure to identify economically meaningful objects. In our case, it is possible to construct an externalizing factor that maps to depression and which has a substantially smaller effect on earnings due to the negative correlation between depres-

[39] We re-estimate the model under all possible combinations of dedicated measurements as described in Table 3.

[40] Additional findings using alternative dedicated factor assumptions are reported in Appendix C.1.

sion and earnings. However, doing so appears to contradict the standard interpretation of externalizing behavior. Alternatively, we can construct an externalizing factor that does not map to depression, loads heavily onto outwardly expressed aggressive behaviors, and which has a positive impact on earnings. The benchmark model imposes the latter assumption.

5.2 Testing the Three-Factor Assumption

The grouping of the factors into cognition, externalizing behavior and internalizing behavior has been previously validated in the literature, as described in Section 2. However, it is still possible that our results are influenced by additional factors that determine both choices and outcomes and that have been omitted in our analysis. We test for this possibility in two ways.

First, if an important fourth factor has been omitted, then the model with only three factors should make poor predictions on sample covariances between outcomes and choices. In Tables 19 and 20, we present the simulated covariances between schooling levels and outcomes (i.e. wages and hours worked) against their sample counterparts, for males and females respectively. As is clear from the tables, our model with three factors has a good sample fit, suggesting that the benchmark model adequately accounts for the observed relationship between choices and outcomes.

Second, we implement an extension of the benchmark model that allows for a fourth latent factor. Results are reported in Appendix C.2. In both male and female samples, we find that this fourth factor is insignificant in the schooling equation but does have some predictive power in the wages and hours equations. However, including the fourth factor does not affect the estimated relationship between externalizing behavior, schooling and earnings. In particular, even after controlling for the fourth factor, externalizing behavior still significantly reduces educational attainment and increases wages for males, while for females externalizing behavior continues to increase hours worked despite having no impact on schooling.

5.3 Imposing Independence of Factors

If one is willing to assume independence across factors, we can relax other assumptions on the measurement system. In particular, we can dispense with two of the three dedicated measures and let most measurements load on all three factors. In this scenario, identification of the three factors still requires one measurement dedicated to a single factor and a second dedicated to two of the three factors. We refer to Williams (2018) for a detailed discussion on the identification of linear latent factor models. We present results from an alternative model with independent factors in Appendix C.3. For this analysis, we chose verbal ability as a dedicated measure for cognition and depression as a semi-dedicated measure for internalizing behavior and cognition. Results under this specification resemble those from our benchmark model that permits correlation across factors.

5.4 Joint Estimation of the Measurement System, Choices and Outcomes

In our benchmark model, we estimated the model in two steps. We estimated the measurement system in a first step, and the educational choice and labor market outcome equations in a second step. In Appendix C.4 we present results where we estimate the measurement system jointly with choices and outcomes. Results remain largely unchanged. Heckman, Humphries, and Veramendi (2018) discuss the relative merits of a two-stage estimation and a joint estimation (see Appendix (A.12) of their paper). The main motivation for us to pursue the two-stage estimation strategy in the benchmark model is that it makes interpreting the factors easier. In the two-stage estimation, the factors are solely identified from the measurement system, and can be clearly interpreted as underlying skills that account for the classroom misbehavior and tests measured at age 11.

5.5 Alternative Models for the Educational Choice

In the benchmark model, we adopt a very flexible model of schooling, a multinomial logit model with six schooling levels. One concern is that the model might be too flexible to pick up the negative effect of externalizing behavior on female schooling choices. In Appendix C.5, when we use more restrictive models such as a linear regression of years of schooling or a multinomial logit with four coarser schooling levels, we show that the results from these alternative specifications are largely consistent with the results from the benchmark model. That is, the externalizing behavior has a significant negative impact on schooling for boys, but the effect is insignificant for girls.

5.6 Alternative Set of Controls

One possible concern is that our results are sensitive to the controls used in the schooling and labor market outcome equations. To test that, in Appendix C.6, we report an alternative set of estimates where we modify the set of controls for labor market outcomes while keeping the controls for education constant. We show that the positive relationship between externalizing behavior and earnings emerges as soon as we control for internalizing behavior and cognition. Also, we report an alternative set of estimates where we assume the sets of controls Z and X are empty. We show that the negative relationship between externalizing behavior and educational attainment and the positive relationship between externalizing behavior and earnings remain in a model without any observable controls. While we continue to control for these variables in our benchmark model to eliminate potential biases, our main results are qualitatively robust to excluding them.

6 Subgroup Analysis: Childhood Poverty, Misbehavior and Earnings

Studying a sample of disadvantaged black children in the U.S., Heckman, Pinto, and Savelyev (2013) find that an early childhood education program increased earnings in part by reducing externalizing behavior. In contrast, we show that externalizing behavior can be valuable in the labor market. In this section, we explore whether differences in findings are explained by differences in the socioeconomic status of the group being analyzed. One possibility is that children born into poorer families face a higher likelihood of criminality or police involvement for the same level of externalizing behavior.

We estimate two variations of our benchmark econometric model. First, we include a measure of police involvement at age 16 as an additional outcome equation and as an additional explanatory variable in the schooling, wage and hours equations. Second, we estimate the model on a sub-sample of our analytic sample, which is selected to resemble the family characteristics of the sample studied in Heckman, Pinto, and Savelyev (2013). In particular, we construct a subsample of our analytic sample consisting of subjects who faced financial difficulties during childhood. We refer to this sample as the "Low SES" subsample. Recall, this occurs if the interviewer reported that the household appeared to be experiencing poverty in 1965 or if a member of the household self-reported having financial difficulties in the 12 months prior to being observed in either 1969 or 1974.[41] We estimate the benchmark econometric model separately for the "Low SES" subsample and for all other subjects in our analytic sample, which we call the "High SES" subsample.[42]

[41] An alternative would be to use family income. However, perhaps surprisingly, the NCDS does not collect information on family income or parental pay in the first three surveys. In the fourth survey, when children were 16 years old, categorical information was collected on each parent's work pay. However, this information on parental pay is missing for over 20% of our sample. Therefore, we decided to use the available information about financial difficulties instead.

[42] Summary statistics for the sub-samples are found in Tables S4 and S5 in Appendix A. The "Low SES"

Estimating separate models by childhood SES, we find that many patterns are similar to the main model.[43] However, we also find some important differences by childhood SES. First, we estimate a larger penalty for externalizing behavior for educational attainment among individuals that grew up in low-SES households (Table 21). This finding is broadly consistent with results in Ramey (2014), who shows that externalizing blacks in the U.S. face a higher likelihood of punishment by suspension in comparison to similarly externalizing whites. This could be because schools that serve low-SES children in the UK (or black children in the U.S.) have fewer resources to address externalizing behavior and therefore react to it through suspensions or expulsions.[44]

Perhaps most importantly, we find that the labor market returns to externalizing behavior fail to extend to the "Low SES" subsample. First, individuals that grew up with financial difficulties experience less than a third of the externalizing wage premium than individuals that did not (Table 22). Second, in the hours equation (Table 23), the coefficient is negative and insignificant for the low-SES group (versus 0.047 and significant for the high-SES group). Wage returns to the other skills are similar across the two groups, as are the returns to education. On the other hand, there are some differences in the influence of internalizing behavior and cognition for the hours worked decision. Internalizing behavior decreases hours worked for the high-SES group but not for the low-SES group and cognition increases hours worked for the low-SES group only. However, one important caveat to the results presented in this section is that we cannot statistically differentiate the returns to externalizing behavior for the two socioeconomic groups because the standard errors in the estimates for the Low SES sample are large.[45]

Following the results in Heckman, Pinto, and Savelyev (2013), one possible explanation for differences in results by childhood SES status is that low-SES individuals are at a higher risk of criminal behavior for a given level of externalizing behavior. In line with this possibility, we find that low-SES individuals are more likely to have some police involvement (the

sample completes less education, earns lower wages, and are less likely to be employed, though when employed hours are similar across groups. They also score higher on all BSAG maladjustment syndromes, on average. We estimate the measurement system separately by group since it is possible that underlying skills map to observed behaviors differently by group. Similarly, to study black-white differentials in labor market outcomes in the U.S., Urzua (2008) allows the distribution and impact of underlying skills to vary by race.

[43] In results available upon request, we separated our sample into four groups by gender and financial difficulty status. The main patterns remain largely similar. However, the standard errors for the group in financial difficulties, when divided by gender, were too large for any useful inference to be made.

[44] There are also some differences in the returns to family characteristics, such as the father's occupation.

[45] One possibility is that differences across groups are driven by non-linearities in the relationship between externalizing behavior and labor market outcomes. We test for (and rule-out) non-linearities within the reduced-form (crude) econometric framework in Appendix A.4. In results available upon request, we also test for non-linearities separately for each of the two SES groups. The different returns to externalizing behavior for the two groups is present even when we allow for non-linearities in the relationship.

estimated constant in the police involvement equation is 0.090 for the high-SES group and 0.179 for the low-SES group). However, the relationship between externalizing behavior and police involvement is stronger for the high-SES group ((see Table 24).[46] Interestingly, we find little evidence that police involvement is related to worse labor market outcomes for either SES subgroup (see the coefficients for police involvement in Tables 22 and 23). In other words, while externalizing behavior predicts higher police involvement, police involvement does not appear to derail labor market prospects among individuals in the British sample we study, including those who grew up in families facing financial difficulties. These results raise the possibility that the returns to externalizing behavior might be negative in a context where police involvement is highly penalized in the labor market. This is the sort of context studied in Heckman, Pinto, and Savelyev (2013), who examine a sample composed mostly of at-risk black youths in the U.S. However, for our sample, police involvement cannot explain why low-SES individuals in the British sample we study receive little payoff to externalizing behavior.

Therefore, despite our initial results showing that externalizing behavior is associated with better labor market outcomes, this positive association does not extend to individuals who faced poverty during childhood. In other words, the payoffs to socio-emotional skills are context-dependent, as argued in Lundberg (2013). To explain differences in returns to skills across socioeconomic groups, we are therefore left with at least two distinct, but related possibilities. The first is that there are true differences in the productivity of externalizing behavior across groups. For example, children born into wealthier families may be better able to channel aggressive tendencies into productive activities.[47] A second possibility is that high-externalizing individuals from lower classes face different selection rules than their higher-SES counterparts, but these rules are not observed by the econometrician. For example, managers or co-workers may view high-externalizing individuals from high-SES families as ambitious leaders and be willing to hire them in high-wage positions or to promote them. In contrast, high-externalizing individuals from lower SES families may find their advancement thwarted if they are viewed as disruptive, aggressive or impolite. If so, high-externalizing individuals from low-SES families are not unproductive *per se*, but instead sort into jobs where they earn less. In both cases, childhood SES and externalizing behavior exhibit complementarities. Seen another way, these results suggest the concerning possibility that children from poorer families are unable to unleash the potential of skills that are valuable and lucrative for children born into wealthier families.

[46] Interestingly, internalizing behavior and cognition are associated with less police involvement, though the coefficients are much larger in magnitude for high-SES individuals.

[47] See, for example, Doyle et al. (2009) on the timing of investments to decrease inequality.

7 Robustness Across Datasets

In this section, we will show that key results from the benchmark model extend to other datasets following individuals from different cohorts from Great Britain and the U.S. The datasets that we explore include the 1970 British Cohort Study, the National Education Longitudinal Study of 1988, the Panel Study of Income Dynamics, and the National Longitudinal Survey of Youth 1979 Children and Young Adults. The latter three are U.S. datasets.

7.1 Additional Datasets and Key Variables

This section briefly introduces each dataset along with scales used to construct the key variables in each dataset. For further details on these datasets, we refer the interested reader to Appendix D. We include in our analysis four major longitudinal studies that follow individuals over the lifecycle with measurements of behavior during childhood in school and early-career labor market outcomes. Compared to the 1958 cohort in the NCDS, these datasets cover younger cohorts born in the 70s and 80s. One noticeable difference is that in the 1958 British cohort, men exhibit higher education on average than women, but the gender education gap has reversed among younger cohorts — in both Great Britain and the U.S. Average years of schooling in our NCDS sample is 12.73 for men and 11.25 for women. For a younger 1970 British cohort, these two numbers are 12.44 for men and 12.46 for women (Table S43). For the 1973-6 U.S. cohort in NELS, these two numbers are 14.03 and 14.17 (Table S48). Despite the reversal of the gender education gap, the gender wage gap persists in all of these datasets. Appendix D contains additional summary statistics for each of these samples.

A concern when comparing estimates across datasets is that each dataset uses a different scale to measure child behaviors and cognition. Thus, it is difficult to distinguish between differences in estimates arising from context-dependence versus differences due to how skills are measured. We cannot fully address this concern, but attempt to ensure that we rely on measures that have been validated in earlier research. Another concern is that there are important differences in parental and teachers' reports of children's behavior (Achenbach, McConaughy, and Howell, 1987; De Los Reyes and Kazdin, 2005; Ronda, 2017). To make sure the new measures are reasonably comparable to those in the NCDS, we rely on measures constructed from teachers' reports when possible. This is possible in 3 of the 4 replication datasets, since the CNLSY did not interview children's teachers. Finally, to facilitate comparison across datasets, we measure the education outcome by years of schooling.

The 1970 British Cohort Study (BCS). The BCS follows the lives of more than 17,000 individuals born in the Great Britain in a single week of 1970. The survey design of the BCS is similar to the NCDS we use in the main analysis. It surveys these cohort members at ages 0, 5, 10, 16, 26, 30, 34, 38 and 42.

In the BCS, cognition is measured by aggregating over a number of test scores from tests taken at age 10. We construct measures of externalizing and internalizing behaviors by factor analyzing a number of descriptions of classroom behaviors at the age of 10 reported by teachers; for example, "Displays outbursts of temper" relates to externalizing behavior, while "Worried and anxious" relates to internalizing behavior. The factor analysis identifies eight factors that explain the variation in children's behaviors in the data. We name these eight factors externalizing behavior, internalizing behavior, cognition, grit, clumsiness, cognition, hand coordination, and two miscellaneous factors. The interpretation of the factors is motivated by magnitudes of estimated factor loadings. To construct each factor, we sum up the scores of the descriptions that belong to each factor. Next, we run regressions that link years of schooling and earnings to these factors. In the main paper, we report estimates with only 3 factors (cognition, externalizing behavior and internalizing behavior). Results are robust to the inclusion of additional factors (see Appendix D). Finally, we construct weekly earnings from yearly earnings reported at age 30.

The National Education Longitudinal Study of 1988 (NELS). The NELS is a nationally representative, longitudinal study of 8th graders in the U.S. in 1988, who belong to the 1973-6 birth cohorts. The students are interviewed in four follow-up surveys in 1990, 1992, 1994, and 2000.

Following the analysis in Farkas (2011), which uses the NELS to construct the 3 skills, externalizing behavior is measured as the weighted average (by factor analysis) of two teachers' responses to the question, "Is this student frequently disruptive?" in the 8th grade (i.e. when the child is around age 12) and one teacher's response in a 5 point scale to the question, "How often is this student disruptive?" in the 10th grade (i.e. when the child is around age 14). Internalizing behavior is measured by the weighted average of two 8th grade teachers' and one 10th grade teacher's responses to the question "Is this student exceptionally passive or withdrawn?" Cognition is measured by reading and math test scores in the 8th grade. Weekly earnings are measured in the 2000 survey, when these individuals are between 24 and 27 years old.

The Panel Study of Income Dynamics (PSID). The Child Development Supplement (CDS) of the PSID is a longitudinal study of children's developmental outcomes within the context of family, neighborhood, and school environments. We draw our sample from the first wave of the CDS, conducted in 1997, which collects extensive child-specific developmental data from families with children aged 0-12. These subjects are PSID individuals too and participate in PSID surveys. We restrict our attention to children aged 7-13 interviewed in the 1997 CDS, which corresponds to members of the 1984-1990 cohorts.

The National Longitudinal Survey of Youth (NLSY) 1979: Children and Young Adults (CNLSY). The CNLSY is a longitudinal study that follows the children of the women in the NLSY 1979, who were a nationally representative sample of women aged 14 to 21 on December 31, 1978. The first interview of the Adults Survey of the NLSY was conducted in 1986, when these women were 21 to 28 years old. The first interview of the Young Adults Survey of the NLSY was conducted in 1994 for children aged 15 and above born to the NLSY women.

In the PSID and CNLSY, we rely on measures for externalizing and internalizing behavior from the Behavior Problems Index (BPI). These measures were developed by Peterson and Zill (1986) and have been used extensively in the literature.[48] In the PSID and CNLSY, cognition is measured by reading and applied problems test scores. The key difference between the two datasets is that in the PSID, the scores for externalizing and internalizing behaviors were constructed from both teachers' reports and mothers' reports, while in the CNLSY the measures are only available from maternal reports. In the main paper, we report results using the teachers' reports in the PSID and mothers' reports in the CNLSY. We report additional analysis using maternal reports in the PSID in Appendix D. Labor market earnings was measured at ages 25 and 26 in the PSID and 29 or 30 in the CNLSY.

7.2 Externalizing, Schooling, and Earnings Across Datasets

In each dataset, we link the measure of externalizing behavior to schooling and earnings in a manner similar to the crude analysis described in Section 2.3. In particular, we run an OLS regression of years of schooling on measures of externalizing behavior, internalizing behavior and cognition, controlling for individual characteristics such as gender and race as well as family background information such as father's and mother's education status and employment status. Results are summarized in Table 25.

[48]The Behavior Problems Index was originally developed from the Achenback Behavior Problems Checklist to measure the incidence and severity of child behavior problems. The BPI scale is based on a set of 32 problems describing whether a behavior is often, sometimes, or never true of the targeted child. These items are divided into two subscales: 1) a measure of externalizing or aggressive behavior and 2) a measure of internalizing, withdrawn or sad behavior, with the group confirmed by factor analysis by the survey team.

Pooling boys and girls, in all datasets, externalizing behavior is associated with fewer years of schooling. This negative effect is strongly significant, with the exception of the PSID where the negative coefficient is significant at the 10% level. Separately by gender, the negative effect of externalizing behavior on education outcomes is particularly relevant for boys. Compared to the NCDS sample, the point estimates of the correlation between externalizing behavior and years of schooling in the samples of younger cohorts tend to be bigger, suggesting an externalizing penalty in school that persists across cohorts.

Next, for each dataset, we run OLS regressions of log weekly earnings for workers in their young adulthood on the same measures of externalizing and internalizing behaviors along with cognition, controlling for gender, race and educational attainment.[49] Results are summarized in Table 26. Externalizing behavior is significantly associated with higher earnings in the two British datasets, the 1958 and the 1970 cohort, and two U.S. datasets, NELS and PSID. In the full sample, the point estimate of the impact of externalizing on earnings from the NCDS lies between estimates obtained from other datasets, suggesting that the externalizing earnings premium does not vary systematically across countries or over time. Interestingly, when we stratify the sample by gender, the positive impact from externalizing on earnings seems to be stronger for young female workers in the UK, while it is stronger for young male workers in the U.S.

The CNLSY is the only dataset where we do not find a significant relationship between externalizing behavior and earnings. There are two reasons that might explain the inconclusive result. To be in the CNLSY analytic sample, the child's labor market performance at the age of 25-26 must be observed. Recall that these individuals are the children of women from the original NLSY 79 sample, who were born between 1957 and 1965. This implies that we only observe earnings of individuals born from relatively young mothers. The average age of the mothers at the birth of the individuals in our analytical sample is 21.8. Using the PSID and the NELS, we can verify that having a young mother is often associated with less financially stable home environments. It is thus possible that our findings using the CNLSY arise from sample selection towards children born into poorer households, which aligns with the lack of evidence of an externalizing premium among low-SES families from the NCDS (see Section 6). Another important difference between the CNLSY and the other datasets we examine is that the CNLSY only has parents' report of children's behavior and previous research has highlighted important differences in parental and teachers' reports of children's behavior (Achenbach, McConaughy, and Howell, 1987; De Los Reyes and Kazdin, 2005), and evidence of bias in maternal reports (Boyle and Pickles, 1997; Najman et al., 2000) .

[49]Results are qualitatively robust to omitting educational attainment. We report these additional specifications in Appendix D.

In sum, we conclude that our main results using a 1958 British cohort — that childhood externalizing behavior negatively affects schooling, especially among females, while positively affecting labor market outcomes — are fairly consistent across datasets. Mixed returns of this particular socio-emotional skill appear to be an empirical regularity that is evident in at least two major economies and in datasets from several different time periods.

8 Discussion and Conclusion

Few would argue against the idea that stronger cognition or better health would improve outcomes on almost any conceivable economic dimension. Some socio-emotional skills also appear to have positive returns in many sectors. In this paper, we illustrate that it is generally not meaningful to think of socio-emotional skill as either good or bad *per se*. Using several methods and datasets, we have demonstrated that the same socio-emotional skill, externalizing behavior, can be productive in one context and counterproductive in another. Our study also illustrates the idea that skills, characteristics or attributes that are valuable during childhood may not be valuable during adulthood. Investments in human capital should be evaluated in light of this possibility. In particular, mixed effects of externalizing behavior suggest caution in devising policies that target children with apparently undesirable behaviors or so-called negative socio-emotional skills. Such policies may be beneficial in the short-run by improving educational outcomes, but may also be costly in the long-run by stifling a potentially productive labor market skill. Our results are particularly salient given recent efforts to include measures of socio-emotional skills as part of schools' and teachers' performance ratings. Many school systems are poised to enact policies that evaluate schools on character skills development (West et al., 2016).

Yet, the externalizing premium we identify does not provide justification for policies encouraging externalizing behavior. A primary reason is that there is little guidance in the literature on how to promote externalizing behavior in a way that would leave other skills unchanged. For example, it is possible that externalizing and internalizing behaviors are inextricably linked — indeed, our findings on the correlations between the two skills suggest this is a possibility. If so, attempts to modify one skill could modify the other in a way that harms children.

Even if it were possible to modify externalizing behavior without modifying other skills, there are at least three reasons why it is unclear if doing so would increase welfare. First, we do not know if externalizing behavior raises productivity, and thus increases aggregate output, or if it instead allows individuals to secure a larger share of a constant amount of resources (e.g., to bargain for higher wages at the expense of co-workers' earnings). In the latter case, it is not clear what would be gained by encouraging externalizing behavior. The second reason is tied to our results on differences in returns by socioeconomic group. Leaving aside important questions about the source of such differences in returns (e.g., a higher propensity for police involvement due to biases in criminal justice systems), policies that promote externalizing behavior could harm groups already at a disadvantage. A third reason is the possibility of negative spillover effects in the classroom if externalizing children are disruptive and limit other students' learning (Henneberger, Coffman, and Gest, 2016). Recall, our results show that externalizing behavior loads heavily onto the maladjustment syndrome "hostility towards children." Given documented negative impacts of bullying on education, policies increasing hostility among schoolchildren are likely to be unproductive (Brown and Taylor, 2008; Carrell and Hoekstra, 2010; Carrell, Hoekstra, and Kuka, 2018).

In future research, it would be useful to explore policies that increase schooling without stifling a valuable labor market skill. Policies could accommodate externalizing behavior rather than penalizing or simply attempting to eliminate it. Notice, the idea is to shift the price of a skill in certain contexts rather than the skill itself. In making this distinction, we draw on pedagogical research that discusses "control-oriented" teaching methods, which are designed to reduce externalizing behavior versus "relationship-oriented" methods, which are designed to strengthen the learning environment for externalizing children.[50] A simple example illustrates the difference in the two approaches. Young students who often initiate conversations with teachers at inopportune times could be punished for interrupting a lesson. Instead, they could be given a "raincheck" and invited to initiate a discussion at an appropriate time. The effectiveness of such practices is demonstrated by a randomized controlled trial of the My Teaching Partner-Secondary program (MTP-S), in which a web-mediated program on improving teacher-student in-class interaction has produced reliable gains in student achievement (Allen et al. (2011)).

Another direction for future research would aim to better understand the externalizing premium, addressing whether high-externalizing individuals produce more or, alternatively, are simply better able to secure larger payoffs for the same amount of output compared to low-externalizing individuals. Future work could also explore heterogeneity in the effects of socio-emotional skills across groups. For example, Ramey (2014) studies a cohort of individuals born in the U.S. in the 1980s and 1990s. He shows that high-externalizing black students are more likely to be suspended from school when compared to similarly externalizing white students. This could lead to differences in the returns to externalizing behavior across

[50]For an overview of pedagogical techniques that foster a caring and positive student-teacher relationship, in particular, in dealing with student misbehavior, see Hamre and Pianta (2006).

racial groups since suspensions are associated with lower schooling attainment and lower earnings. Extending the findings in Ramey (2014) to consider labor market outcomes could help to clarify whether differences in returns to the same socio-emotional skill help to explain stubbornly persistent inequality across socioeconomic and racial groups.

References

Achenbach, Thomas M., Stephanie H. McConaughy, and Catherine T. Howell. 1987. "Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Information Correlations for Situational Specificity." *Psychological Bulletin* 101 (2):213–232.

Agan, Amanda Y. 2011. "Non-Cognitive Skills and Crime." University of Chicago.

Aizer, Anna. 2009. "Neighborhood Violence and Urban Youth." In *The Problems of Disadvantaged Youth: An Economic Perspective*, edited by Jonanthan Gruber. University of Chicago Press, 275–308.

Allen, Joseph P., Robert C. Pianta, Anne Gregory, Amori Yee Mikami, and Janetta Lun. 2011. "An Interation-Based Approach to Enhancing Secondary School Instruction and Student Achievement." *Science* 333 (6045):1034–1037.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Personality Psychology and Economics." *Handbook of the Economics of Education* 4 (1).

Barry, Bruce and Raymond A Friedman. 1998. "Bargainer Characteristics in Distributive and Integrative Negotiation." *Journal of Personality and Social Psychology* 74 (2):345.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse. 2012. "The Relationship Between Economic Preferences and Psychological Personality Measures." *Annual Review of Economics* 4 (1):453–478.

Becker, Gary Stanley. 1967. *Human Capital and the Personal Distribution of Income: An Analytical Approach*. Institute of Public Administration.

Bertrand, Marianne and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1):32–64.

Blanden, Jo, Paul Gregg, and Lindsey Macmillan. 2007. "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education." *Economic Journal* 117 (519):C43–C60.

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.

Boyle, Michael H and Andrew R Pickles. 1997. "Influence of Maternal Depressive Symptoms on Ratings of Childhood Behavior." *Journal of Abnormal Child Psychology* 25 (5):399–412.

Brown, Sarah and Karl Taylor. 2008. "Bullying, Education and Earnings: Evidence from the National Child Development Study." *Economics of Education Review* 27 (4):387–401.

Card, David. 2012. "Earnings, Schooling, and Ability Revisited." In *35th Anniversary Retrospective (Research in Labor Economics, Volume 35)*, edited by Solomon W. Polachek and Konstantinos Tatsiramos. Emerald Group Publishing Limited, 111–136.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review* 108 (11):3377–3415.

Carrell, Scott E and Mark L Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics* 2 (1):211–228.

Chase-Lansdale, P Lindsay, Andrew J Cherlin, and Kathleen E Kiernan. 1995. "The Long-term Effects of Parental Divorce on the Mental Health of Young Adults: A Developmental Perspective." *Child Development* 66 (6):1614–1634.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.

De Los Reyes, Andres and Alan E Kazdin. 2005. "Informant Discrepancies in the Assessment of Childhood Psychopathology: a Critical Review, Theoretical Framework, and Recommendations for Further Study." *Psychological Bulletin* 131 (4):483.

Doyle, Orla, Colm P Harmon, James J Heckman, and Richard E Tremblay. 2009. "Investing in Early Human Development: Timing and Economic Efficiency." *Economics & Human Biology* 7 (1):1–6.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly. 2007. "Grit: Perseverance and Passion for Long-Term Goals." *Journal of Personality and Social Psychology* 92 (6):1087.

Duncan, Greg J and Rachel Dunifon. 2012. "'Soft-Skills' and Long-Run Labor Market Success." In *5th Anniversary Retrospective (Research in Labor Economics, Volume 35)*, edited by Solomon W. Polachek and Konstantinos Tatsiramos. Emerald Group Publishing Limited, 313–339.

Duncan, Greg J and Katherine Magnuson. 2011. "The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems." In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation, 47–69.

Ehrler, David J, J Gary Evans, and Ron L McGhee. 1999. "Extending Big-Five Theory into Childhood: A Preliminary Investigation into the Relationship Between Big-Five Personality Traits and Behavior Problems in Children." *Psychology in the Schools* 36 (6):451–458.

Farkas, George. 2011. "Middle and High School Skills, Behaviors, Attitudes, and Curriculum Enrollment, and Their Consequences." In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation, 71–90.

Farmer, Elizabeth MZ. 1993. "Externalizing Behavior in the Life Course The Transition From School to Work." *Journal of Emotional and Behavioral Disorders* 1 (3):179–188.

———. 1995. "Extremity of Externalizing Behavior and Young Adult Outcomes." *Journal of Child Psychology and Psychiatry* 36 (4):617–632.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2012. "Long-term Effects of Class Size." *Quarterly Journal of Economics* 128 (1):249–285.

Fronstin, Paul, David H Greenberg, and Philip K Robins. 2005. "The Labor Market Consequences of Childhood Maladjustment." *Social Science Quarterly* 86 (s1):1170–1195.

Ghodsian, M. 1977. "Children's Behaviour and the BSAG: Some Theoretical and Statistical Considerations." *British Journal of Social and Clinical Psychology* 16 (1):23–28.

Gregory, Mary K. 1977. "Sex Bias in School Referrals." *Journal of School Psychology* 15 (1):5–8.

Hamre, Bridget K. and Robert C. Pianta. 2006. "Student-Teacher Relationships." In *Children's Needs III: Development, Prevention and Intervention*, edited by G. Bear and Kathleen M. Minke, chap. 5. National Association of School Psychologists, 49–60.

Heckman, James J and Flavio Cunha. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2):31–47.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2018. "Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking." *Journal of Political Economy* 126 (S1):197–246.

Heckman, James J and Tim Kautz. 2014. "Fostering and Measuring Skills: Interventions that Improve Character and Cognition." In *The Myth of Achievement Tests*, edited by James J. Heckman, John Eric Humphries, and Tim Kautz. University of Chicago Press, 341–430.

Heckman, James J, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6):2052–86.

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

Henneberger, Angela K, Donna L Coffman, and Scott D Gest. 2016. "The Effect of Having Aggressive Friends on Aggressive Behavior in Childhood: Using Propensity Scores to Strengthen Causal Inference." *Social Development* 26 (2):295–309.

Jackson, Michelle. 2006. "Personality Traits and Occupational Attainment." *European Sociological Review* 22 (2):187–199.

Judge, Timothy A, Beth A Livingston, and Charlice Hurst. 2012. "Do Nice Guys and Gals Really Finish Last? The Joint Effects of Sex and Agreeableness on Income." *Journal of Personality and Social Psychology* 102 (2):390.

Kendler, Kenneth S. and John Myers. 2014. "The Boundaries of the Internalizing and Externalizing Genetic Spectra in Men and Women." *Pschological Medicine* 44 (3):647–655.

Kessler, Ronald C, Cindy L Foster, William B Saunders, and Paul E Stang. 1995. "Social Consequences of Psychiatric Disorders, I: Educational Attainment." *American Journal of Psychiatry* 152 (7):1026–1032.

Levine, Ross and Yona Rubinstein. 2017. "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?" *Quarterly Journal of Economics* 132 (2):963–1018.

Lundberg, Shelly. 2013. "The College Type: Personality and Educational Inequality." *Journal of Labor Economics* 31 (3):421–441.

McLeod, Jane D and Karen Kaiser. 2004. "Childhood Emotional and Behavioral Problems and Educational Attainment." *American Sociological Review* 69 (5):636–658.

Najman, Jake M, Gail M Williams, Jane Nikles, SUE Spence, William Bor, Michael O'callaghan, Robyne Le Brocque, and Margaret J Andersen. 2000. "Mothers' Mental Illness and Child Behavior Problems: Cause-Effect Association or Observation Bias?" *Journal of the American Academy of Child & Adolescent Psychiatry* 39 (5):592–602.

Neal, Derek A and William R Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5):869–895.

Peterson, James L and Nicholas Zill. 1986. "Marital Disruption, Parent-child Relationships, and Behavior Problems in Children." *Journal of Marriage and the Family* :295–307.

Ramey, David M. 2014. *The Social Control of Childhood Behavior via Criminalization or Medicalization: Why Race Matters*. Ph.D. thesis, The Ohio State University.

Regier, Darrel A., Emily A. Kuhl, and David J. Kupfer. 2013. "The DSM-5: Classification and Criteria Changes." *World Psychiatry* 12 (2):92–98.

Ronda, Victor. 2017. "When Mothers and Teachers Disagree: Observer Reports and Children's Noncognitive Skills." Mimeo, Johns Hopkins University.

Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2):135–146.

Segal, Carmit. 2013. "Misbehavior, Education, and Labor Market Outcomes." *Journal of the European Economic Association* 11 (4):743–779.

Spurk, Daniel and Andrea E Abele. 2011. "Who Earns More and Why? A Multiple Mediation Model From Personality to Salary." *Journal of Business and Psychology* 26 (1):87–103.

Stott, Denis Herbert, Emily G Sykes, and NC Marston. 1974. *The Social Adjustment of Children: Manual to the Bristol Social Adjustment Guides*. Hodder & Stoughton.

Urzua, Sergio. 2008. "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices." *Journal of Human Resources* 43 (4):919–971.

West, Martin R, Matthew A Kraft, Amy S Finn, Rebecca E Martin, Angela L Duckworth, Christopher FO Gabrieli, and John DE Gabrieli. 2016. “Promise and Paradox Measuring Students Non-Cognitive Skills and the Impact of Schooling.” *Educational Evaluation and Policy Analysis* 38 (1):148–170.

Williams, Benjamin. 2018. “Identification of the Linear Factor Model.” *forthcoming at Econometrics Reviews*.

Willis, Robert J. and Sherwin Rosen. 1979. “Education and Self-Selection.” *Journal of Political Economy* 87 (5):S7–S36.

Wolfson, Johanna, Joel H Fields, and Susan A Rose. 1987. “Symptoms, Temperament, Resiliency, and Control in Anxiety-Disordered Preschool Children.” *Journal of the American Academy of Child & Adolescent Psychiatry* 26 (1):16–22.

Yeung, W Jean, Miriam R Linver, and Jeanne Brooks-Gunn. 2002. “How Money Matters for Young Children’s Development: Parental Investment and Family Processes.” *Child Development* 73 (6):1861–1879.

9 Tables and Figures

Table 1: Summary Statistics of Demographics, Education, and Labor Market Outcomes

	Both	Males	Females	Diff
Female	0.507			
	(0.500)			
No Formal Education	0.112	0.103	0.121	*
	(0.316)	(0.304)	(0.327)	
CSE	0.128	0.113	0.142	***
	(0.334)	(0.316)	(0.349)	
O Level	0.345	0.305	0.384	***
	(0.475)	(0.460)	(0.486)	
A Level	0.147	0.191	0.104	***
	(0.354)	(0.393)	(0.305)	
Higher Education	0.146	0.150	0.143	
	(0.354)	(0.357)	(0.350)	
Higher Degree	0.122	0.138	0.106	***
	(0.327)	(0.345)	(0.308)	
Hourly Wage	6.636	7.638	5.457	***
	(3.053)	(2.967)	(2.712)	
Weekly Hours Worked	36.36	43.54	27.91	***
	(12.67)	(7.772)	(12.09)	
Weekly Earnings	252.5	329.0	162.3	***
	(152.5)	(134.5)	(119.6)	
In Paid Work	0.804	0.919	0.692	***
	(0.397)	(0.273)	(0.462)	
Employee	0.675	0.740	0.612	***
	(0.468)	(0.439)	(0.487)	
Financial Difficulty	0.160	0.155	0.165	
	(0.367)	(0.362)	(0.371)	
London Before 16	0.355	0.352	0.359	
	(0.479)	(0.478)	(0.480)	
London at 33	0.298	0.292	0.304	
	(0.457)	(0.455)	(0.460)	
Observations	7241	3573	3668	7241

Notes: This table lists the summary statistics of demographics, education, and labor market outcomes for the analytic sample of 7,241 individuals. For education categories and employment, entries are in the form of percentages divided by 100. Wages and weekly earnings are measured in 1992 British pounds. Employee means the percentage of individuals in the sample that are in paid work and not self-employed. Statistics are reported separately for all individuals (Column [1]), for males (Column [2]) and for females (Column [3]). In Column [4], *, ** and *** mean that differences between males and females are significant at the 10, 5 and 1 percent levels, respectively.

Table 2: Summary Statistics of BSAG Syndromes, Test Scores, and Crude Measures of Unobserved Skills

	Both	Males	Females	Diff
Hostility Towards Adults	0.763	0.889	0.641	***
	(1.753)	(1.858)	(1.635)	
Hostility Towards Children	0.239	0.265	0.215	**
	(0.718)	(0.777)	(0.655)	
Anxiety for Acceptance by Adults	0.515	0.483	0.546	*
	(1.152)	(1.097)	(1.203)	
Anxiety for Acceptance by Children	0.298	0.401	0.197	***
	(0.761)	(0.898)	(0.580)	
Restlessness	0.194	0.242	0.147	***
	(0.520)	(0.575)	(0.455)	
Inconsequential Behavior	1.262	1.674	0.861	***
	(1.869)	(2.152)	(1.433)	
Depression	0.932	1.085	0.784	***
	(1.454)	(1.536)	(1.353)	
Withdrawal	0.308	0.374	0.243	***
	(0.772)	(0.878)	(0.646)	
Unforthcomingness	1.477	1.537	1.419	*
	(2.034)	(2.009)	(2.057)	
Writing Off of Adults and Adult Standards	0.908	1.124	0.697	***
	(1.588)	(1.786)	(1.334)	
Verbal Ability	23.21	22.17	24.22	***
	(8.952)	(9.171)	(8.615)	
Reading Ability	16.59	16.61	16.57	
	(5.977)	(6.232)	(5.717)	
Non-Verbal Ability	21.76	21.59	21.93	*
	(7.310)	(7.424)	(7.194)	
Math Ability	17.71	18.02	17.42	*
	(10.07)	(10.32)	(9.812)	
Externalizing	0.000	0.155	-0.151	***
	(1.000)	(1.107)	(0.858)	
Internalizing	0.000	0.113	-0.110	***
	(1.000)	(1.058)	(0.927)	
Misbehavior	0.000	0.154	-0.150	***
	(1.000)	(1.083)	(0.887)	
Cognition	0.000	-0.0309	0.0301	**
	(1.000)	(1.030)	(0.969)	
Observations	7241	3573	3668	7241

Notes: This table lists the summary statistics of the BSAG maladjustment syndromes and the test scores for the analytic sample of 7,241 individuals. The BSAG syndromes are constructed using teachers' reports of misbehavior in school. For each maladjustment syndrome, a child receives a score, which is an integer between 0 and 15, with 15 indicating a persistent display of the behavior described by the syndrome. In the table, entries are averages for each syndrome for the analytic sample. To construct crude measures of unobserved skills, we sum up all variables used to measure that skill in the preliminary analysis (see Table 3) and then normalize each unobserved skill to have mean zero and standard deviation one. Statistics are reported separately for all individuals (Column [1]), for males (Column [2]) and for females (Column [3]). In Column [4], *, ** and *** mean that differences between males and females are significant at the 10, 5 and 1 percent levels, respectively.

Table 3: MEASUREMENTS USED TO IDENTIFY UNOBSERVED SKILLS

Unobserved Skill	Measures
Externalizing Behavior	◇ Hostility Towards Adults ◇ Hostility Towards Children ◇ Anxiety for Acceptance by Adults ◇ Anxiety for Acceptance by Children ◇ Restlessness ◇ Inconsequential Behavior ◇ Writing Off of Adults and Adult Standards
Internalizing Behavior	◇ Depression ◇ Withdrawal ◇ Unforthcomingness ◇ Writing Off of Adults and Adult Standards
Cognition	◇ Reading Comprehension Test Score ◇ Mathematics Test Score ◇ Non Verbal Score on General Ability Test ◇ Verbal Score on General Ability Test

Notes: This table lists the three unobserved skills used in the empirical analysis (externalizing behavior, internalizing behavior and cognition) and the observed variables used to identify them. Measures for externalizing and internalizing behaviors are drawn from the BSAG maladjustment variables derived from teachers' reports of misbehavior. For cognition, a series of aptitude test scores are used as measures. See Section 2.2 for further details.

Table 4: Additional Control Variables Used in the Analysis

Variables	Measurement System	Schooling Choices	Labor Outcomes
Class Size	x		
Percentage of Students Taking GCE exams	x		
Local Education Authority Dummy	x		
Number of Full-Time Teachers	x		
Financial Difficulties		x	x
London Dummy		x	x
Mother Education		x	
Father Education		x	
No Father Info.		x	
Father in Skilled Oc.		x	
Father in Managerial Oc.		x	
Working Mother		x	

Notes: This table summarizes the additional control variables we use in the measurement equations, the schooling choice equations and the labor outcome equations.

Table 5: Preliminary Analysis: Educational Attainment

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Misbehavior	-0.316	-0.119							
	(0.013)	(0.014)							
Externalizing			-0.277	-0.105	-0.197	-0.077	-0.083	-0.105	-0.050
			(0.014)	(0.014)	(0.015)	(0.016)	(0.016)	(0.019)	(0.026)
Internalizing					-0.164	-0.060	-0.056	-0.036	-0.082
					(0.015)	(0.016)	(0.016)	(0.020)	(0.025)
Cognition		0.748		0.757		0.748	0.673	0.652	0.704
		(0.016)		(0.015)		(0.016)	(0.016)	(0.022)	(0.024)
Mother Education							0.274	0.226	0.321
							(0.035)	(0.049)	(0.050)
Father Education							0.256	0.187	0.323
							(0.031)	(0.044)	(0.044)
No Father Info.							0.218	0.176	0.265
							(0.078)	(0.111)	(0.109)
Father in Skilled Oc.							0.145	0.183	0.109
							(0.033)	(0.046)	(0.047)
Father in Managerial Oc.							0.377	0.434	0.321
							(0.042)	(0.060)	(0.059)
Working Mother							0.017	-0.001	0.035
							(0.026)	(0.037)	(0.037)
London Dummy	0.075	0.019	0.079	0.020	0.073	0.018	-0.071	-0.081	-0.062
	(0.025)	(0.026)	(0.025)	(0.026)	(0.025)	(0.026)	(0.026)	(0.038)	(0.037)
Financial Difficulties	-0.577	-0.357	-0.597	-0.361	-0.582	-0.358	-0.284	-0.285	-0.288
	(0.036)	(0.037)	(0.035)	(0.037)	(0.036)	(0.037)	(0.038)	(0.052)	(0.055)
Female	-0.293	-0.326	-0.279	-0.322	-0.293	-0.327	-0.329		
	(0.025)	(0.026)	(0.025)	(0.026)	(0.025)	(0.026)	(0.026)		
Cutoff 1	-1.536	-1.828	-1.515	-1.824	-1.536	-1.829	-1.578	-1.505	-1.320
	(0.028)	(0.031)	(0.028)	(0.031)	(0.028)	(0.031)	(0.043)	(0.056)	(0.057)
Cutoff 2	-0.971	-1.137	-0.958	-1.135	-0.972	-1.138	-0.879	-0.857	-0.572
	(0.025)	(0.026)	(0.024)	(0.026)	(0.025)	(0.026)	(0.040)	(0.052)	(0.053)
Cutoff 3	0.031	0.084	0.032	0.085	0.029	0.083	0.374	0.264	0.819
	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.039)	(0.052)	(0.054)
Cutoff 4	0.461	0.594	0.459	0.594	0.459	0.593	0.901	0.927	1.208
	(0.022)	(0.024)	(0.022)	(0.024)	(0.022)	(0.024)	(0.040)	(0.054)	(0.056)
Cutoff 5	1.041	1.302	1.033	1.302	1.038	1.301	1.640	1.649	1.964
	(0.025)	(0.026)	(0.025)	(0.026)	(0.025)	(0.026)	(0.042)	(0.057)	(0.059)
Obs.	7241	7241	7241	7241	7241	7241	7241	3573	3668

Notes: This table contains parameter estimates from an ordered probit model used to link unobserved skills to educational attainment. We estimate the ordered probability of choosing one of six schooling levels on a set of observable variables along with crude measures of unobserved skills. To construct the crude measures of the three unobserved skills, we sum up all variables used to measure that skill according to Table 3 and then normalize each unobserved skill to have mean zero and standard deviation one. Models [1]-[7] include all individuals and a gender dummy, Model [8] includes only males and Model [9] only females. Standard errors in parentheses.

Table 6: Preliminary Analysis: Log Weekly Earnings

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Misbehavior	-0.089	-0.024							
	(0.008)	(0.009)							
Externalizing			-0.059	-0.013	-0.000	0.025	0.032	0.020	0.041
			(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)	(0.020)
Internalizing				-0.091		-0.055	-0.047	-0.055	-0.031
				(0.010)		(0.009)	(0.009)	(0.008)	(0.020)
Cognition		0.189			0.197	0.189	0.079	0.067	0.103
		(0.009)			(0.009)	(0.009)	(0.010)	(0.009)	(0.021)
CSE							0.075	0.049	0.091
							(0.035)	(0.035)	(0.058)
O Level							0.197	0.134	0.247
							(0.031)	(0.031)	(0.053)
A Level							0.333	0.170	0.511
							(0.035)	(0.033)	(0.073)
Higher Education							0.505	0.280	0.727
							(0.035)	(0.034)	(0.060)
Higher Degree							0.634	0.370	0.953
							(0.039)	(0.036)	(0.071)
Female	-0.927	-0.914	-0.917	-0.926	-0.908	-0.913	-0.867		
	(0.018)	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)		
London Dummy	0.242	0.215	0.245	0.240	0.215	0.214	0.202	0.211	0.185
	(0.019)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)	(0.015)	(0.034)
Financial Difficulties	-0.162	-0.079	-0.176	-0.164	-0.082	-0.079	-0.038	-0.040	-0.033
	(0.024)	(0.023)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)	(0.023)	(0.039)
Constant	5.674	5.656	5.673	5.674	5.654	5.656	5.346	5.476	4.374
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.029)	(0.029)	(0.048)
Obs.	4888	4888	4888	4888	4888	4888	4888	2643	2245

Notes: This table contains parameter estimates from OLS regressions used to link socio-emotional and cognitive skills to earnings. We regress log earnings of workers on a set of observable variables along with crude measures of unobserved skills. To construct the crude measures of the three unobserved skills, we sum up all variables used to measure that skill according to Table 3 and then normalize each unobserved skill to have mean zero and standard deviation one. Models [1]-[7] include all individuals and a gender dummy, Model [8] includes only males and Model [9] only females. Standard errors in parentheses.

Table 7: Measurement System: Latent Factor Correlation Matrix

	Males		
	Externalizing	Internalizing	Cognition
Externalizing	1.000	0.776	-0.164
Internalizing	0.776	1.000	-0.471
Cognition	-0.164	-0.471	1.000
	Females		
Externalizing	1.000	0.807	-0.135
Internalizing	0.807	1.000	-0.415
Cognition	-0.135	-0.415	1.000

Notes: This table lists the estimated correlation matrix of the three latent skills from the measurement system, separately by gender.

Table 8: Measurement System: From Skills to Misbehaviors and Test Scores, Males

	Exter.	Inter.	Cog.	C.Size	P.GCE	LEA	N.Teac.	Con.
Hostile Towards Children	1.000	0.000	0.000	-0.002	-0.001	0.039	-0.000	0.242
	(.)	(.)	(.)	(0.001)	(0.000)	(0.059)	(0.001)	(0.063)
Hostile Towards Adults	1.639	0.206	0.026	-0.003	-0.003	0.009	-0.001	0.629
	(0.082)	(0.043)	(0.014)	(0.001)	(0.001)	(0.081)	(0.002)	(0.088)
Anxiety Towards Children	1.614	-0.335	-0.053	-0.001	-0.001	0.071	-0.001	0.252
	(0.074)	(0.036)	(0.010)	(0.001)	(0.000)	(0.060)	(0.002)	(0.065)
Anxiety Towards Adults	1.079	-0.328	-0.067	-0.002	-0.002	-0.001	0.002	0.338
	(0.068)	(0.045)	(0.012)	(0.001)	(0.001)	(0.058)	(0.002)	(0.066)
Inconsequential Behavior	2.100	0.074	-0.145	-0.002	-0.003	0.125	0.003	0.802
	(0.091)	(0.047)	(0.014)	(0.002)	(0.001)	(0.093)	(0.002)	(0.100)
Restless Behavior	0.778	-0.053	-0.058	-0.001	-0.001	0.054	-0.000	0.167
	(0.042)	(0.023)	(0.008)	(0.001)	(0.000)	(0.049)	(0.001)	(0.054)
Depression	0.000	1.000	0.000	-0.002	-0.003	0.146	-0.001	0.629
	(.)	(.)	(.)	(0.001)	(0.001)	(0.085)	(0.002)	(0.093)
Withdrawal	-0.983	1.132	0.137	0.001	-0.001	0.034	-0.000	0.234
	(0.078)	(0.050)	(0.010)	(0.001)	(0.000)	(0.069)	(0.002)	(0.075)
Unforthcomingness	-2.117	1.899	0.195	0.001	-0.002	0.134	0.003	0.588
	(0.143)	(0.088)	(0.018)	(0.002)	(0.001)	(0.105)	(0.002)	(0.114)
Write Off Adults and Standards	0.082	1.077	0.069	0.001	-0.002	0.068	-0.001	0.543
	(0.072)	(0.052)	(0.013)	(0.001)	(0.001)	(0.093)	(0.002)	(0.101)
Verbal Ability	0.000	0.000	1.000	0.020	0.017	0.051	-0.011	-1.204
	(.)	(.)	(.)	(0.002)	(0.001)	(0.107)	(0.003)	(0.120)
Reading Ability	-0.116	-0.070	0.832	0.020	0.016	-0.183	-0.006	-0.920
	(0.108)	(0.065)	(0.021)	(0.002)	(0.001)	(0.109)	(0.003)	(0.127)
Non-Verbal Ability	-0.149	0.068	0.898	0.018	0.014	-0.039	-0.011	-0.871
	(0.099)	(0.062)	(0.020)	(0.002)	(0.001)	(0.101)	(0.003)	(0.114)
Math Ability	-0.087	-0.107	0.889	0.018	0.019	0.014	-0.006	-1.074
	(0.087)	(0.054)	(0.019)	(0.002)	(0.001)	(0.109)	(0.003)	(0.124)

Notes: This table lists the parameter estimates of the measurement system (equation (9)) for the subsample of males. Standard errors in parentheses.

Table 9: MEASUREMENT SYSTEM: FROM SKILLS TO MISBEHAVIORS AND TEST SCORES, FEMALES

	Exter.	Inter.	Cog.	C.Size	P.GCE	LEA	N.Teac.	Con.
Hostile Towards Children	1.000	0.000	0.000	-0.001	-0.002	-0.038	0.001	0.282
	(.)	(.)	(.)	(0.001)	(0.000)	(0.037)	(0.001)	(0.043)
Hostile Towards Adults	1.672	0.166	0.009	-0.005	-0.004	-0.003	0.000	0.615
	(0.080)	(0.042)	(0.012)	(0.001)	(0.001)	(0.065)	(0.002)	(0.076)
Anxiety Towards Children	1.288	-0.283	-0.033	-0.001	-0.001	-0.043	-0.000	0.240
	(0.066)	(0.035)	(0.008)	(0.001)	(0.000)	(0.032)	(0.001)	(0.039)
Anxiety Towards Adults	1.466	-0.380	-0.085	-0.002	-0.001	0.036	0.000	0.361
	(0.086)	(0.054)	(0.013)	(0.001)	(0.001)	(0.062)	(0.002)	(0.072)
Inconsequential Behavior	1.509	0.150	-0.107	-0.003	-0.004	-0.102	0.003	0.771
	(0.079)	(0.044)	(0.013)	(0.001)	(0.001)	(0.060)	(0.002)	(0.070)
Restless Behavior	0.554	-0.005	-0.055	0.000	-0.001	-0.034	0.000	0.156
	(0.036)	(0.022)	(0.007)	(0.001)	(0.000)	(0.032)	(0.001)	(0.039)
Depression	0.000	1.000	0.000	-0.004	-0.004	0.074	0.001	0.641
	(.)	(.)	(.)	(0.001)	(0.001)	(0.062)	(0.002)	(0.072)
Withdrawal	-0.957	0.935	0.099	-0.000	-0.001	-0.002	0.001	0.221
	(0.072)	(0.045)	(0.009)	(0.001)	(0.000)	(0.037)	(0.001)	(0.046)
Unforthcomingness	-2.750	2.161	0.187	-0.001	-0.002	0.219	0.002	0.544
	(0.178)	(0.106)	(0.019)	(0.002)	(0.001)	(0.086)	(0.002)	(0.096)
Write Off Adults and Standards	-0.341	1.086	0.069	-0.002	-0.003	-0.026	0.001	0.589
	(0.077)	(0.051)	(0.012)	(0.001)	(0.001)	(0.057)	(0.002)	(0.068)
Verbal Ability	0.000	0.000	1.000	0.020	0.014	-0.248	-0.009	-0.642
	(.)	(.)	(.)	(0.002)	(0.001)	(0.092)	(0.002)	(0.106)
Reading Ability	-0.101	-0.040	0.820	0.019	0.014	-0.522	-0.005	-0.508
	(0.130)	(0.075)	(0.021)	(0.002)	(0.001)	(0.087)	(0.003)	(0.102)
Non-Verbal Ability	-0.103	0.008	0.928	0.014	0.014	-0.235	-0.007	-0.546
	(0.122)	(0.071)	(0.021)	(0.002)	(0.001)	(0.089)	(0.003)	(0.103)
Math Ability	0.051	-0.189	0.879	0.016	0.017	-0.219	-0.006	-0.817
	(0.113)	(0.065)	(0.020)	(0.002)	(0.001)	(0.089)	(0.003)	(0.107)

Notes: This table lists the parameter estimates of the measurement system (equation (9)) for the sub-sample of females. Standard errors in parentheses.

Table 10: Education Attainment, Marginal Effects

	Males						Females					
	No Qual.	CSE	O-lvl	A-lvl	H.Edu	H.Deg	No Qual.	CSE	O-lvl	A-lvl	H.Edu	H.Deg
Externalizing Behavior	0.014	0.021	0.004	-0.006	-0.022	-0.011	0.003	-0.025	0.013	0.006	0.009	-0.005
	(0.003)	(0.010)	(0.019)	(0.017)	(0.015)	(0.007)	(0.003)	(0.012)	(0.021)	(0.013)	(0.016)	(0.006)
Internalizing Behavior	0.003	0.004	0.009	-0.005	-0.008	-0.003	0.014	0.034	-0.015	-0.015	-0.019	0.002
	(0.004)	(0.011)	(0.021)	(0.019)	(0.017)	(0.008)	(0.004)	(0.012)	(0.021)	(0.014)	(0.017)	(0.007)
Cognition	-0.063	-0.075	-0.064	0.051	0.069	0.081	-0.071	-0.102	0.003	0.055	0.057	0.058
	(0.005)	(0.008)	(0.014)	(0.012)	(0.010)	(0.007)	(0.006)	(0.009)	(0.015)	(0.009)	(0.011)	(0.006)
Mother Education	-0.027	-0.033	-0.046	0.027	0.049	0.031	-0.034	-0.044	-0.082	0.059	0.057	0.044
	(0.009)	(0.015)	(0.024)	(0.020)	(0.016)	(0.007)	(0.009)	(0.017)	(0.024)	(0.013)	(0.017)	(0.007)
Father Education	-0.031	-0.034	-0.006	-0.015	0.033	0.053	-0.029	-0.067	-0.050	0.048	0.060	0.039
	(0.012)	(0.018)	(0.028)	(0.023)	(0.018)	(0.009)	(0.011)	(0.020)	(0.027)	(0.015)	(0.018)	(0.007)
No Father Info.	-0.021	-0.026	-0.014	0.063	-0.044	0.042	-0.029	-0.095	0.061	0.065	-0.041	0.040
	(0.018)	(0.033)	(0.062)	(0.052)	(0.058)	(0.021)	(0.028)	(0.050)	(0.065)	(0.032)	(0.055)	(0.015)
Father in Skilled Oc.	-0.017	-0.049	-0.051	0.058	0.051	0.008	-0.018	-0.039	0.011	0.000	0.026	0.020
	(0.006)	(0.012)	(0.024)	(0.023)	(0.020)	(0.011)	(0.006)	(0.015)	(0.025)	(0.016)	(0.020)	(0.010)
Father in Managerial Oc.	-0.051	-0.092	-0.086	0.095	0.082	0.052	-0.052	-0.087	0.002	0.011	0.081	0.046
	(0.012)	(0.019)	(0.031)	(0.028)	(0.024)	(0.012)	(0.012)	(0.022)	(0.032)	(0.019)	(0.024)	(0.011)
Working Mother	-0.003	-0.006	0.015	0.004	-0.001	-0.010	-0.007	0.001	0.014	-0.017	0.010	-0.001
	(0.006)	(0.011)	(0.020)	(0.017)	(0.014)	(0.007)	(0.006)	(0.013)	(0.020)	(0.012)	(0.015)	(0.005)
London Dummy	-0.005	0.048	-0.000	-0.008	-0.013	-0.022	-0.012	0.059	0.013	-0.011	-0.046	-0.003
	(0.006)	(0.011)	(0.020)	(0.018)	(0.015)	(0.007)	(0.007)	(0.013)	(0.020)	(0.012)	(0.015)	(0.005)
Financial Difficulties	0.037	0.053	0.054	-0.054	-0.051	-0.039	0.049	0.050	-0.036	-0.063	0.018	-0.017
	(0.007)	(0.014)	(0.027)	(0.026)	(0.024)	(0.013)	(0.008)	(0.015)	(0.028)	(0.021)	(0.022)	(0.011)
Constant	-0.037	-0.031	0.280	-0.023	-0.096	-0.093	-0.038	-0.034	0.386	-0.091	-0.113	-0.110
	(0.007)	(0.014)	(0.027)	(0.023)	(0.021)	(0.010)	(0.007)	(0.017)	(0.027)	(0.015)	(0.020)	(0.010)

Notes: This table lists marginal effects estimates from a multinomial logit model used to link socio-emotional and cognitive skills to educational attainment. We estimate educational attainment on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 11: Log Hourly Wages

	Males		Females	
	[1]	[2]	[1]	[2]
Externalizing Behavior	0.055	0.064	0.023	0.013
	(0.018)	(0.017)	(0.024)	(0.020)
Internalizing Behavior	-0.099	-0.096	-0.043	-0.021
	(0.020)	(0.018)	(0.027)	(0.022)
Cognition	0.106	0.025	0.163	0.044
	(0.011)	(0.011)	(0.015)	(0.013)
CSE	.	0.035	.	0.062
		(0.032)		(0.043)
O-Level	.	0.163	.	0.182
		(0.029)		(0.036)
A-Level	.	0.222	.	0.330
		(0.030)		(0.045)
Higher Education	.	0.340	.	0.569
		(0.032)		(0.041)
Higher Degree	.	0.470	.	0.729
		(0.037)		(0.046)
London Dummy	0.208	0.200	0.172	0.149
	(0.017)	(0.016)	(0.021)	(0.018)
Financial Difficulties	-0.093	-0.048	-0.097	-0.045
	(0.022)	(0.020)	(0.028)	(0.024)
Constant	1.888	1.671	1.544	1.266
	(0.010)	(0.026)	(0.015)	(0.035)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to hourly wages. We regress log hourly wages on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 12: Log Weekly Hours Worked

	Males		Females	
	[1]	[2]	[1]	[2]
Externalizing Behavior	0.012	0.015	0.047	0.047
	(0.008)	(0.008)	(0.024)	(0.025)
Internalizing Behavior	-0.014	-0.018	-0.023	-0.020
	(0.010)	(0.009)	(0.027)	(0.026)
Cognition	-0.015	-0.007	0.078	0.021
	(0.005)	(0.006)	(0.016)	(0.017)
CSE	.	0.009	.	0.037
		(0.019)		(0.045)
O-Level	.	-0.016	.	0.098
		(0.017)		(0.040)
A-Level	.	-0.030	.	0.226
		(0.019)		(0.058)
Higher Education	.	-0.027	.	0.199
		(0.019)		(0.049)
Higher Degree	.	-0.047	.	0.301
		(0.021)		(0.057)
London Dummy	0.014	0.015	0.049	0.041
	(0.009)	(0.009)	(0.025)	(0.025)
Financial Difficulties	-0.002	-0.008	-0.020	0.004
	(0.010)	(0.011)	(0.030)	(0.030)
Constant	3.755	3.776	3.206	3.078
	(0.005)	(0.016)	(0.016)	(0.037)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to hours worked. We regress log weekly hours worked on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 13: Log Hourly Wages, Varying Controls of Intermediate Outcomes

	Males					Females				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Externalizing Behavior	0.064	0.052	0.052	0.049	0.049	0.013	0.023	0.034	0.040	0.036
	(0.017)	(0.016)	(0.016)	(0.015)	(0.015)	(0.020)	(0.022)	(0.020)	(0.020)	(0.018)
Internalizing Behavior	-0.096	-0.084	-0.082	-0.079	-0.073	-0.021	-0.032	-0.038	-0.042	-0.033
	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.022)	(0.024)	(0.022)	(0.022)	(0.021)
Cognition	0.025	0.027	0.027	0.027	0.009	0.044	0.042	0.034	0.029	0.021
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
CSE	0.035	0.040	0.041	0.028	0.006	0.062	0.063	0.057	0.036	0.005
	(0.032)	(0.031)	(0.031)	(0.031)	(0.030)	(0.043)	(0.043)	(0.040)	(0.039)	(0.034)
O-Level	0.163	0.154	0.157	0.142	0.099	0.182	0.181	0.148	0.122	0.052
	(0.029)	(0.029)	(0.029)	(0.029)	(0.028)	(0.036)	(0.036)	(0.034)	(0.033)	(0.030)
A-Level	0.222	0.216	0.219	0.211	0.144	0.330	0.329	0.272	0.261	0.149
	(0.030)	(0.030)	(0.031)	(0.031)	(0.030)	(0.045)	(0.045)	(0.041)	(0.040)	(0.037)
Higher Education	0.340	0.328	0.332	0.325	0.209	0.569	0.568	0.516	0.487	0.274
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.041)	(0.041)	(0.037)	(0.036)	(0.035)
Higher Degree	0.470	0.459	0.462	0.509	0.379	0.729	0.729	0.647	0.689	0.448
	(0.037)	(0.037)	(0.037)	(0.039)	(0.038)	(0.046)	(0.046)	(0.042)	(0.041)	(0.039)
Partner Dummy	.	0.148	0.137	0.128	0.107	.	0.020	0.085	0.076	0.067
		(0.022)	(0.024)	(0.024)	(0.023)		(0.028)	(0.026)	(0.026)	(0.024)
Number of Children	.	.	0.009	0.007	0.011	.	.	-0.108	-0.081	-0.066
			(0.007)	(0.007)	(0.007)			(0.008)	(0.008)	(0.007)
Experience	.	.	.	0.001	0.001	.	.	.	0.002	0.001
				(0.000)	(0.000)				(0.000)	(0.000)
Skilled Manual Occu.	.	.	.	.	0.061	.	.	.	.	0.073
					(0.022)					(0.035)
Skilled Non-manual Occu.	.	.	.	.	0.203	.	.	.	.	0.176
					(0.027)					(0.023)
Managerial Occu.	.	.	.	.	0.257	.	.	.	.	0.381
					(0.023)					(0.025)
London Dummy	0.200	0.202	0.203	0.204	0.178	0.149	0.149	0.133	0.132	0.126
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.018)	(0.018)	(0.017)	(0.017)	(0.016)
Financial Difficulties	-0.048	-0.047	-0.048	-0.045	-0.035	-0.045	-0.045	-0.026	-0.020	-0.018
	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.024)	(0.024)	(0.022)	(0.022)	(0.020)
Constant	1.671	1.544	1.540	1.369	1.287	1.266	1.249	1.382	1.149	1.086
	(0.026)	(0.033)	(0.033)	(0.046)	(0.045)	(0.035)	(0.043)	(0.041)	(0.045)	(0.044)

Notes: This table listss parameter estimates from a linear model used to link socio-emotional and cognitive skills to hourly wages with different sets of controls. We regress log hourly wages of workers on a set of observable variables along with the three skills. The goal is to understand how the relationship between socio-emotional skills and wages changes as we change the set of additional regressors. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 14: Log Weekly Hours Worked, Varying Controls of Intermediate Outcomes

	Males					Females				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Externalizing Behavior	0.015	0.012	0.015	0.012	0.013	0.047	0.061	0.078	0.086	0.083
	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.025)	(0.026)	(0.022)	(0.023)	(0.022)
Internalizing Behavior	-0.018	-0.014	-0.017	-0.014	-0.015	-0.020	-0.045	-0.045	-0.053	-0.045
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.026)	(0.027)	(0.024)	(0.026)	(0.024)
Cognition	-0.007	-0.006	-0.007	-0.006	-0.004	0.021	0.019	0.007	0.007	-0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.017)	(0.017)	(0.015)	(0.015)	(0.014)
CSE	0.009	0.010	0.010	0.009	0.006	0.037	0.032	0.019	0.006	-0.018
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.045)	(0.044)	(0.038)	(0.038)	(0.036)
O-Level	-0.016	-0.017	-0.015	-0.017	-0.020	0.098	0.089	0.023	0.005	-0.040
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.040)	(0.040)	(0.034)	(0.034)	(0.033)
A-Level	-0.030	-0.030	-0.029	-0.030	-0.033	0.226	0.214	0.097	0.086	0.003
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.058)	(0.058)	(0.048)	(0.047)	(0.046)
Higher Education	-0.027	-0.028	-0.027	-0.028	-0.032	0.199	0.186	0.081	0.059	-0.112
	(0.019)	(0.020)	(0.019)	(0.020)	(0.020)	(0.049)	(0.049)	(0.041)	(0.042)	(0.044)
Higher Degree	-0.047	-0.049	-0.046	-0.043	-0.049	0.301	0.292	0.124	0.141	-0.045
	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.057)	(0.056)	(0.048)	(0.048)	(0.052)
Partner Dummy	.	0.022	0.014	0.013	0.011	.	-0.156	-0.020	-0.026	-0.032
		(0.015)	(0.016)	(0.016)	(0.016)		(0.035)	(0.030)	(0.030)	(0.029)
Number of Children	.	.	0.006	0.006	0.006	.	.	-0.222	-0.207	-0.193
			(0.004)	(0.004)	(0.004)			(0.012)	(0.012)	(0.011)
Experience	.	.	.	0.000	0.000	.	.	.	0.001	0.001
				(0.000)	(0.000)				(0.000)	(0.000)
Skilled Manual Occu.	.	.	.	.	0.025	.	.	.	.	0.234
					(0.013)					(0.042)
Skilled Non-manual Occu.	.	.	.	.	-0.023	.	.	.	.	0.131
					(0.016)					(0.025)
Managerial Occu.	.	.	.	.	0.015	.	.	.	.	0.316
					(0.014)					(0.032)
London Dummy	0.015	0.015	0.016	0.016	0.015	0.041	0.039	0.006	0.005	0.000
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.025)	(0.025)	(0.022)	(0.021)	(0.021)
Financial Difficulties	-0.008	-0.007	-0.008	-0.007	-0.007	0.004	0.001	0.040	0.044	0.038
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.030)	(0.030)	(0.026)	(0.026)	(0.025)
Constant	3.776	3.757	3.755	3.739	3.735	3.078	3.220	3.491	3.369	3.295
	(0.016)	(0.022)	(0.021)	(0.026)	(0.027)	(0.037)	(0.049)	(0.045)	(0.053)	(0.051)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to weekly work hours with different sets of controls. We regress log weekly hours worked of workers on a set of observable variables along with the three skills. The goal is to understand how the relationship between socio-emotional skills and hours changes as we change the set of additional regressors. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 15: Number of Children

	Males	Females
Externalizing Behavior	0.112	0.102
	(0.037)	(0.040)
Internalizing Behavior	-0.131	-0.056
	(0.044)	(0.046)
Cognition	-0.042	-0.037
	(0.029)	(0.028)
CSE	-0.127	-0.036
	(0.084)	(0.072)
O-Level	-0.101	-0.243
	(0.072)	(0.065)
A-Level	-0.174	-0.403
	(0.082)	(0.086)
Higher Education	-0.174	-0.455
	(0.088)	(0.078)
Higher Degree	-0.351	-0.626
	(0.096)	(0.088)
London Dummy	-0.066	-0.055
	(0.043)	(0.039)
Financial Difficulties	0.115	0.064
	(0.055)	(0.052)
Constant	1.491	1.877
	(0.067)	(0.060)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to the number of children. We model the number of children as a linear function of a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 16: OCCUPATION DECISION, MULTINOMIAL LOGIT

	Males			Females		
	Sk. Manual	Sk. Non-Manual	Managerial	Sk. Manual	Sk. Non-Manual	Managerial
Externalizing Behavior	0.143	-0.045	0.126	0.260	0.077	0.099
	(0.080)	(0.115)	(0.090)	(0.130)	(0.087)	(0.107)
Internalizing Behavior	-0.181	-0.254	-0.363	-0.375	-0.323	-0.366
	(0.097)	(0.137)	(0.113)	(0.158)	(0.105)	(0.127)
Cognition	-0.034	0.533	0.389	-0.133	0.203	0.101
	(0.078)	(0.109)	(0.087)	(0.109)	(0.068)	(0.082)
CSE	0.565	0.998	0.561	0.050	0.653	0.816
	(0.170)	(0.373)	(0.250)	(0.245)	(0.163)	(0.256)
O-Level	1.142	1.780	1.464	0.641	1.469	1.722
	(0.157)	(0.341)	(0.222)	(0.222)	(0.154)	(0.235)
A-Level	1.724	2.563	2.420	1.275	1.706	2.771
	(0.209)	(0.371)	(0.257)	(0.318)	(0.224)	(0.287)
Higher Education	1.210	2.664	3.361	1.284	0.964	3.945
	(0.268)	(0.402)	(0.290)	(0.314)	(0.240)	(0.272)
Higher Degree	0.610	3.276	4.705	0.272	1.578	4.891
	(0.545)	(0.564)	(0.471)	(0.704)	(0.356)	(0.377)
London Dummy	0.123	0.188	0.669	0.001	0.242	0.352
	(0.133)	(0.169)	(0.141)	(0.173)	(0.106)	(0.122)
Financial Difficulties	-0.399	-0.640	-0.522	0.212	-0.281	0.008
	(0.131)	(0.205)	(0.160)	(0.180)	(0.126)	(0.155)
Constant	-0.016	-1.938	-1.060	-1.758	-0.589	-2.096
	(0.145)	(0.319)	(0.206)	(0.205)	(0.145)	(0.221)

Notes: This table lists parameter estimates from a multinomial logit model used to link socio-emotional and cognitive skills to the occupation decision at age 33. We estimate the relationship between occupation sorting and a set of observable variables along with the unobserved skills. The base category is the "unskilled manual occupation." The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 17: Employment Decision, Multinomial Logit

	Males		Females	
	Self-Emp	Employee	Self-Emp	Employee
Externalizing Behavior	0.123	-0.057	0.379	0.223
	(0.113)	(0.100)	(0.125)	(0.078)
Internalizing Behavior	-0.359	-0.245	-0.354	-0.300
	(0.132)	(0.115)	(0.156)	(0.091)
Cognition	0.095	0.198	-0.109	-0.016
	(0.105)	(0.094)	(0.099)	(0.058)
CSE	0.674	0.720	0.558	0.256
	(0.280)	(0.232)	(0.291)	(0.147)
O-Level	0.767	0.562	0.622	0.364
	(0.236)	(0.201)	(0.270)	(0.134)
A-Level	1.215	1.165	0.797	0.173
	(0.294)	(0.260)	(0.330)	(0.174)
Higher Education	0.673	1.069	0.686	0.669
	(0.329)	(0.287)	(0.326)	(0.165)
Higher Degree	0.534	0.920	0.765	0.521
	(0.367)	(0.323)	(0.360)	(0.189)
London Dummy	0.126	-0.051	-0.079	-0.327
	(0.167)	(0.149)	(0.151)	(0.083)
Financial Difficulties	-0.388	-0.360	-0.064	0.248
	(0.190)	(0.162)	(0.209)	(0.108)
Constant	0.201	1.641	-1.911	0.405
	(0.226)	(0.192)	(0.254)	(0.123)

Notes: This table lists parameter estimates from a multinomial logit model used to link socio-emotional and cognitive skills to the employment decision. We model the the employment decision as a linear function of a set of observable variables along with the unobserved skills. The base category is "not employed" at age 33. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors in parentheses.

Table 18: Log Weekly Earnings, Imputing Missing Earnings

	Males		Females	
	[1]	[2]	[3]	[4]
Externalizing Behavior	0.071	0.066	0.050	0.050
	(0.016)	(0.017)	(0.037)	(0.033)
Internalizing Behavior	-0.111	-0.105	-0.034	-0.054
	(0.018)	(0.019)	(0.043)	(0.037)
Cognition	0.020	0.039	0.061	0.039
	(0.012)	(0.013)	(0.024)	(0.022)
CSE	0.049	0.066	0.102	0.064
	(0.032)	(0.034)	(0.065)	(0.058)
O-Level	0.147	0.142	0.285	0.284
	(0.029)	(0.030)	(0.060)	(0.053)
A-Level	0.192	0.219	0.564	0.540
	(0.032)	(0.034)	(0.076)	(0.066)
Higher Education	0.311	0.323	0.774	0.746
	(0.034)	(0.037)	(0.072)	(0.062)
Higher Degree	0.420	0.447	1.039	0.988
	(0.039)	(0.044)	(0.085)	(0.070)
London Dummy	0.215	0.204	0.190	0.216
	(0.017)	(0.018)	(0.033)	(0.029)
Financial Difficulties	-0.056	-0.053	-0.042	-0.064
	(0.020)	(0.021)	(0.043)	(0.039)
Constant	5.443	5.365	4.339	4.258
	(0.027)	(0.029)	(0.055)	(0.049)

Notes: This table lists parameter estimates from a linear regression used to link socio-emotional and cognitive skills to weekly earnings under alternative specifications. We regress log weekly earnings of workers on a set of observable variables along with the three skills. In Model [1], the dependent variable is reported gross weekly earnings for males that were working at age 33. In Model [2], we impute weekly earnings for males that were not working at age 33 using self-reported weekly earnings from previous jobs and include those observations in the regression. In Model [3], the dependent variable is reported gross weekly earnings for females that were working at age 33. In Model [4], we impute weekly earnings for females that were not working at age 33 using self-reported weekly earnings from previous jobs and include those observations in the regression. With the imputation, we manage to compute earnings for 92% of the individuals in our sample. Standard errors in parentheses.

Table 19: Sample Covariance vs Simulated Covariance, Males

	Log-Hourly Wages		Log-Hours Worked	
	Sample	Simu.	Sample	Simu.
No Formal Edu.	-0.028	-0.025	0.002	0.002
CSE	-0.025	-0.019	0.003	0.003
O-Level	-0.022	-0.021	0.001	0.002
A-Level	0.003	0.005	-0.002	-0.002
Higher Education	0.024	0.022	-0.001	-0.001
Higher Degree	0.048	0.038	-0.004	-0.003

Notes: This table compares the data and the benchmark-model simulated variance-covariance matrix of choices and outcomes for males.

Table 20: Sample Covariance vs Simulated Covariance, Females

	Log-Hourly Wages		Log-Hours Worked	
	Sample	Simu.	Sample	Simu.
No Formal Edu.	-0.040	-0.034	-0.017	-0.015
CSE	-0.035	-0.029	-0.015	-0.013
O-Level	-0.039	-0.030	-0.012	-0.008
A-Level	0.007	0.008	0.010	0.009
Higher Education	0.048	0.044	0.012	0.012
Higher Degree	0.058	0.041	0.022	0.015

Notes: This table compares the data and the benchmark-model simulated variance-covariance matrix of choices and outcomes for females.

Table 21: Subgroup Analysis: Educational Attainment, Multinomial Logit

	[High SES]					[Low SES]				
	CSE	O-lvl	A-lvl	H.Edu.	H.Deg.	CSE	O-lvl	A-lvl	H.Edu.	H.Deg.
Externalizing Behavior	-0.084	0.021	0.052	-0.037	-0.346	-0.156	-0.198	-0.514	-0.293	-0.031
	(0.079)	(0.065)	(0.079)	(0.084)	(0.110)	(0.201)	(0.199)	(0.322)	(0.294)	(0.576)
Internalizing Behavior	-0.018	-0.248	-0.379	-0.367	-0.223	-0.075	-0.245	0.002	-0.176	-0.667
	(0.100)	(0.087)	(0.102)	(0.105)	(0.135)	(0.227)	(0.209)	(0.343)	(0.298)	(0.642)
Cognition	0.757	1.555	2.004	2.095	3.490	0.939	1.449	1.876	1.858	2.945
	(0.092)	(0.082)	(0.093)	(0.096)	(0.123)	(0.167)	(0.160)	(0.210)	(0.216)	(0.387)
Mother Education	0.239	0.408	0.836	0.798	1.122	0.544	0.617	0.665	1.183	1.665
	(0.186)	(0.170)	(0.183)	(0.182)	(0.194)	(0.360)	(0.350)	(0.444)	(0.428)	(0.557)
Father Education	0.098	0.515	0.720	0.759	1.314	0.353	0.073	-0.247	0.791	1.041
	(0.236)	(0.212)	(0.223)	(0.224)	(0.233)	(0.412)	(0.427)	(0.592)	(0.487)	(0.616)
No Father Info.	0.426	1.091	1.296	0.618	1.668	-0.028	0.250	1.613	0.520	0.162
	(0.457)	(0.410)	(0.454)	(0.493)	(0.492)	(0.757)	(0.703)	(0.745)	(0.983)	(1.566)
Father in Skilled Oc.	0.030	0.233	0.439	0.658	0.515	-0.268	0.188	0.579	-0.121	0.211
	(0.136)	(0.126)	(0.152)	(0.161)	(0.204)	(0.211)	(0.206)	(0.325)	(0.287)	(0.518)
Father in Managerial Oc.	0.094	0.608	0.938	1.314	1.508	1.852	2.383	3.178	1.995	2.679
	(0.225)	(0.205)	(0.228)	(0.231)	(0.263)	(0.960)	(0.969)	(1.053)	(1.035)	(1.119)
Working Mother	0.081	0.113	0.019	0.122	0.050	0.074	0.134	0.338	0.416	-0.029
	(0.129)	(0.117)	(0.132)	(0.134)	(0.148)	(0.211)	(0.203)	(0.306)	(0.290)	(0.458)
London Dummy	0.626	0.185	0.145	0.083	-0.014	0.690	0.526	0.286	-0.256	-0.105
	(0.133)	(0.125)	(0.139)	(0.141)	(0.156)	(0.235)	(0.228)	(0.343)	(0.340)	(0.485)
Police Involvement	-0.798	-1.109	-1.451	-1.634	-1.976	-0.903	-1.044	-1.533	-1.557	-14.047
	(0.230)	(0.208)	(0.264)	(0.309)	(0.387)	(0.374)	(0.371)	(0.584)	(0.693)	(3.977)
Police Info. Missing	-0.568	-0.695	-0.857	-0.875	-0.965	-0.245	-0.566	-0.684	-0.811	-1.009
	(0.135)	(0.123)	(0.142)	(0.143)	(0.163)	(0.229)	(0.216)	(0.331)	(0.319)	(0.539)
Female	0.034	-0.156	-1.100	-0.635	-0.852	-0.625	-0.822	-1.900	-0.697	-1.288
	(0.128)	(0.119)	(0.138)	(0.137)	(0.162)	(0.231)	(0.228)	(0.334)	(0.316)	(0.552)
Constant	0.954	2.411	1.765	1.248	-0.122	0.529	0.986	-0.400	-0.325	-2.403
	(0.191)	(0.172)	(0.196)	(0.205)	(0.250)	(0.293)	(0.294)	(0.377)	(0.385)	(0.588)

Notes: This table lists parameter estimates from a multinomial logit model used to link socio-emotional and cognitive skills to educational attainment, by high-SES and low-SES subsamples. We estimate educational attainment on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors are in parentheses.

Table 22: Subgroup Analysis: Log Hourly Wages

	[High SES]	[Low SES]
Externalizing Behavior	0.052	0.015
	(0.013)	(0.029)
Internalizing Behavior	-0.074	-0.066
	(0.015)	(0.032)
Cognition	0.052	0.046
	(0.009)	(0.019)
CSE	0.050	0.032
	(0.031)	(0.050)
O-Level	0.158	0.121
	(0.028)	(0.042)
A-Level	0.263	0.126
	(0.030)	(0.059)
Higher Education	0.421	0.416
	(0.030)	(0.054)
Higher Degree	0.534	0.515
	(0.034)	(0.080)
London Dummy	0.172	0.211
	(0.012)	(0.033)
Police Involvement	-0.018	-0.012
	(0.028)	(0.057)
Police Info. Missing	-0.007	0.011
	(0.013)	(0.032)
Female	-0.322	-0.377
	(0.012)	(0.030)
Constant	1.650	1.615
	(0.027)	(0.040)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to hourly wages, by high-SES and low-SES subsamples. We regress log hourly wages on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors are in parentheses.

Table 23: SUBGROUP ANALYSIS: LOG WEEKLY HOURS WORKED

	[High SES]	[Low SES]
Externalizing Behavior	0.047	-0.030
	(0.013)	(0.032)
Internalizing Behavior	-0.040	0.043
	(0.015)	(0.036)
Cognition	0.002	0.050
	(0.010)	(0.023)
CSE	0.015	0.038
	(0.026)	(0.049)
O-Level	0.032	0.036
	(0.024)	(0.043)
A-Level	0.072	0.072
	(0.030)	(0.080)
Higher Education	0.066	0.155
	(0.028)	(0.071)
Higher Degree	0.093	0.116
	(0.032)	(0.108)
London Dummy	0.025	0.037
	(0.013)	(0.037)
Police Involvement	0.061	0.041
	(0.042)	(0.074)
Police Info. Missing	0.011	0.017
	(0.014)	(0.034)
Female	-0.520	-0.579
	(0.017)	(0.046)
Constant	3.690	3.687
	(0.028)	(0.058)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to hours worked, by high-SES and low-SES subsamples. We regress log weekly hours worked on a set of observable variables along with the unobserved skills. The coefficients on the three skills have been standardized to represent a 1 standard deviation effect. Standard errors are in parentheses.

Table 24: Subgroup Analysis: Police Involvement at 16

	[High SES]	[Low SES]
Externalizing Behavior	0.159	0.020
	(0.024)	(0.015)
Internalizing Behavior	-0.042	-0.000
	(0.015)	(0.017)
Cognition	-0.025	-0.018
	(0.004)	(0.016)
Mother Education	0.004	-0.051
	(0.007)	(0.042)
Father Education	0.000	-0.006
	(0.009)	(0.037)
No Father Info.	0.005	0.060
	(0.017)	(0.045)
Father in Skilled Oc.	-0.018	-0.030
	(0.006)	(0.018)
Father in Managerial Oc.	-0.026	-0.087
	(0.009)	(0.088)
Working Mother	0.004	0.009
	(0.006)	(0.018)
London Dummy	-0.017	0.011
	(0.006)	(0.021)
Female	-0.051	-0.129
	(0.007)	(0.029)
Constant	0.090	0.179
	(0.019)	(0.037)

Notes: This table lists parameter estimates from a linear model used to link socio-emotional and cognitive skills to police involvement at age 16, by high-SES and low-SES subsamples. We regress log weekly hours worked on a set of observable variables along with the unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. Standard errors are in parentheses.

Table 25: Robustness Across Datasets: Years of Education

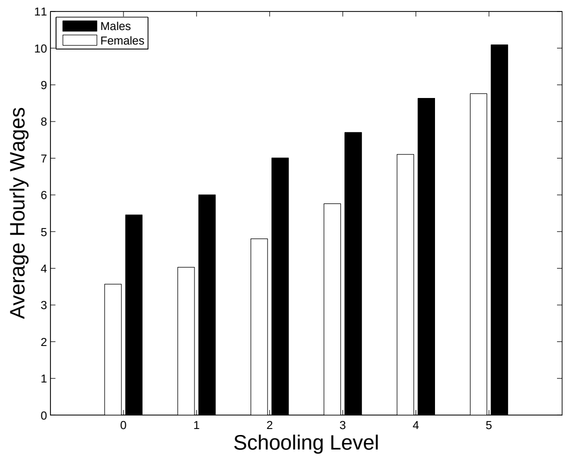
	[NCDS]	[BCS]	[NELS]	[PSID]	[CNLSY]
Males & Females					
Externalizing	-0.074	-0.122	-0.161	-0.176	-0.136
	(0.027)	(0.026)	(0.023)	(0.110)	(0.030)
Internalizing	-0.069	0.019	-0.165	-0.037	0.015
	(0.028)	(0.027)	(0.021)	(0.105)	(0.028)
Cognition	1.088	0.587	0.637	0.770	0.220
	(0.031)	(0.027)	(0.026)	(0.092)	(0.023)
N	7241	5789	5052	468	1597
Males					
Externalizing	-0.115	-0.148	-0.170	-0.370	-0.085
	(0.035)	(0.034)	(0.028)	(0.136)	(0.039)
Internalizing	-0.059	0.059	-0.166	0.109	0.030
	(0.038)	(0.039)	(0.032)	(0.134)	(0.039)
Cognition	1.169	0.585	0.548	0.646	0.198
	(0.041)	(0.038)	(0.038)	(0.127)	(0.031)
N	3573	2808	2373	216	737
Females					
Externalizing	0.004	-0.089	-0.148	-0.042	-0.197
	(0.042)	(0.039)	(0.039)	(0.199)	(0.046)
Internalizing	-0.085	-0.023	-0.161	-0.116	0.015
	(0.042)	(0.038)	(0.028)	(0.177)	(0.042)
Cognition	0.995	0.588	0.720	0.871	0.244
	(0.045)	(0.038)	(0.036)	(0.133)	(0.034)
N	3668	2981	2679	252	860

Notes: This table lists estimates from OLS regressions used to link socio-emotional and cognitive skills to years of education across datasets. For each dataset, we regress years of education on a set of observable variables along with proxies for the unobserved skills. Standard errors are in parentheses.

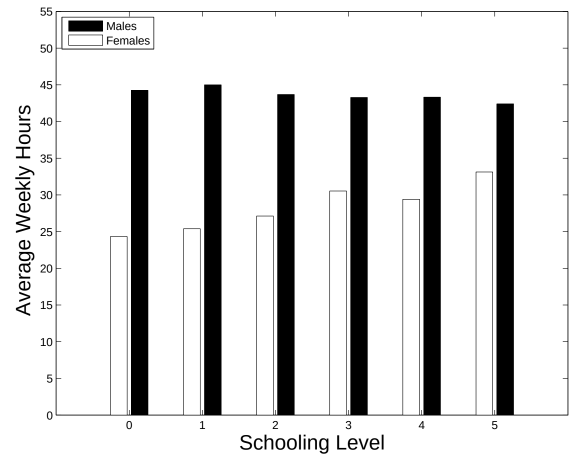
Table 26: Robustness Across Datasets: Log Earnings

	[NCDS]	[BCS]	[NELS]	[PSID]	[CNLSY]
Males & Females					
Externalizing	0.032	0.020	0.028	0.068	0.002
	(0.009)	(0.011)	(0.009)	(0.034)	(0.024)
Internalizing	-0.047	-0.033	-0.040	-0.090	-0.066
	(0.009)	(0.011)	(0.009)	(0.033)	(0.025)
Cognition	0.079	0.064	0.019	0.044	0.077
	(0.010)	(0.011)	(0.011)	(0.025)	(0.019)
N	4888	5140	5161	249	1269
Males					
Externalizing	0.020	0.012	0.028	0.089	-0.027
	(0.008)	(0.013)	(0.011)	(0.047)	(0.035)
Internalizing	-0.055	-0.029	-0.046	-0.136	-0.076
	(0.008)	(0.014)	(0.014)	(0.049)	(0.039)
Cognition	0.067	0.061	0.011	0.065	0.057
	(0.009)	(0.013)	(0.015)	(0.033)	(0.027)
N	2643	2665	2457	118	593
Females					
Externalizing	0.041	0.028	0.028	0.042	0.034
	(0.020)	(0.018)	(0.018)	(0.050)	(0.034)
Internalizing	-0.031	-0.035	-0.034	-0.048	-0.047
	(0.020)	(0.018)	(0.012)	(0.042)	(0.031)
Cognition	0.103	0.070	0.021	0.030	0.099
	(0.021)	(0.017)	(0.015)	(0.037)	(0.027)
N	2245	2475	2704	131	676

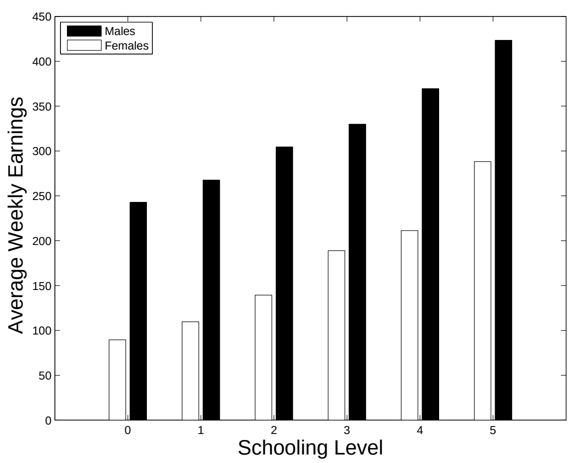
Notes: This table compares estimates from OLS regressions used to link socio-emotional and cognitive skills to log earnings in early adulthood across datasets. For each dataset, we regress log weekly earnings on education attainment along with proxies for the unobserved skills. Standard errors in parentheses.



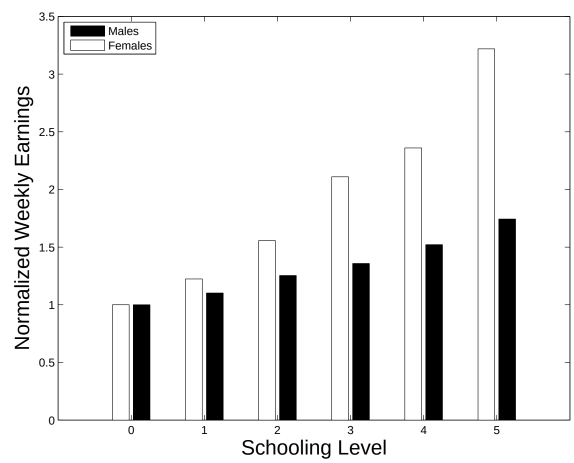
(a) Wages by schooling



(b) Hours by schooling

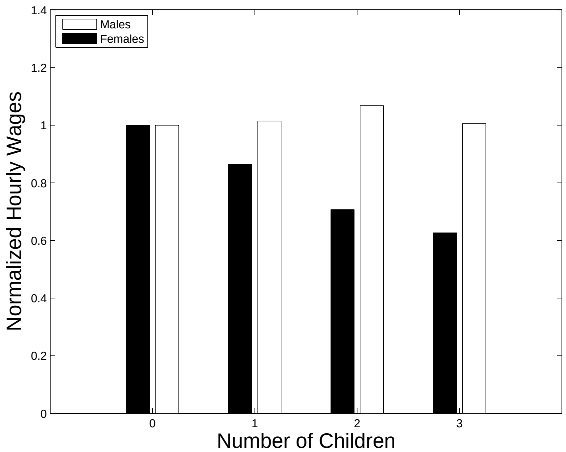


(c) Earnings by schooling

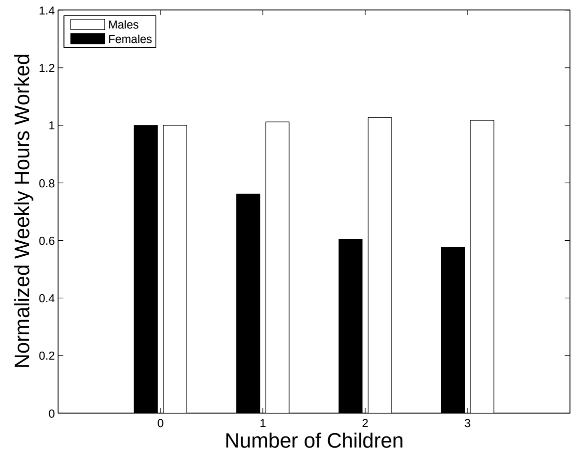


(d) Normalized earnings by schooling

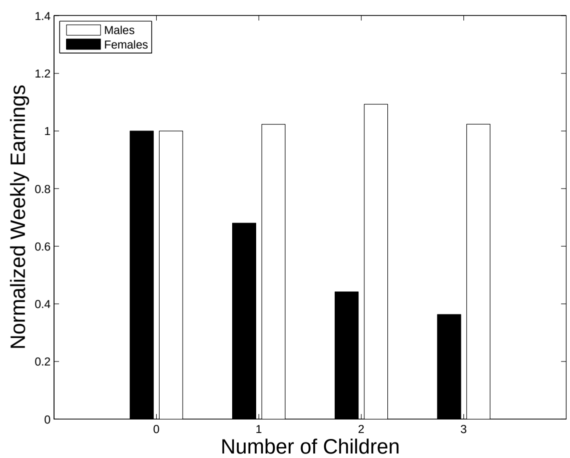
Figure 1: Gender Differences in Labor Market Outcomes by Schooling: Figure 1(a) compares hourly wages by schooling level and gender, Figure 1(b) compares weekly hours worked by schooling level and gender, and Figures 1(c) and 1(d) compare weekly earnings and normalized weekly earnings by schooling level and gender.



(a) Normalized wages by fertility

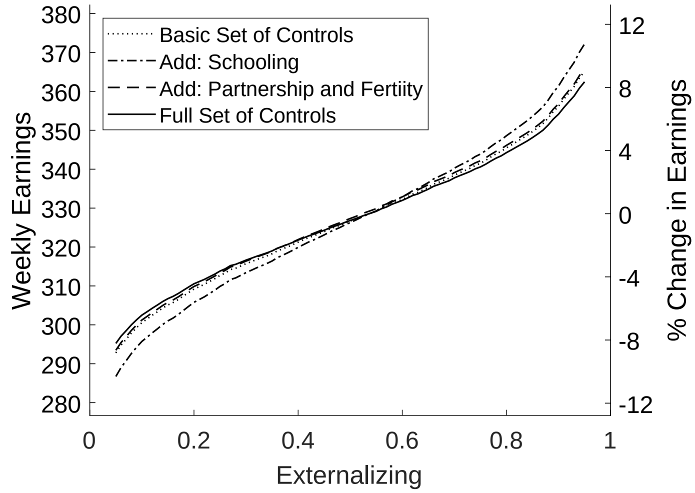


(b) Normalized hours by fertility

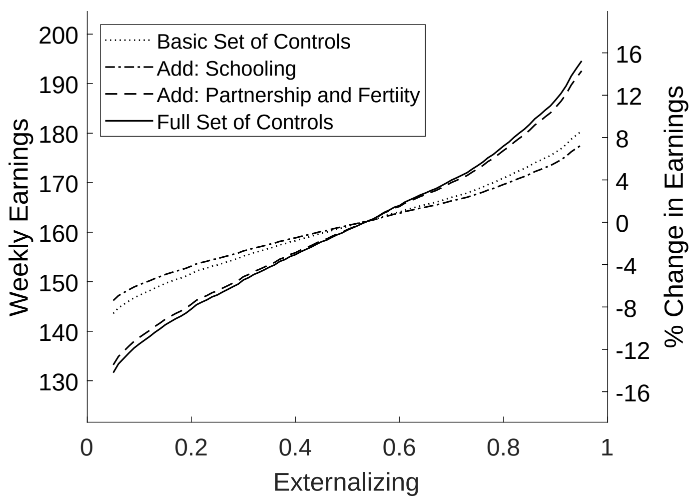


(c) Normalized earnings by fertility

Figure 2: Gender Differences in Labor Market Outcomes by Fertility: Figure 2(a) compares hourly wages by number of children and gender, Figure 2(b) compares weekly hours worked by number of children and gender, and Figure 2(c) compares normalized weekly earnings by number of children and gender.

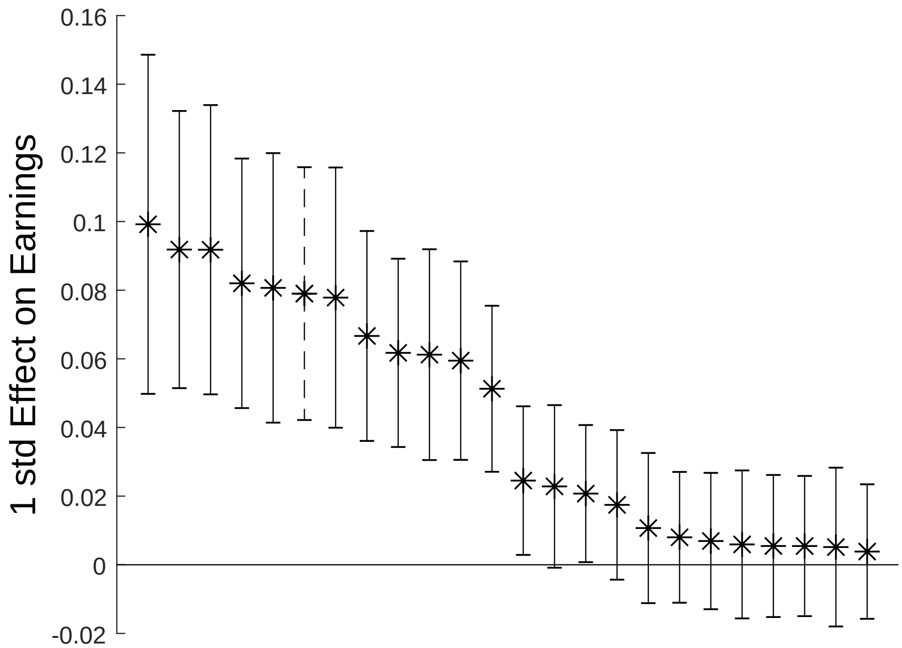


(a) Males

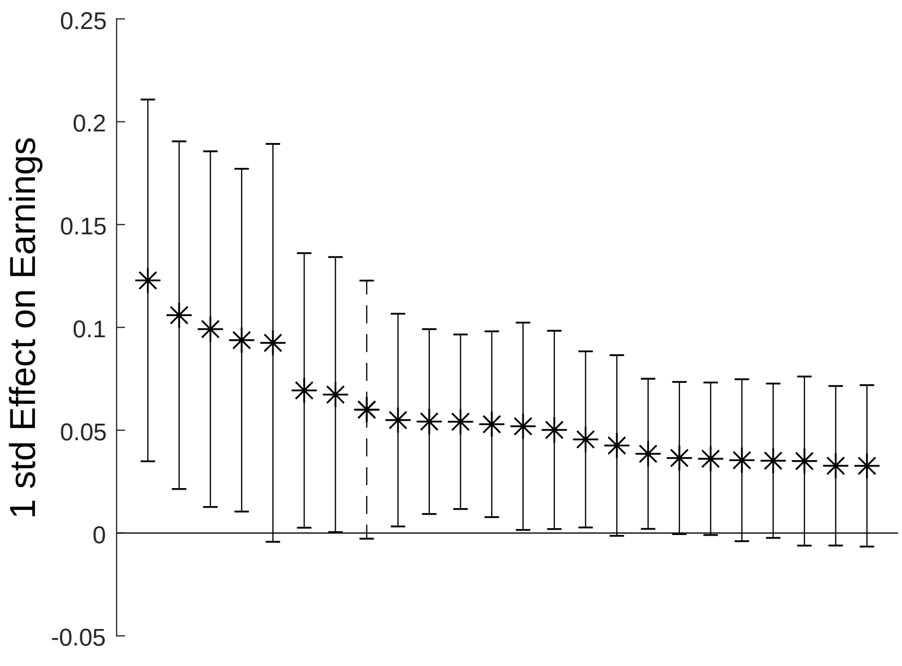


(b) Females

Figure 3: Decomposition of Effects of Externalizing on Weekly Earnings: Figure 3 visualizes the results from regressing weekly earnings on a varying set of controls presented in Tables 13 and 14. It illustrates how the predicted weekly earnings in regression models with different sets of controls vary, when we increase externalizing behavior from the lowest 5th percentile to the highest 95th percentile, keeping other latent skills and covariates at the population median.



(a) Males



(b) Females

Figure 4: Distribution of Effects of Externalizing on Earnings: Figure 4 visualizes the effects on weekly earnings from 1 standard deviation increase in externalizing behavior from specifications that span all possible combinations of the dedicated measurements for externalizing and internalizing behaviors. It summarizes the results reported in Tables S19 and S20. The dashed bars indicate results from our benchmark model.