The Economic Value of *Breaking Bad*: Misbehavior, Schooling and the Labor Market*

Nicholas W. Papageorge†
Department of Economics, Johns Hopkins University, IZA and NBER

Victor Ronda‡
Uber Technologies, Inc.

Yu Zheng§
School of Economics and Finance, Queen Mary University of London and CEPR

ABSTRACT: Prevailing research argues that childhood misbehavior in the classroom is bad for schooling and, presumably, bad overall. We examine lifecycle impacts of a widely studied socio-emotional skill, externalizing behavior, which captures childhood misbehavior in school and is linked to aggression, hyperactivity and lower educational attainment. Externalizing behavior has been the focus of hundreds of papers across several fields and its negative impact on educational attainment has justified policy to address or discourage it. Aligned with prior work, we find that externalizing behavior lowers educational attainment for males. However, we also provide novel evidence that it increases earnings for both males and females. The earnings premium holds across genders and occupations, is replicated in several data sets and is robust to a host of modeling approaches. That a single skill can be both helpful and harmful raises concerns about policies surrounding skill acquisition or modification, especially if the skill has opposite effects across childhood and adulthood. For example, well-meaning policies to improve schooling can have negative repercussions over the lifecycle. Ironically, such policies risk undermining a crucial purpose of education, which is to develop children's human capital so that they are more likely to succeed as adults. More broadly, our results illustrate how skill prices based on measurements from limited sets of economic outcomes can be misleading and potentially motivate policy with unintended or harmful consequences.

KEYWORDS: Labor, Education, Socio-Emotional Skills
JEL CLASSIFICATION: J10 J20 I20

*First draft: January 31, 2014. Current draft: October 28, 2024. 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, Stony Brook University, University of Cyprus, the Econometric Society World Congress, SOLE 2016, Barcelona GSE Summer Forum (Structural Microeconometrics). We acknowledge excellent research assistance from Mikkel Houmark. The usual caveats apply.

†[Corresponding author] papageorge@jhu.edu.
‡victor.ronda.econ@gmail.com.
§yu.zheng@qmul.ac.uk

1 Introduction

Economic research has shown that human capital consists of multiple factors that affect lifecycle behavior and outcomes and that the returns to these factors can vary across economic sectors and contexts. For example, to explain career choices, Willis and Rosen (1979) emphasize differences across occupations in returns to manual versus academic skill.[1] While many components of human capital, such as academic or mechanical skill, may be rewarded differently across sectors, it is difficult to imagine a context in which they have a direct negative impact. In contrast, socio-emotional skills—increasingly viewed as crucial components of human capital (Heckman and Rubinstein, 2001)—are linked to such a wide array of behaviors and tendencies, that it is natural to ask whether they can be rewarded in some contexts and penalized in others. This possibility has generally been ignored in research on human capital. However, it suggests potential unintended consequences of policies designed to either curb or promote specific socio-emotional skills. Of particular concern are skills that have opposite impacts in childhood and adulthood. In this case, well-intentioned interventions designed, for example, to improve school performance by targeting children's socio-emotional skills (which are relatively malleable (Heckman and Kautz, 2014)) could have negative repercussions over the lifecycle.

In this paper, we examine a widely-studied pair of socio-emotional skills measured by teachers among schoolchildren to capture misbehavior: *externalizing behavior* and *internalizing behavior*.[2] Externalizing behavior is linked to aggression and hyperactivity, while internalizing behavior captures anxiety, depression, shyness, unassertiveness and fearfulness. The conceptual development of externalizing and internalizing behaviors dates back to Achenbach (1978). Since then, they have been measured in dozens of data sets and have been the subject of hundreds of studies.[3] Nearly all focus on the negative impact of externalizing behavior on educational outcomes, which has led to a multitude of interventions and programs in classrooms and schools, mostly designed to address, discourage or curb it.[4] The key empirical fact we establish is

[1]This point has its origins in Roy (1951) and Mandelbrot (1962), which are later developed into a model of comparative advantage and self-selection in the labor market by Willis and Rosen (1979), Heckman and Sedlacek (1985), and many papers thereafter.

[2]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.

[3]In fact, as shown in Figure 1, when we compare the number of publications in the PubMed database that mention internalizing or externalizing to those that mention the "Big Five" personality traits, we find that since 1980 there are strong increasing trends in the number of both and that the number of publications related to internalizing and externalizing is roughly double the number of those related to the "Big Five".

[4]The constructs of externalizing and internalizing behaviors are well established in the developmental psychology literature (Ghodsian, 1977; Campbell, Shaw, and Gilliom, 2000; Eisenberg et al., 2001; Duncan and Magnuson, 2011). Stage and Quiroz (1997) and O'Connor and Hayes (2020) conduct meta-analysis of 99 and 17 journal articles, respectively, about interventions to target and address disruptive behavioral problems in public education, mostly in the US. Similar interventions have also been introduced in Europe (e.g., Humphrey et al. (2010) and Närhi, Kiiski, and Savolainen (2017) in the UK and Sørlie and Ogden (2015) in Norway). Externalizing and in-

that, while externalizing behavior lowers educational attainment for boys (with no significant impact for girls), the story changes once these children reach adulthood and enter the labor market: externalizing behavior increases earnings for both men and women. In other words, a well-studied socio-emotional skill, generally seen as a problem due to a vast set of studies showing its negative relationship to education, leads to higher earnings. *Breaking Bad* can be good—or at least lucrative.[5]

Our main analysis uses a longitudinal dataset from Britain, the National Child Development Survey (NCDS), which consists of the universe of children born in a single week in 1958 in Great Britain and has followed them through adulthood, allowing us to link childhood behavior in school, educational attainment, and earnings. The NCDS collected teachers' reports of children's misbehavior, and it was through analyses of these reports that the constructs of externalizing and internalizing behaviors initially emerged. We replicate earlier findings showing that a composite measure of childhood classroom misbehavior is negatively associated with both schooling and earnings (e.g., Segal (2013)). However, once we follow earlier literature and separate misbehavior into two separate skills, externalizing and internalizing, we find that externalizing, despite lowering educational attainment, increases earnings. To our knowledge, we are the first to study how externalizing and internalizing behaviors jointly affect schooling and earnings.[6] While the key patterns are evident in preliminary analyses of the data, our main approach is to estimate a flexible econometric model that approximates schooling, hours of work and wages using linear-in-parameters equations. The model includes a formal measurement system to map the teachers' reports described above and multiple aptitude test scores to cognition, externalizing and internalizing behavior, which enter the model as three latent factors that permit correlation across equations. Given gender differences in schooling, earnings, and childhood misbehavior (see, e.g., Bertrand and Pan (2013)), all parameters can vary by gender.

Our main estimates show that a 1 SD rise in externalizing leads to 9.8% higher earnings for men and 6.4% for women. The relationship is nearly the same whether or not we adjust for completed schooling, i.e., externalizing behavior increases earnings even when we account for its negative impact on schooling. Moreover, the estimated externalizing premium is robust to a host of modeling assumptions, including variations on the measurement system and various control sets, and is not driven by

ternalizing behaviors have also been studied in economic research (see, e.g., Neidell and Waldfogel (2010); Duncan and Dunifon (2012); Bertrand and Pan (2013); Gertler et al. (2014); Heckman, Pinto, and Savelyev (2013); Doyle (2020)).

[5]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 a television show in which the protagonist is an unsuccessful chemist with a talent for producing illicit drugs, illustrating how certain skills or behaviors may lead to low productivity in one sector and high productivity in another.

[6]While most of the literature on externalizing and internalizing behaviors has focused on its impacts on schooling outcomes and ignore potential impacts on the labor market, most of the literature on the importance of socio-emotional skills for labor market outcomes has focused on the Big 5 personality traits. We also show that the impact of externalizing and internalizing behaviors on earnings cannot be substantially explained by the Big 5 personality traits.

obvious sources of selection (e.g., into employment), nor reflect a correlation between externalizing and other potentially correlated factors that have been shown to affect earnings (e.g., the "Big 5" personality traits typically measured during adulthood). The NCDS data allow us to investigate whether the earnings premium is driven by hours versus wages. We find that high-externalizing females work more hours, while high-externalizing men command higher wages. This may suggest externalizing leads to higher productivity for men and higher non-pecuniary benefits of work for women, though additional analyses find little scope for changes in hours for men or wages for women, which may reflect the structure of the labor market relevant to our sample. Our interpretation is that externalizing behavior may thus capture a kind of proactiveness, resourcefulness or high energy, whereby individuals find ways to obtain what they want (e.g., higher earnings) under the constraints, social norms, and circumstances they face. Such tendencies among schoolchildren could easily be perceived as disobedience or a challenge to authority.

We conduct several analyses to investigate the generalizability of our findings. First, we explore family structures, asking if the externalizing premium is driven by people who focus on their careers and choose not to marry or have children, leading to higher earnings. While externalizing predicts higher marriage and fertility rates, neither of these outcomes explains the externalizing premium. Second, we ask if the externalizing premium is concentrated among certain occupations and instead find that it extends across tasks and virtually all occupations. This finding is noteworthy as it contrasts with earlier work that finds a positive correlation between teenage rule-breaking and earnings for a specific and small sector of the labor market: incorporated entrepreneurs (Levine and Rubinstein, 2017)).[7] Finally, we ask if the externalizing premium is limited to NCDS respondents, i.e., individuals born in Great Britain in the 1950s. We are able to replicate the externalizing penalty in school and premium at work across multiple cohorts (e.g. 1970 British Cohort Study (BCS)) and across countries (e.g. Panel Study of Income Dynamics (PSID) and National Educational Longitudinal Study of 1988 (NELS88)).

Additional analyses are prompted by earlier findings related to externalizing behavior in Heckman, Pinto, and Savelyev (2013), who show that an early childhood intervention (the Perry Preschool Program focused on low-income Black children in the U.S.) raised earnings and that about 20% of this increase is attributable to a reduction in externalizing behavior, i.e., that lowering externalizing can help individuals on the labor market. In contrast, we find that, for a 1958 representative British cohort, externalizing behavior increases earnings. To reconcile findings, we consider a sub-sample of the NCDS British cohort we select to mimic the financially disadvantaged group studied in Heckman, Pinto, and Savelyev (2013). Among individuals in this selected sample, externalizing behavior carries no significant earnings premium.

[7] The authors focus on people with incorporated small businesses and high cognition who used illicit substances or engaged in delinquency in high school, roughly 3.4% of their sample. Another contrast is that Levine and Rubinstein (2017) relate observed behaviors during adolescence, while we discuss an underlying skill, identified from teacher reports of childhood (age 7) behavior. Never-the-less, their paper, as ours, raises questions about the wisdom of viewing certain tendencies as only beneficial or only detrimental.

One possible reason is selection into criminality (Aizer, 2009; Heckman, Pinto, and Savelyev, 2013). We show that externalizing behavior indeed predicts measures of criminality, such as having interactions with the police due to delinquent behavior. However, these measures do not explain differences in returns to externalizing behavior across socioeconomic groups in our sample and, indeed, do not seem to explain earnings at all in the British cohort we study. Rather, externalizing is simply priced differently in the labor market depending on an individual's background, which is in line with Lundberg (2013), who demonstrates that the payoff to non-cognitive skills varies by socioeconomic status. Context dependence of the externalizing premium is troubling as it suggests that disadvantaged individuals miss out on the benefits to skills that their relatively privileged individuals enjoy.

Our work contributes to a large literature in economics that continues to expand our understanding of what constitutes human capital beyond traditional factors such as cognition, education and work experience to include, for example, health (Grossman, 2000) and socio-emotional skills (Heckman and Rubinstein, 2001).[8] Specifically, our work suggests how measurement of skill prices can be misleading if performed on a specific context or a convenient sample. For example, Duckworth and Quinn (2009) explore how a socio-emotional skill, grit, can be good, but the analysis is limited to highly selected groups (e.g., Air Force cadets) for whom focused attention on a single, long-term goal might be beneficial. Subsequent research uses a laboratory experiment to show that grit may also capture a lack of flexibility and lead to suboptimal behavior (Alaoui and Fons-Rosen, 2021), casting doubt on the external validity of earlier findings on grit. Yet, schools have engaged in attempts to develop grit in students as a "desirable trait" with little thought of potential downsides (see, e.g., Bashant (2014)). Related, much policy surrounding externalizing behavior is based on measuring its impact on one outcome, schooling, without considering others, or by measuring negative returns in a selected sample as in Heckman, Pinto, and Savelyev (2013), who examine disadvantaged Black children in the U.S. in the 1960s, for whom, according to our own analyses of a similarly selected sample, externalizing may not be rewarded. In contrast, when we study externalizing behavior in a larger and more representative population, and across economic contexts, we show that it has positive average labor market returns. A broad takeaway from our study is, thus, that skill prices measured in specific samples, contexts, or sectors may not be externally valid or generalizable. If they are none-the-less taken as such, they can translate to policy interventions with unintended and potentially harmful consequences.

More broadly, our findings on externalizing behavior challenge a large literature on child development focused on skills measured during childhood and their impact on schooling. The implicit assumption in much of this literature is that skills bad for schooling must be bad overall. Based on this assumption, researchers have generally not examined long-run impacts of skills shown to be detrimental for schooling. This is problematic in light of the vast literature on externalizing behavior in school leading

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

to the broad consensus that it is a problem to be solved (see Malti and Rubin (2018) for a review). Indeed, ignoring potential downsides, school districts, states and regions across the world have adopted programs that are meant to develop socio-emotional skills that promote behaviors deemed good for school (O'Connor and Hayes, 2020; Nangle, Erdley, and Schwartz-Mette, 2020).[9] Yet, our findings suggest that interventions aiming to improve schooling outcomes by addressing externalizing behavior may instead undermine a key role of schooling, which is to improve students' labor market outcomes. Indeed, such interventions may not be ethically defensible. Would parents agree to such an intervention if they were made aware that it might lead their children to earn less?

Section 2 introduces the NCDS dataset and presents a preliminary data analysis. Section 3 outlines the econometric model. Section 4 describes our main results, sensitivity analyses and robustness tests. Section 5 investigates the generalizability of our findings. Section 6 explores socioeconomic status, criminality and policy involvement. Section 7 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 several 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, 44, 46, 50, 55 and 62 (2020).

[9] Some well-known interventions include the Check-In/Check-Out targeted intervention, a.k.a the Behavior Education Program (Crone, Hawken, and Horner, 2003; Campbell and Anderson, 2011), and the Fast Track prevention program (CPPRG, 1992). Some elements of the intervention are designed to curb misbehavior, e.g., setting a goal of "not leaving my seat once without permission" under the Check-In/Check-Out system, while others focus on encouraging prosocial behaviors, e.g., providing social skills training under the Fast Track prevention program. However, when it comes to the evaluation of interventions, reduction in occurrences of misbehavior is invariably the most common measure.

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.[10]

2.2 Key Variables and Summary Statistics

2.2.1 Education, Labor Market and Other Lifecycle 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.[11] 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 and other lifecycle outcomes are found in Table 1. 51% of our sample is female. Females in our sample are less educated compared to males. On average, females are significantly less likely to be employed, while employed females' wages are 29% lower, hours are 36% lower, and earnings are 51% lower than those reported by employed males. Females on average have more children by age 33 than males, though they are similarly likely to have a

[10]There is significant attrition over time. Of the original 18,555 births, only 11,364 individuals were surveyed in 1991, at age 33. In results available upon request, we show that results are similar if we examine earnings at higher ages. To further assess whether sample attrition affects our main 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, more likely to be employed, receives slightly lower wages and is less likely to live in London. However, none of these differences are statistically significant. Summary statistics for both samples are reported in Appendix A.

[11]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.

partner. Due to observed gender differences in schooling and labor market outcomes, we allow all econometric model parameters 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 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.[12] 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.[13] 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 also 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

[12] 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. However, externalizing and internalizing have been measured more recently in other data sets and do not rely on such inventories, but are measured directly with fewer questions about behavior. As we show in Section 4.4, our main results extend to these other data sets, which alleviates the concern that our results could be driven by the particular inventories originally used in the NCDS.

[13] While most individuals score near zeros on most maladjustment syndromes, the median student has a score of 4 summing across all syndromes and few students (15.5%) score zero on all of them. This means that our results are not driven by a small percentage of very poorly behaved students.

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 the distributions of 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 skills, we follow Ghodsian (1977), who proposed dividing up the BSAG syndromes into two groups based on apparent differences among what behaviors the syndromes capture.[14] 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" and "restlessness" among others, and expresses anxious, aggressive, and outwardly-expressed behavior. The second variable, *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.[15] In addition, for use in our preliminary analyses, 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 a 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

In Table 1, we also report the summary statistics for the control variables we include in our subsequent analyses. 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

[14]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 our analysis sample (results available upon request).

[15]We report the assignment of the syndromes to the measures of socio-emotional skills in Appendix A (Table A1). 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).

member of the household self-reported having financial difficulties in the 12 months prior to being observed in either 1969 or 1974, and zero otherwise. 16% of our sample experienced financial difficulty in their childhood. The second control variable is an indicator variable for living in London either during childhood before age 16 or when labor market outcomes are measured at age 33. Including a London dummy is common practice using the NCDS given possible London-specific differences in schooling or labor outcomes.[16] 36% of the sample lives in or around London before age 16 versus 30% at age 33. Other control variables include a set of family background variables: 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 aged 11.

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 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 describe in Section 5.3. 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.

For the preliminary analysis, we use years of schooling as the education outcome. Formally, defining s_i as years of schooling for individual i, we estimate regressions of

[16] 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.

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 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. Z_i is a vector of control variables, including dummies for experiencing childhood financial difficulty and living in London before age 16. One concern is that, while externalizing behavior could capture a productive skill on the labor market, it 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 the set of family background variables introduced in Section 2.2.3 in the schooling equation but exclude them from the earnings equations (Heckman, Humphries, and Veramendi, 2018). Finally, e_i^S is a normally distributed disturbance.

Estimates of equation (1) are reported in Table 3. We start by regressing years of schooling on the crude measure of generic misbehavior and cognition, pooling males and females, and estimate a negative relationship in Column [1]. When we disaggregate misbehavior into the externalizing and internalizing behaviors to be included in the regression, we find both externalizing and internalizing behaviors independently lower education attainment in Column [2]. Repeating the same analysis for the male and female samples separately, we find broadly similar patterns of negative impacts from the two socio-emotional skills and a positive impact from cognition (Columns [3] and [4]). The controls, which we omit from the table for brevity, all have expected signs. For example, higher parental education has a positive impact on the respondent's own education.

To explain earnings, we regress log weekly earnings at age 33, conditional on being employed, onto measures of socio-emotional and cognitive skills (Table 3). 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 if using the pooled sample, experiencing childhood financial difficulty and living in London at age 33) and may or may not include schooling outcomes.

Column [5] contains estimates using the single measure of misbehavior, controlling for cognition and the basic set of controls for the pooled sample. In line with previous research (see, 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 (Column [6]). The single measure of misbehavior masks countervailing effects on earnings of the two socio-emotional skills. While internalizing behavior continues to be associated with an earnings penalty, the positive price of externalizing behavior emerges as soon as we control for internalizing behavior and cognition. In Column [7], we further control for schooling. When we do so, the positive coefficient on externalizing behavior becomes even larger. This change in the size of the coefficient reflects how externalizing behavior has a direct positive association with earnings along with an indirect negative association with earnings through less schooling. When we do not control for schooling, we measure the net of these two, but when we control for schooling, we measure the direct association, which is thus larger. 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 but that higher levels of externalizing behavior have a negative net effect on labor market outcomes through schooling that overwhelms the direct positive effect on earnings. In Columns [8] and [9], we repeat the exercise for males and females separately and conclude that the earnings premium of externalizing behavior exists for both genders in our sample.

The results from the preliminary analysis presented in Table 3 provide initial evidence that a socio-emotional skill that is productive on the labor market is not productive in school. The positive association between externalizing behavior during childhood and earnings in adulthood has generally not been recognized in previous literature on the economic consequences of childhood misbehavior. There are several reasons for this lack of recognition. First, 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. One of such studies is Fronstin, Greenberg, and Robins (2005) which 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 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 and highlight the importance of recognizing that misbehavior reflects distinct socio-emotional skills with potentially different returns in the labor market.

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 Main Econometric 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 and 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 a formal latent factor model that essentially relaxes some of these restrictions. The model treats observed maladjustment syndromes and test scores as measures with error of underlying skills. The formal model allows each measure, including test scores, to provide information about more than one factor and 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. Using this framework, we are able to secure identification of the impact of underlying skills imposing relatively few assumptions. We present the benchmark econometric model in Section 3.1 and explain how we identify and estimate model parameters in Sections 3.2 and 3.3.

3.1 Model

3.1.1 Measurement System for Unobserved Skills

Let the vector of unobserved non-cognitive skills and cognition be denoted $\mathbf{f}$, which can nevertheless be proxied by a set of observable measurements such as the ten BSAG maladjustment syndromes and the four aptitude test scores measured at the age of 11. 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.

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

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. Given many zeros on BSAG maladjustment scores, we use the logarithm of each BSAG score plus one as the relevant measurement in the measurement system. The error terms capturing mismeasurements, ε's, are assumed to follow a Poisson distribution for the BSAG syndromes and a normal distribution for the test scores. The latent skills follow a joint normal distribution, with mean μ and variance-covariance matrix $\mathbf{\Sigma}$:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \sim N(\mu, \mathbf{\Sigma}) = N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \right). \tag{4}$$

As suggested by Williams (2020), to identify the measurement system, we assume three "dedicated measures," one for each skill. That is, for each skill we choose one measure that is only affected by that skill. 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), in which case the relevant measurement equations are reduced to

$$\begin{aligned} M_1 &= m_1 + 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3 + W\delta_1 + \varepsilon_1 \\ M_2 &= m_2 + 0 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3 + W\delta_2 + \varepsilon_2 \\ M_3 &= m_3 + 0 \cdot f_1 + 0 \cdot f_2 + 1 \cdot f_3 + W\delta_3 + \varepsilon_3 \end{aligned} \tag{5}$$

For the remaining 11 measurements, we allow all three skills to load on them.

The choice of dedicated measures is a matter of judgement and 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). Finally, as factors do not have a natural scale, we normalize the coefficients of the dedicated measurements to unity as is commonly done in this literature. In Section 4.4, where we discuss robustness and sensitivity, we show that our main results are robust to reasonable alternative assumptions related to the measurement system.

3.1.2 Parameterizations of the Schooling Decision Rule and Labor Market Outcomes

We approximate the schooling decision with a linear-in-parameters model of years of schooling:

$$s = \underbrace{Z}_{\substack{\text{observed by}\\ \text{econometrician}}} \cdot \beta_S + \underbrace{\eta_S}_{\substack{\text{unobserved by}\\ \text{econometrician}}}, \tag{6}$$

where Z is a vector of variables observed by the econometrician that affect the schooling decision. It includes two basic controls, a dummy for the child experiencing financial difficulty at home and a London dummy, along with variables describing family backgrounds such as parents' education and occupation status, as in the preliminary analysis. β_S is a vector of parameters mapping variables in Z to schooling attainment, and η_S describes shocks to education attainment that are unobserved by the econometrician. We impose separability between the observed and unobserved variables in the representation of the schooling decision rule.

We summarize the labor market outcome by earnings at age 33 for individuals who are employed at that time.[17] More specifically, the log earnings at age 33, y, is represented by the following equation:

$$y = \underbrace{X}_{\substack{\text{observed by}\\ \text{econometrician}}} \cdot \beta_Y + \gamma_{s,Y} \cdot \underbrace{s}_{\substack{\text{observed by}\\ \text{econometrician}}} + \underbrace{U_Y}_{\substack{\text{unobserved by}\\ \text{econometrician}}} \tag{7}$$

X is the set of basic controls and the β_Y is the vector of associated coefficients. s is years of schooling with the associated coefficient $\gamma_{s,Y}$. U_Y is unobserved determinants of earnings.

We assume that all dependences across the unobserved components in the schooling and earnings outcome equations, η_S and U_Y, are generated by the vector of skills, $\mathbf{f}$, unobserved by the econometrician. More specifically, suppose

$$\eta_S = \mathbf{f}'\alpha_{\mathbf{S}} + \nu_S, \tag{8}$$
$$U_Y = \mathbf{f}'\alpha_{\mathbf{Y}} + \omega_Y, \tag{9}$$

where the α's 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 is a normal idiosyncratic error term for the earnings outcome.

[17] To avoid biases due to non-random attrition, we focus on age 33 earnings. In descriptive analyses (results of which are available upon request), we find similar returns to externalizing behavior among individuals when they are of age 42 and 50, suggesting that the externalizing premium extends over the lifecycle.

3.2 Identifying Assumptions

The key identifying assumption is that conditional on $\mathbf{f}$, Z, and X, choices and outcomes are statistically independent. Formally, we assume that,

$$\nu_S \perp\!\!\!\perp \omega_Y. \tag{10}$$

In addition, we array the measurement errors, ε_k, $k \in \{1, ..., K\}$ into a vector $\varepsilon = (\varepsilon_1, ..., \varepsilon_K)$ and assume that,

$$\varepsilon_k \perp\!\!\!\perp \varepsilon_{k'}, \forall k \neq k', \tag{11}$$

$$(\nu_S, \omega_Y) \perp\!\!\!\perp \varepsilon. \tag{12}$$

Assumptions (11) and (12) maintain that the measurement errors are independent from each other, and independent from the shocks. Last, we assume that,

$$(\nu_S, \omega_Y, \varepsilon) \perp\!\!\!\perp (\mathbf{f}, Z, X), \tag{13}$$

$$\mathbf{f} \perp\!\!\!\perp (X, Z). \tag{14}$$

Assumption (13) assumes independence of all the shocks and measurement errors with respect to factors and observables, and Assumption (14) assumes independence of factors with respect to observables.[18]

3.3 Estimation

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

$$\Phi = (\beta, \gamma, \alpha, \Xi) \tag{15}$$

where β denotes the set of coefficients on the vectors of observables absent the schooling attainment equation (6)-(7), $\gamma_{s,Y}$ is the coefficient governing the returns to schooling, α is the set of coefficients governing the returns to unobserved skills and Ξ are parameters of the measurement system described in equations (3) and (4).

We estimate the model by simulated maximum likelihood in two stages and allow all parameters to differ by gender. In the first stage, we estimate the measurement system for unobserved skills. For each suggestion of 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 $T = 500$ times and, for each draw of the factors, compute the probability of observing each measurement. 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

[18] Williams (2020) discusses these assumptions in more detail. In particular,Williams (2020) describes conditions under which Assumption (14) can be relaxed.

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). The simulated log likelihood function is computed as the sum of the log of each individual's average likelihood contribution taken over the T draws:

$$\mathcal{L}_1^{(g_1)} = \sum_{i=1}^{I} \log \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{f}_{it}^{\mathbf{M},(g_1)} (M_i) \right) \tag{16}$$

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, given the parameter estimates $\widehat{\Xi}$ found in the first step, we estimate the remaining structural parameters, (β, γ, α). 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_i, $\left(\mathbf{f}_{it}^{\mathbf{S},(g_2)}(s_i)\right)$, and the probability of observing earnings y_i, $\left(\mathbf{f}_{it}^{\mathbf{Y},(g_2)}(y_i)\right)$, for individual i, draw t and parameter suggestion (g_2). The simulated log likelihood in the second stage is given by:

$$\mathcal{L}_2^{(g_2)} = \sum_{i=1}^{I} \log \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{f}_{it}^{\mathbf{M},(\widehat{\Xi})}(M_i) \times \mathbf{f}_{it}^{\mathbf{S},(g_2)}(s_i) \times \mathbf{f}_{it}^{\mathbf{Y},(g_2)}(y_i)^{\mathbf{1}(e_i=1)} \right)$$

where e_i is the observed employment status (with employed taking the value 1) in the data.[19]

4 Main Results on Schooling and Earnings

Our benchmark model provides a framework to characterize the unobserved heterogeneity as three latent factors, and under the assumptions laid out in the previous section, we can interpret the effects of the socio-emotional and cognitive skills on outcomes as causal. In this section, we report the estimation of key coefficients of the measurement system, the education equation, and the earnings equation in Sections 4.1, 4.2 and 4.3. We report the full estimation results in Appendix B. We show in a number of robustness and sensitivity analyses, which we report in Section 4.4, that our main findings are robust to including additional unobserved factors or possible confounding variables. These results make us confident in assigning a causal

[19] 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.

interpretation to our main findings. That said, it does not mean there is a clear policy implication to alter externalizing behavior as it is unclear how to implement policy to target one specific socio-emotional skill without affecting others. Lastly, in Section 4.5, we investigate whether the earnings premium is driven by hours versus wages. We document that high-externalizing females work more hours, while high-externalizing men receive higher wages. These results are a reflection of the structure of the relevant labor market experience by the individuals in our sample. We interpret these results as possible evidence that externalizing behavior may capture a kind of resoluteness or doggedness, helping individuals figure out alternative ways to achieve their goals (e.g. higher earnings) under the constraints, norms, and circumstances they face. While helpful in the labor market, similar tendencies could easily be perceived as disobedience or a challenge to authority in the classroom.

4.1 Mapping Unobserved Skills to Observed Misbehaviors

We start by reporting the joint distribution of unobserved skills. We find a positive correlation between externalizing and internalizing behavior of 0.486 for males and 0.323 for females, along with negative correlations between the two socio-emotional skills and cognition. These correlations could reflect the distribution of skill endowments at birth or early childhood investments if the same environments that promote externalizing and internalizing behaviors also slow cognitive development (Heckman and Cunha, 2007).[20] 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 overestimate the degree to which each socio-emotional skill negatively affects schooling.

In Table 4, we report estimates of factor loadings mapping latent skills to BSAG maladjustment syndromes and aptitude test scores. Estimates are reported for the pooled sample as well as 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, and inconsequential 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.

[20]An example would be childhood poverty, which we investigate in Section 6.1. The positive correlation 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).

4.2 Schooling

The estimation results of the schooling equation are reported in Table 5. There is a significant negative relationship between externalizing behavior and educational attainment for boys in the sense that externalizing significantly reduces education attainment as measured by years of schooling. The relationship between externalizing and schooling for females is much weaker. In other words, high-externalizing females are better able to finish school in comparison to high-externalizing males. This finding is generally consistent with earlier literature showing that the negative impact of externalizing behavior on schooling is more salient for boys than for girls (Bertrand and Pan, 2013). Indeed, teachers are more likely to punish boys versus girls 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).

The sizes of the effects of socio-emotional skills in the schooling model are much smaller than those of cognition, which predicts schooling at similar magnitudes across genders, but are of comparable magnitudes as the effects of family characteristics, which all have the expected signs. 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.

Our estimates for the schooling model are broadly consistent with the interpretation that externalizing children face higher costs, such as a higher effort to conform to expected in-classroom behavior, that make it difficult for them to obtain higher levels of schooling. There is a basis for this interpretation in earlier literature. 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.

4.3 Earnings

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 the earnings equation conditional on employment are reported in Table 5. Columns [4], [6] and [8] report results without conditioning on schooling, while Columns [5], [7] and [9] report results from our benchmark conditioning on schooling. For males, a one-standard-deviation increase in externalizing behavior predicts a statistically significant 9.8% increase in earnings. For females, a one-standard-deviation increase in externalizing behavior predicts a statistically significant 6.4% increase in earnings.

One explanation for this finding relates to our interpretation of the negative impact of externalizing on schooling. To compensate for difficulties attaining high levels of education, such as high effort costs, high-externalizing students could be positively selected on other dimensions that lead to higher earnings. This dynamic would pertain to males, who face the schooling penalty. A leading contender would be positive selection in the form of higher cognition, which leads to higher earnings. Since our model explicitly controls for cognition, this does not drive our results. As part of our sensitivity analyses (summarized in Section 4.4.1), we allow for an additional (fourth) unobserved factor, which would capture additional sources of positive selection. Results remain largely unchanged.

We also investigate whether the "net effect" of externalizing behavior is lucrative. Our benchmark model includes educational attainment in the labor market outcome equations, which captures the negative impact of externalizing on earnings through lower educational attainment. If we omit education, the estimated coefficient mapping externalizing to earnings includes both the direct effect of externalizing on earnings and the indirect effect working through schooling, the net of which could be negative. Net effects are reported in Columns [4], [6] and [8]. Consistent with the earlier result that externalizing reduces schooling for boys but not girls, omitting education reduces the point estimate of the effect of externalizing for boys but not girls. Notably, the coefficient is positive whether or not we include schooling, suggesting that more externalizing workers make higher earnings despite the negative impact of externalizing on schooling. Consistent with the intuition of ability bias, omitting education increases the point estimates of the effect of cognition for both males and females.

Internalizing behavior is negatively related to earnings. For males, a one-standard-deviation increase in internalizing behavior predicts a very significant 14.9% decrease in earnings. For females, the counterpart coefficient is negative but insignificant. We also find that cognition significantly increases earnings (by 3.4% for males and as much as 11.7% for females). The remaining parameters follow conventional wisdom. Higher educational attainment increases earnings. Individuals living in or around London earn significantly more, while individuals who experienced financial difficulties in childhood earn less.

Our findings demonstrate a more nuanced relationship between childhood misbehavior 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 Robustness and Sensitivity of Main Results

The positive impact of externalizing behavior on earnings and its mixed effects on schooling versus earnings for males are novel findings that have not been explored in earlier literature. We constructed our main econometric model to account for obvious sources of selection. Nevertheless, securing our estimates requires a number of assumptions and pertains to a particular cohort in one time and place. Our next step is to assess whether findings are robust to alternative modeling assumptions, when accounting for selection into employment or when controlling for personality traits typically measured in adulthood. In this section, we summarize our findings from the various sensitivity and robustness analyses we painstakingly conducted that are documented in more detail in our online appendix.

4.4.1 Alternative Modelling Assumptions

In the benchmark model, we assume there are three unobservable skills, externalizing behavior, internalizing behavior and cognition, which are identified from measures of childhood classroom misbehavior and test scores. Identification of measurement system parameters requires that, for each skill, we designate one particular measurement (the dedicated measurement) that is not a measurement of the other two skills. Which measurement to dedicate to each skill is ultimately a choice. A benefit of our approach is that we can re-estimate the model iterating over all possible candidates for the dedicated measurements of the two socio-emotional skills to assess sensitivity. We report the full analysis that spans all possible combinations of the dedicated measurements for externalizing and internalizing behaviors in Appendix C.1 and summarize our findings here.

While different dedicated measurement choices imply different magnitudes of the effects on earnings, in the majority of cases externalizing behavior has a significantly positive earnings premium for both genders. In no specification do we find significant evidence against our main result. In particular, the earnings premium of the externalizing factor loses significance precisely when the factor starts to be contaminated by reflecting depressive syndromes in alternative specifications. This happens when we choose withdrawal or unforthcomingness to be the dedicated measure for internalizing behavior, in which case depression loads heavily on the "externalizing" factor. In contrast, we can construct an externalizing factor that does not map to depression, loads heavily onto outwardly expressed aggressive behaviors, which is what we do in the benchmark model, and that externalizing factor has a positive impact on earnings. Broadly, this analysis illustrates the fundamental identification problem in measuring underlying traits as discussed in Almlund et al. (2011). The researcher faces a trade-off between letting the data guide the analysis versus imposing just enough structure to identify economically meaningful objects.

Another concern related to our measurement system is whether results are robust to relaxing the assumption of only three factors. In Appendix C.2, we re-estimate the model under the assumption that there are four unobserved factors underlying childhood classroom misbehaviors and allow this fourth factor, in addition to externalizing behavior, internalizing behavior, and cognition, to influence the schooling, wages and hours worked equations.[21] The full results are reported in Appendix C.2, and we summarize our findings here. The fourth factor is significantly positively correlated with schooling, but adding the factor does not affect the impact of the

[21] We decompose weekly earnings into weekly wages and hours worked since we needed three outcome functions to identify the fourth factor.

externalizing factor on schooling. If anything, the negative effect of externalizing behaviors for females becomes larger, but remains statistically insignificant. The fourth factor is significantly negatively correlated with hours, but its effect on wage is positive for males and negative for females. However, including the fourth factor again does not change the returns to externalizing behaviors in any material way. We still find that externalizing behaviors are positively related to hourly wages and hours worked for both males and females when we allow for the fourth factor in our model. These results reassure us that our main empirical findings are remarkably robust against varying modeling or identifying assumptions.[22]

4.4.2 Selection into Employment

In the benchmark model, as we identify unobserved cognition as a latent factor, selection based on cognition is therefore controlled for. But there might be other types of selection that could explain our finding of an earnings premium for externalizing behavior. In this section, we examine the possibility that the earnings premium is driven by selection into employment.

Recall that the earnings regression is 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. 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.

We start by estimating 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.[23] We report the full estimation results from this section in Appendix D. 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.[24] For males, externalizing behavior is weakly negatively related to unemployment. Moreover, men and women with high levels of internalizing behavior are significantly more likely to be unemployed. Cognition is not statistically associated with the employment decision for males and females. The main impact of cognition on employment

[22] We conducted additional robustness checks related to our measurement system. In Appendix C.3, we present estimates from an alternative specification where we estimate the measurement system jointly with outcomes. Results remain largely unchanged. In results available upon request, we also allow for correlation in the error terms for anxiety towards children and anxiety towards adults, and for hostility towards children and hostility towards adults. In both cases and for both genders, the estimated correlation is zero.

[23] 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.

[24] 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.

likely works through schooling, which we control for 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).[25] 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 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. For both genders, including earnings for the unemployed does change the estimated impact of externalizing on earnings or its significance. 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.

4.4.3 Stress Hormones and the Big 5

In this section, we consider the possibility that externalizing behavior might be correlated with other psychological or physiological factors. The full analyses are in Appendix E.

The Big Five One possible interpretation of our findings is that externalizing behavior might be correlated with other well-known constructs of non-cognitive skills, which are known to have labor market impacts, for example the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness and neuroticism). Several studies have examined the relationship between externalizing and internalizing behaviors and 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).[26] Could it be the ex-

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

[26]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

ternalizing individuals are simply less agreeable people? We test for this possibility by adding the "Big 5" traits to the descriptive earnings equation in Section 2.3 as controls. Controlling for the "Big 5" traits reduces the effect of externalizing behavior on earnings by about 30% (from 0.026 to 0.018) and increases the negative effect on education by about 15% (from -0.091 to -0.107). 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.[27]

The Stress Hormones Similarly, our main findings could be explained by externalizing individuals having higher levels of stress hormones (cortisol), which recent research has shown to relate to risk-taking behaviors that could potentially make schooling difficult but be productive at work (Shirtcliff et al., 2005). Again, we do not find evidence that this is the case. While externalizing is positively correlated with the salivary cortisol measure collected at age 44 in NCDS, controlling for cortisol does not significantly change the relationship between externalizing behavior, schooling and earnings in our sample. In general, our results are not explained by correlations between externalizing behavior and well-known constructs that have been examined in earlier literature.

4.5 Wage versus Hours

So far, we have focused on annual earnings as a key labor market outcome, but externalizing behavior can induce higher earnings by increasing wage or hours or both. If externalizing behavior's earnings premium is mostly achieved by lengthening hours worked, then the gain from acquiring externalizing (if possible) will naturally be diminished for those who already work full time.

After disaggregating earnings to wages and hours separately, we find that externalizing behavior impacts wages and hours differently for males and females. We estimate the benchmark econometric model replacing the earnings equation by the analogous hourly wage and weekly hours worked equations, conditional on employment. The results are in Table 6.

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.

[27]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), which we describe in more detail in Section 5.3. 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.

For males, a one-standard-deviation increase in externalizing behavior predicts a statistically significant 5.7% increase in hourly wages, while for females, the effect on hourly wages is halved at 3% and only marginally significant. In contrast, a one-standard-deviation increase in externalizing behavior predicts a 2.5% increase in hours worked for males and a 5.4% increase in hours worked for females, both statistically significant. In other words, externalizing increases earnings for males mainly through its positive effect on hourly wages while for females it works mainly through longer hours.

Given these different paths to higher earnings, it may be tempting to infer that externalizing behavior raises productivity for male workers but not for female workers, while it reduces the disutility of labor more for female than for male workers. However, the evidence for this is not conclusive. Returning to Table 1, there is some evidence that the kinds of job options available to men versus women in our sample differed. The average number of hours worked is 43.5 for men with a standard deviation of 7.8, while it is 27.9 for women with a standard deviation of 12.1. This suggests that there is relatively limited scope to change hours for externalizing men compared to women. In comparison, wages exhibit a lower mean and a lower variance, suggesting limited score for women to increase their earnings through higher wages. This leads us to conclude, though speculatively, that high-externalizing individuals are proactive, energetic or resourceful and thus figure out a way to get what they want along the dimensions that prevailing constraints, norms, and circumstances permit. This tendency is distinct from grit, which captures a single-minded pursuit of a long-term goal, or the Big Five personality traits, as we show in the last section.

We also note that internalizing behavior is negatively related to both wages and hours worked for both genders. For males, a one-standard-deviation increase in internalizing behavior predicts a significant 10.0% decrease in hourly wage, and for females, it implies a significant 4.4% reduction. In terms of hours, a one-standard deviation increase in internalizing behavior reduces males' as well as females' hours by 3.1%, though the effect for males is significant and for females is not. In contrast to externalizing behavior, internalizing behavior reduces earnings mainly through reducing wage for both genders. Finally, we find that cognition significantly increases hourly wages (by 5.4% for males and 7.1% for females), but marginally reduces hours by 1.4% for males while significantly increases hours by 4.8% for females.

5 Generalizability of Findings

In this section, we present results from several analyses that investigate the generalizability of our findings. We consider three different dimensions. First, in Section 5.1, we ask whether the externalizing premium is dependent on having a certain type of family structure. While, perhaps surprisingly given the positive effects on earnings, externalizing behavior predicts higher rates of marriage and fertility for women, we show that the positive relationship between externalizing and earnings does not decrease after adjusting for marriage and fertility. Second, in Section 5.2, we ask whether the externalizing premium is limited to certain occupations or occupational tasks. We find that the premium is evident across virtually all occupations and tasks. This finding is noteworthy as it contrasts with earlier work that finds some positive benefits to potentially problematic behavior in small slivers of the labor market (e.g., rule-breaking and successful entrepreneurship by Levine and Rubinstein (2017)).[28] Finally, in Section 5.3, we ask if the externalizing premium is limited to the Great Britain or to cohorts born in the 1950s. It is entirely possible that labor market structures or teacher training in different times and places would change. In particular, we replicate our main analysis in four other studies and find that the externalizing premium replicates across these data sets.

5.1 Marriage and Fertility

Differences in effects of externalizing behavior on hourly wages and weekly hours worked for males and females suggest that externalizing might work through different channels across genders. To further probe our findings, we examine two additional lifecycle outcomes that are especially relevant for the age group under study: marriage and fertility. For example, it is possible that high-externalizing individuals are less likely to be in committed relationships or to have children, which could free up time and energy to boost their earnings.

We assess how estimated coefficients change when we add endogenous intermediate outcome variables to the earnings equation, while keeping the measurement system mapping latent skills to observed measurements of misbehavior as in the benchmark model. The results are in Table 7. We find that while having a partner has a positive effect on earnings for males, having children has a strong negative effect on earnings 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 and find that externalizing males and females are both more likely to have a larger number of children by age 33. However, since fertility only lowers earnings for females, this channel operates to counteract the direct positive effects of externalizing in the labor market for females but not males. It is noteworthy that these patterns remain after we control for further outcomes such as months of experience and occupations.

Controlling for partnership and fertility lowers the coefficient of externalizing on earnings slightly from 0.098 in the benchmark to 0.079 for males, but it increases the coefficient significantly for females from 0.064 in the benchmark to 0.101. Married men with children appear to enjoy an earnings premium in the labor market, and therefore these intermediate outcomes partly explain externalizing's earnings premium we found earlier for men. However, the story is different for women. Despite the fact that more externalizing women tend to have more children which reduces

[28] The authors show that individuals who engage in illicit behaviors as teenagers report high earnings in incorporated self-employment. This means that the gain from being smart and illicit accrues to a small fraction of the labor force, 3.4% of their CPS sample and 1.5% of their NLSY79 sample.

their earnings, the net effect of externalizing is still positive. In other words, controlling for these intermediate outcomes only reveals an even higher earnings premium for externalizing women.

To visualize these results, we plot earnings against different levels of externalizing separately for men and women in Figure 2. The slope of the curve represents the impact of externalizing behavior on earnings. To generate the figure, we simulate weekly earnings 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 while conditioning on the partnership and fertility intermediate outcome variables. For males, conditioning on these outcomes does not change the slope very much (Panel (a)). For females, the slope increases noticeably when we condition on partnership and fertility (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 highly externalizing women who do not have children.

5.2 Returns Across Occupational Tasks

Some dimensions of human capital are more productive for the completion of certain tasks, as different tasks in life require different skills in different degrees (see, e.g., Roy, 1951; Mandelbrot, 1962; Willis and Rosen, 1979; Heckman and Sedlacek, 1985; Heckman, Stixrud, and Urzúa, 2006). Building on this idea, a reasonable hypothesis is that externalizing behavior is productive for a limited set of tasks and is thus lucrative in a subset of possible occupations. This would have policy implications. For example, if results are driven by a very small number of tasks, it might be that low-externalizing people could be trained on just these tasks at relatively low cost (or could avoid occupations that require such tasks) and that the general thrust of policies aiming to reduce externalizing behavior are not necessarily the wrong approach.

We extend our labor market model to allow for the returns to skills to vary with occupational tasks using the O*NET[29] task-intensity scales as in Acemoglu and Autor (2011). We focus on two well-studied measures: the abstract/social task intensity and the routine/manual task intensity. The task intensities are composite measures of O*NET Work Activities and Work Context Importance scales.[30] The

[29] The O*NET is an American classification system, and the NCDS collected detailed information on individual occupations in the ISCO-88 classification system. We rely on the methodology in Hardy, Keister, and Lewandowski (2018) to link the NCDS individuals' occupations to the O*NET classification.

[30] The abstract/social task measure is a normalized composite scale of six O*NET subscales: "analyzing data/information," "thinking creatively," "interpreting information for others," "establishing and maintaining personal relationships," and "guiding, directing, and motivating subordinates and coaching and developing others." The routine/manual task measure is a normalized composite scale of six O*NET subscales: "importance of repeating the same tasks," "importance of being exact or accurate," "structured versus unstructured work," "controlling machines and processes," "keeping

two composite scales were constructed using factor analysis and are standardized to have mean zero and variance equal to one. The results are found in Appendix F.

While we find some heterogeneity in the returns to externalizing behavior across tasks, the externalizing behavior labor market premium is predominantly positive. Mainly, for males, we find that the returns to externalizing behavior are smaller in occupations that are intensive in abstract and social tasks and larger in occupations that are intensive in manual and routine tasks. This heterogeneity is, however, very small. Only for jobs with routine tasks below the 2.5th percentile (2 standard deviations below the mean) or with abstract tasks in the 97.5th percentile (2 standard deviations above the mean) or both (as measured in the NCDS) would we expect to find an overall negative return to externalizing in the labor market. Roughly 5% of individuals in our sample have an occupation meeting one of these criteria. Individuals in our sample in occupations requiring sufficiently high levels of abstract tasks to meet this threshold include senior government officials and managers of personnel departments. Those requiring sufficiently low levels of manual tasks include fashion models.

This points to the generality of our main message to the vast majority of working adults. This is in contrast to the finding in Levine and Rubinstein (2017), who show that the combination of being smart and illicit during youth appears to be productive in a very small sector of the economy, incorporated entrepreneurship. One could argue that there is little reason to foster or accommodate individuals who engaged in illicit behavior as doing so might bring costs and benefits are limited to a small sliver of the labor market. Our findings on externalizing behavior suggest a much different trade-off since the benefits are widespread.

5.3 Replication in Other Data Sets

Another possible concern is that our findings are specific to Great Britain in the 1950s. We thus explore other data sets in more contemporary settings and in different social contexts. We replicate our main analysis in 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. data sets. These are the major longitudinal studies that follow individuals over the lifecycle with measurements of both behavior in school during childhood and labor market outcomes in adulthood for the same individuals. Detailed descriptions of the data sets and the full empirical results are found in Appendix G.

In each dataset, we construct crude measures of skills and link these to schooling and earnings. We find that, in all data sets, externalizing behavior is associated with fewer years of schooling. This negative effect is strongly significant, with the exception of in the PSID where the negative coefficient is significant at the 10% level. Compared to the NCDS sample, the point estimates of the correlation between ex-

a pace set by machinery or equipment," and "time spent making repetitive motions."

ternalizing 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. We also show that externalizing behavior is significantly associated with higher earnings in the two British data sets, the 1958 and the 1970 cohort, and two U.S. data sets, NELS and PSID.[31] The point estimate of the impact of externalizing on earnings from the NCDS lies between estimates obtained from other data sets. These results suggest that the documented externalizing earnings premium does not vary systematically across countries or over time.

6 Context Dependence

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 or by the presence of certain life events such as criminality or police involvement.

6.1 Childhood Disadvantage

We construct a subsample of our analytic sample consisting of subjects who faced financial difficulties during childhood to resemble the family characteristics of the sample studied in Heckman, Pinto, and Savelyev (2013), which we refer to as the "Low SES" subsample that accounts for about 16% of our analytic sample.[32] We estimate the model including the measurement system for the "Low SES" subsample and for all other subjects in our analytic sample, which we call the "High SES" subsample, separately.[33] The results on schooling and earnings are in Table 8.

[31]The CNLSY is the only dataset where we do not find a significant relationship between externalizing behavior and earnings. This can be due to two reasons. First, the CNLSY is the only dataset where we rely on parents' report of children's behaviors 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). Second, the CNLSY sample with observed earnings is a selected sample born from young mothers. 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 6.1).

[32]An individual is coded as experiencing financial difficulties during childhood 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. Summary statistics for the sub-samples are found in Appendix A.2.

[33]Note that we estimate the model including 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., Urzúa (2008) allows the distribution and impact of underlying skills to vary by race.

Estimates from this model show patterns that are similar to the main results. However, we find suggestive evidence of differences by childhood SES. First, the point estimate for the schooling penalty appear to be more negative for the low-SES individuals relative to the benchmark result, though due to the small sample size it is very imprecisely estimated. This is however broadly consistent with results in Ramey (2018), who shows that high-externalizing Black students in the U.S. face a higher likelihood of punishment by suspension in comparison to similarly externalizing whites. This could arise because schools that serve low-SES children have fewer resources to address externalizing behavior and therefore react to it through suspensions or expulsions.

Perhaps more importantly, we find suggestive evidence that the labor market returns to externalizing behavior vary by the childhood socio-economic environment. While for the high SES sample we find that a 1 standard deviation increase in externalizing behavior is related to a 14.5% increase in earnings, we show that this positive effect completely disappears for the low-SES sample. If aggressive behavior is deemed unacceptable in jobs limited to advantaged people with social connections, we might expect the externalizing premium to be larger for people from less advantaged backgrounds. We find that returns are, if anything, larger for more advantaged groups. While individuals that grew up with financial difficulties faces an insignificant if not negative earnings premium of externalizing behavior, individuals from more advantageous backgrounds see a significant 14.5% earnings premium from externalizing behavior. This analysis helps to reconcile results with those in Heckman, Pinto, and Savelyev (2013).

Reconciling these results by cutting our sample also highlights concerns related to sample selection, external validity, and thus generalizability when measuring returns to skills in specific groups (Lundberg, 2013). If we had limited our sample to disadvantaged children, we might have erroneously concluded that externalizing has no impact at all, even though it has positive returns for most of the more representative population we study. This type of issue is common. For example, positive measured returns of another skill, grit, influenced pedagogy for years, but may have been a product of the highly selected samples used to measure these returns. Subsequent work has shown a much more nuanced set of returns that are context dependent (i.e., there are contexts in which grit might not be good).

More broadly, differences in returns to externalizing behavior across childhood environments leave us with at least two distinct, but related, possibilities. The first is a labor supply story: there are true differences in the productivity of externalizing behavior across groups. For example, children born into wealthier families may have resources that teach them to channel their hyperactivity into productive activities. Alternatively, a labor demand story would be that the skill may be equally productive across groups, but perceived and thus rewarded differently on the labor market. 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. In either case, our results using a broadly representative sample allows us to provide evidence suggesting the troubling 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.[34]

6.2 Criminal Activity and Police Involvement

We examine next whether, consistent with the story in Heckman, Pinto, and Savelyev (2013), criminality might explain the heterogeneous labor market returns to externalizing. To do that, we expand the benchmark econometric model by including a measure of police involvement at age 16 as an additional outcome equation and as an additional explanatory variable in the earnings equation.

We find that externalizing behavior does predict higher police involvement for both males and females (Table 9). A one-standard-deviation increase in externalizing behavior predicts a 6.6 percentage point increase in the likelihood of police involvement for males and a 2.1 percentage point increase for females. However, police involvement does not appear to derail labor market prospects among individuals in the British sample we study. For males, police involvement leads to an insignificant 3.9% reduction in earnings and for females an insignificant 12.4% increase. Controlling for this particular channel affects only slightly the magnitude of the coefficients of externalizing behavior in the earnings equations for both genders.

These results raise a further concerning possibility: that the returns to externalizing behavior might be negative in a context where police involvement is highly penalized in the labor market such as that studied in Heckman, Pinto, and Savelyev (2013), but not necessarily so in a different context such as in the British sample we study. The labor market price of externalizing behavior, as we show, can be quite different for individuals from different socio-economic backgrounds, and can also depend on the wider institutional context that affects how criminality is viewed in the labor market. These results again underscore the pitfalls of measuring returns to socio-emotional skills (or to any part of human capital) in highly selected samples as doing so can generate misleading estimates of returns that obfuscate complex variation or that simply do not generalize to other groups or contexts.

7 Conclusion

For many factors that comprise human capital, such as health or cognition, it is hard to imagine contexts where more is not better. We provide evidence that a widely studied socio-emotional skill measured during childhood, externalizing behavior, predicts lower education and higher earnings. We probe this main finding by showing it is robust across modeling assumptions and data sets. This result cannot

[34]Indeed, a recent and burgeoning literature shows that factors as fundamental as genetic endowments can have vastly different relationships to key life outcomes depending on contextual factors, such a parents' resources (see e.g., Papageorge and Thom (2020) and Ronda et al. (2021)).

be explained by selection into employment, occupation or family structure, and survives after controlling for other socio-emotional skills known to affect labor market outcomes.

Our findings calls into question earlier work showing positive labor market returns to policies that decrease externalizing behavior. However, that work uses a selected sample of disadvantaged children. We reconcile our findings with these results by selecting a similarly disadvantaged sample, which illustrates not only that skill prices can be context dependent, but also that prices measured using selected samples may not be generalizable.

This paper raises a broader concern that research on child development is incomplete if it focuses solely on schooling as the outcome. To the degree that schooling is an input to other outcomes, such as earnings, a skill that decreases education, but increases earnings, may have value. Put differently, policies that target skills with the aim of increasing education but in doing so lead to lower earnings, are potentially counterproductive. Ironically, such policies risk undermining a crucial purpose of education, which is to develop children's human capital to help them succeed adults. This risk is not theoretical as policies exist that aim to develop some socio-emotional skills thought to be helpful in a schooling environment or to curb tendencies deemed unhelpful at school. The rationale for such policies is the assumption that the same skills that are valuable in childhood are also beneficial in adulthood. This assumption is largely untested and—as our findings show—potentially erroneous. Future research may continue to interrogate this assumption when evaluating policies concerning schooling and the accumulation of human capital.

References

Acemoglu, Daron and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, vol. 4. Elsevier, 1043–1171.

Achenbach, Thomas M. 1978. "The Child Behavior Profile: I. Boys Aged 6-11." *Journal of Consulting and Clinical Psychology* 46:478–488.

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.

Alaoui, Larbi and Christian Fons-Rosen. 2021. "Know When to Fold' Em: The Flip Side of Grit." *European Economic Review* 136:103736.

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.

Bashant, Jennifer. 2014. "Developing Grit in Our Students: Why Grit is Such a Desirable Trait, and Practical Strategies for Teachers and Schools." *Journal for Leadership and Instruction* 13 (2):14–17.

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.

Campbell, Amy and Cynthia M. Anderson. 2011. "Check-In/Check-Out: A Systematic Evaluation and Component Analysis." *Journal of Applied Behavior Analysis* 44 (2):315–326.

Campbell, Susan B, Daniel S Shaw, and Miles Gilliom. 2000. "Early Externalizing Behavior Problems: Toddlers and Preschoolers at Risk for Later Maladjustment." *Development and Psychopathology* 12 (3):467–488.

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.

CPPRG. 1992. "A Developmental and Clinical Model for the Prevention of Conduct Disorders: The FAST Track Program." *Development and Psychopathology* 4:505–527.

Crone, Deanne A., Leanne S. Hawken, and Robert H. Horner. 2003. *Responding to Problem Behavior in Schools: The Behavior Education Program. Second Edition.* Guilford Press.

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. 2020. "The First 2,000 Days and Child Skills." *Journal of Political Economy* 128 (6):2067–2122.

Duckworth, Angela Lee and Patrick D Quinn. 2009. "Development and Validation of the Short Grit Scale (GRIT–S)." *Journal of Personality Assessment* 91 (2):166–174.

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.

Eisenberg, Nancy, Amanda Cumberland, Tracy L. Spinrad, Richard A. Fabes, Stephanie A. Shepard, Mark Reiser, Bridget C. Murphy, Sandra H. Losoya, and Ivanna K. Guthrie. 2001. "The Relations of Regulation and Emotionality to Children's Externalizing and Internalizing Problem Behavior." *Child Development* 72 (4):1112–1134.

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.

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

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M Chang, and Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344 (6187):998–1001.

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.

Grossman, Michael. 2000. "The Human Capital Model of the Demand for Health." In *Handbook of Health Economics*, vol. 1A. Elsevier, 347–408.

Hardy, Wojciech, Roma Keister, and Piotr Lewandowski. 2018. "Educational Upgrading, Structural Change and the Task Composition of Jobs in Europe." *Economics of Transition* 26 (2):201–231.

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.

Heckman, James J and Guilherme Sedlacek. 1985. "Heterogeneity, Aggregation, and Market Wage Functions: an Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy* 93 (6):1077–1125.

Heckman, James J, Jora Stixrud, and Sergio Urzúa. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *The Journal of Labor Economics* 24 (3):411–482.

Humphrey, Neil, Afroditi Kalambouka, Michael Wigelsworth, Ann Lendrum, Clare Lennie, and Peter Farrell. 2010. "New Beginnings: Evaluation of a Short Social–Emotional Intervention for Primary-Aged Children." *Educational Psychology* 30 (5):513–532.

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):390–407.

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.

Malti, Tina and Kenneth H. Rubin. 2018. *Handbook of Child and Adolescent Aggression*. Guilford Press.

Mandelbrot, Benoit. 1962. "Paretian Distributions and Income Maximization." *The Quarterly Journal of Economics* 76 (1):57–85.

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.

Nangle, Douglas W., Cynthia A. Erdley, and Rebecca A. Schwartz-Mette. 2020. *Social Skills Across the Life Span: Theory, Assessment, and Intervention*. Academic Press.

Närhi, Vesa, Tiina Kiiski, and Hannu Savolainen. 2017. "Reducing Disruptive Behaviours and Improving Classroom Behavioural Climate with Class-Wide Positive Behaviour Support in Middle Schools." *British Educational Research Journal* 43 (6):1186–1205.

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.

Neidell, Matthew and Jane Waldfogel. 2010. "Cognitive and Noncognitive Peer Effects in Early Education." *The Review of Economics and Statistics* 92 (3):562–576.

O'Connor, Kim M. and Ben Hayes. 2020. "How Effective are Targeted Interventions for Externalizing Behavior When Delivered in Primary Schools?" *International Journal of School and Educational Psychology* 8 (3):161–173.

Papageorge, Nicholas W and Kevin Thom. 2020. "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study." *Journal of the European Economic Association* 18 (3):1351–1399.

Ramey, David M. 2018. "The Social Construction of Child Social Control via Criminalization or Medicalization: Why Race Matters." *Sociological Forum* 33 (1):139–164.

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, Esben Agerbo, Dorthe Bleses, Preben Bo Mortensen, Anders Børglum, David M. Hougaard, Ole Mors, Merete Nordentoft, Thomas Werge, and Michael Rosholm. 2021. "Family Disadvantage, Gender and the Returns to Genetic Human Capital." *Scandinavian Journal of Economics* .

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.

Shirtcliff, Elizabeth A, Douglas A Granger, Alan Booth, and David Johnson. 2005. "Low Salivary Cortisol Levels and Externalizing Behavior Problems in Youth." *Development and Psychopathology* 17 (1):167–184.

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.

Stage, Scott A. and David R. Quiroz. 1997. "A Meta-Analysis of Interventions to Decrease Disruptive Classroom Behavior in Public Education Settings." *School Psychology Review* 26 (3):333–368.

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

Sørlie, Mari-Anne and Terje Ogden. 2015. "School-Wide Positive Behavior Support – Norway: Impacts on Problem Behavior and Classroom Climate." *International Journal of School and Educational Psychology* 3 (3):202–217.

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

Williams, Benjamin. 2020. "Identification of the Linear Factor Model." *Econometric Reviews* 39 (1):92–109.

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.

8 Tables and Figures

Table 1: Summary Statistics of Demographics and Outcomes

	All	Males	Females	Diff
Female	0.507			
	(0.500)			
Outcomes				
Years of Education	12.619	12.846	12.397	***
	(2.526)	(2.573)	(2.308)	
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)	
Has a Partner	0.873	0.877	0.868	
	(0.333)	(0.328)	(0.338)	
Number of Children	1.475	1.349	1.597	***
	(1.125)	(1.152)	(1.084)	
Controls				
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)	
Father Studied Beyond Min. Schooling Age	0.265	0.266	0.265	
	(0.442)	(0.442)	(0.441)	
Mother Studied Beyond Min. Schooling Age	0.215	0.217	0.213	
	(0.411)	(0.412)	(0.410)	
No Info on Father Figure	0.0254	0.0260	0.0248	
	(0.157)	(0.159)	(0.156)	
Father in Skilled Occupation	0.532	0.530	0.534	
	(0.499)	(0.499)	(0.499)	
Father in Managerial Occupation	0.244	0.246	0.242	
	(0.430)	(0.431)	(0.429)	
Working Mother	0.614	0.610	0.619	
	(0.487)	(0.488)	(0.486)	
Observations	7,241	3,573	3,668	7,241

Notes: This table lists the summary statistics of demographics, education, and labor market outcomes for the analytic sample of 7,241 individuals. For employment, entries are in the form of percentages divided by 100. Wages and weekly earnings are measured in 1992 British pounds. Statistics are reported separately for all individuals, for males and for females. In the last column, *, ** 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

	All	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	7,241	3,573	3,668	7,241

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. Entries in the table are averages for each syndrome. To construct crude measures of unobserved skills, we sum up all variables used to measure that skill according to Table A1 in Appendix A and then normalize each unobserved skill to have mean zero and standard deviation one. Statistics are reported separately for all individuals, for males, and for females. In the last column, *, ** and *** mean that differences between males and females are significant at the 10, 5 and 1 percent levels, respectively.

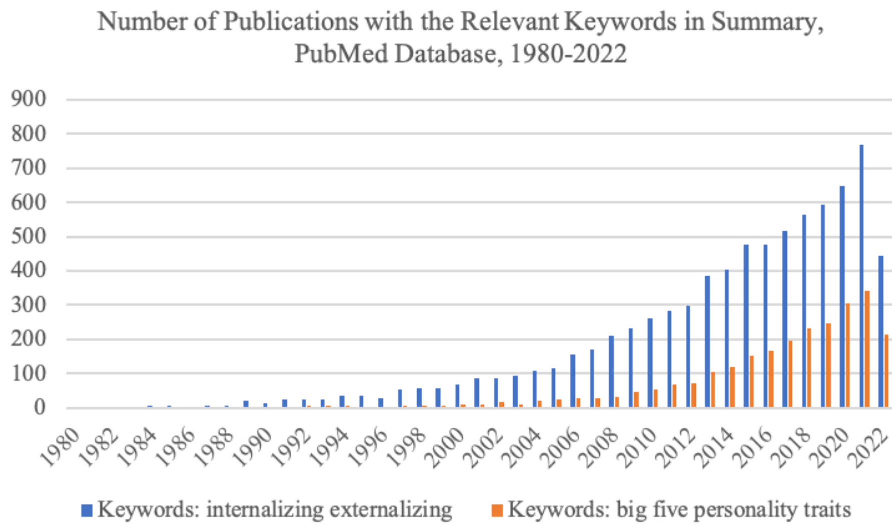


Figure 1: NUMBER OF PUBLICATIONS: The figure shows the number of publications found each year during 1980-2020 using PubMed and using two keyword queries: "internalizing, externalizing" and "big five personality traits."

Table 3: PRELIMINARY ANALYSIS

Outcome	Years of Education				Log Earnings				
	All		Males	Females	All			Males	Females
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Misbehavior	-0.145				-0.026				
	(0.024)				(0.009)				
Externalizing		-0.095	-0.133	-0.026		0.027	0.033	0.019	0.044
		(0.026)	(0.034)	(0.039)		(0.010)	(0.010)	(0.008)	(0.022)
Internalizing		-0.070	-0.057	-0.089		-0.058	-0.052	-0.059	-0.039
		(0.027)	(0.036)	(0.039)		(0.010)	(0.010)	(0.008)	(0.021)
Cognition	1.151	1.152	1.221	1.073	0.199	0.199	0.098	0.079	0.132
	(0.029)	(0.029)	(0.040)	(0.043)	(0.009)	(0.009)	(0.011)	(0.009)	(0.021)
Educational Attainment							(X)	(X)	(X)
Obs.	7,241	7,241	3,573	3,668	4,888	4,888	4,888	2,643	2,245

Notes: This table presents descriptive evidence linking early skills to educational attainment and earnings. Columns (1) to (4) contain parameter estimates from a linear regression model used to link crude measures of unobserved skills to years of education. Columns (5) to (9) present estimates from a linear regression of log earnings on the same crude measures of unobserved skills. We present results separately by gender in Columns (3), (4), (8) and (9). To construct the crude measures of the three unobserved skills, we sum up all variables used to measure that skill according to Table A1 in Appendix A and then normalize each unobserved skill to have mean zero and standard deviation one. Misbehavior is a normalized aggregated measure, where we sum all variables used to measure both externalizing and internalizing behaviors. We report standard errors in parentheses.

Table 4: MEASUREMENT SYSTEM: FROM SKILLS TO MISBEHAVIORS AND TEST SCORES

	All			Males			Females		
	Exter.	Inter.	Cog.	Exter.	Inter.	Cog.	Exter.	Inter.	Cog.
Hostile Towards Children	1.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000
	–	–	–	–	–	–	–	–	–
Hostile Towards Adults	0.797	0.285	-0.023	0.664	0.388	-0.018	0.783	0.600	0.274
	(0.029)	(0.045)	(0.026)	(0.033)	(0.063)	(0.035)	(0.040)	(0.044)	(0.034)
Anxiety Towards Children	1.116	-0.948	-0.407	0.935	-0.820	-0.407	1.103	-0.297	0.185
	(0.045)	(0.071)	(0.040)	(0.051)	(0.094)	(0.052)	(0.063)	(0.067)	(0.055)
Anxiety Towards Adults	1.053	-1.570	-0.756	0.927	-1.652	-0.784	0.850	-0.581	-0.184
	(0.038)	(0.060)	(0.029)	(0.048)	(0.088)	(0.041)	(0.037)	(0.042)	(0.029)
Inconsequential Behavior	0.491	-0.032	-0.322	0.448	-0.027	-0.312	0.417	0.288	-0.156
	(0.019)	(0.034)	(0.021)	(0.023)	(0.045)	(0.027)	(0.025)	(0.033)	(0.027)
Restless Behavior	0.660	-0.147	-0.526	0.596	-0.143	-0.466	0.568	0.243	-0.376
	(0.038)	(0.073)	(0.047)	(0.045)	(0.098)	(0.062)	(0.050)	(0.075)	(0.062)
Depression	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000
	–	–	–	–	–	–	–	–	–
Withdrawal	-0.721	2.721	0.881	-0.626	2.922	0.954	-0.641	1.970	0.170
	(0.049)	(0.096)	(0.045)	(0.060)	(0.138)	(0.061)	(0.062)	(0.099)	(0.052)
Unforthcomingness	-0.790	2.114	0.751	-0.665	2.147	0.762	-0.722	1.475	0.101
	(0.036)	(0.060)	(0.024)	(0.042)	(0.083)	(0.034)	(0.042)	(0.054)	(0.026)
Write Off Adults and Standards	-0.056	1.299	0.297	-0.007	1.322	0.305	-0.065	1.109	0.060
	(0.020)	(0.041)	(0.020)	(0.023)	(0.055)	(0.028)	(0.026)	(0.043)	(0.026)
Verbal Ability	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000
	–	–	–	–	–	–	–	–	–
Reading Ability	-0.019	-0.009	0.845	-0.014	-0.029	0.846	-0.016	-0.029	0.819
	(0.012)	(0.024)	(0.015)	(0.016)	(0.037)	(0.023)	(0.013)	(0.018)	(0.017)
Non-Verbal Ability	-0.017	0.018	0.910	-0.017	0.027	0.891	-0.004	0.027	0.915
	(0.011)	(0.022)	(0.015)	(0.014)	(0.034)	(0.022)	(0.012)	(0.017)	(0.017)
Math Ability	-0.011	-0.045	0.899	-0.013	-0.045	0.901	-0.004	-0.045	0.878
	(0.010)	(0.020)	(0.014)	(0.013)	(0.031)	(0.021)	(0.011)	(0.015)	(0.017)

Notes: This table presents estimates of factor loadings mapping latent skills to BSAG maladjustment syndromes and aptitude test scores. Estimates are reported for the pooled sample as well as separately by gender. We report standard errors in parentheses.

Table 5: Main Results

Outcome	Years of Education			Log Earnings					
	All	Males	Females	All		Males		Females	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Externalizing Factor	-0.183	-0.226	-0.009	0.120	0.119	0.088	0.098	0.068	0.064
	(0.083)	(0.102)	(0.073)	(0.031)	(0.027)	(0.023)	(0.022)	(0.034)	(0.030)
Internalizing Factor	-0.051	-0.059	-0.196	-0.147	-0.130	-0.149	-0.149	-0.093	-0.051
	(0.080)	(0.118)	(0.072)	(0.029)	(0.026)	(0.027)	(0.025)	(0.035)	(0.032)
Cognition	1.112	1.160	0.932	0.157	0.056	0.091	0.034	0.258	0.117
	(0.044)	(0.062)	(0.051)	(0.015)	(0.015)	(0.013)	(0.013)	(0.021)	(0.021)
Years of Education					0.071		0.038		0.113
					(0.004)		(0.004)		(0.008)
Female	-0.460			-0.900	-0.865				
	(0.055)			(0.020)	(0.019)				
Constant	12.379	12.426	11.851	5.642	4.718	5.645	5.148	4.739	3.341
	(0.084)	(0.104)	(0.122)	(0.019)	(0.055)	(0.010)	(0.049)	(0.022)	(0.096)

Notes: This table presents the parameter estimates from the main model linking early skills to educational attainment and earnings. Columns (1) to (3) contain parameters from the educational attainment equation linking years of education to the three unobserved factors. Columns (4) to (9) present estimates from the earnings equation linking log earnings to the three unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. We report standard errors in parentheses.

Table 6: Wages and Hours Decomposition

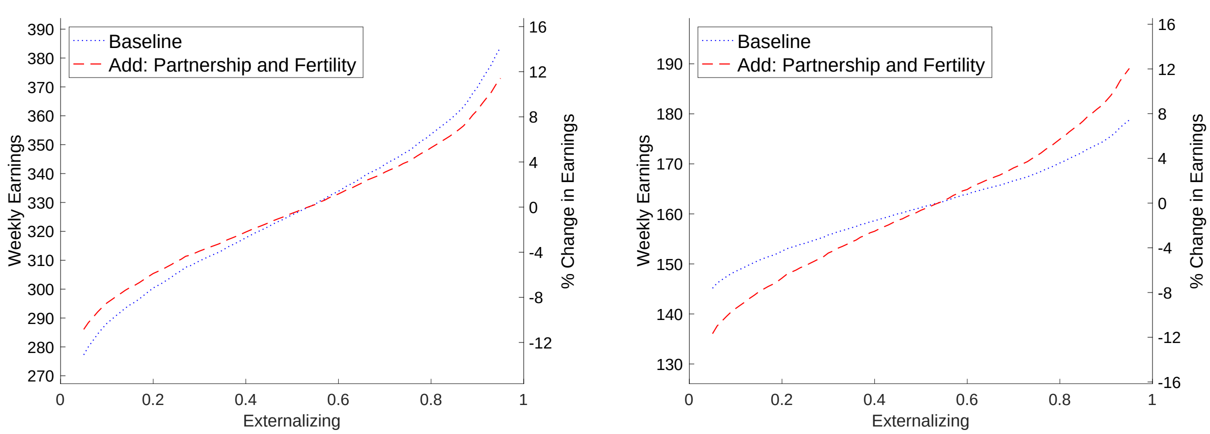
Outcome	Log Hourly Wages			Log Hours Worked		
	All	Males	Females	All	Males	Females
	[1]	[2]	[3]	[4]	[5]	[6]
Externalizing Factor	0.070	0.057	0.030	0.073	0.025	0.054
	(0.019)	(0.020)	(0.017)	(0.019)	(0.011)	(0.022)
Internalizing Factor	-0.089	-0.100	-0.044	-0.061	-0.031	-0.031
	(0.018)	(0.024)	(0.018)	(0.018)	(0.013)	(0.023)
Cognition	0.054	0.054	0.071	-0.001	-0.014	0.048
	(0.010)	(0.012)	(0.012)	(0.010)	(0.007)	(0.016)
Years of Education	0.061	0.043	0.084	0.010	0.603	0.056
	(0.003)	(0.004)	(0.004)	(0.003)	(0.002)	(0.005)
Female	-0.330			-0.535		
	(0.011)			(0.016)		
Constant	1.102	1.334	0.502	3.623	3.822	2.845
	(0.034)	(0.046)	(0.052)	(0.037)	(0.025)	(0.068)

Notes: This table presents the parameter estimates from the main model linking early skills to wages and hours separately. Columns (1) to (3) contain parameters from the wage equation linking log hourly wages to the three unobserved factors. Columns (4) to (6) present estimates from the hours equation linking log weekly hours worked to the three unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. We report standard errors in parentheses.

Table 7: Marriage and Fertility

Outcome	Log Earnings					
	All		Males		Females	
	[1]	[2]	[3]	[4]	[5]	[6]
Externalizing Factor	0.119	0.138	0.098	0.079	0.064	0.101
	(0.027)	(0.026)	(0.022)	(0.022)	(0.030)	(0.027)
Internalizing Factor	-0.130	-0.146	-0.149	-0.125	-0.051	-0.082
	(0.026)	(0.026)	(0.025)	(0.025)	(0.032)	(0.032)
Cognition	0.056	0.041	0.034	0.040	0.117	0.083
	(0.015)	(0.014)	(0.013)	(0.013)	(0.021)	(0.018)
Years of Education	0.071	0.065	0.038	0.038	0.113	0.091
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.006)
Female	-0.865	-0.854				
	(0.019)	(0.017)				
Partnered at 33		0.145		0.158		0.061
		(0.024)		(0.025)		(0.041)
Number of Children at 33		-0.142		0.014		-0.330
		(0.008)		(0.007)		(0.014)
Constant	4.718	4.862	5.148	4.999	3.341	4.021
	(0.055)	(0.058)	(0.049)	(0.053)	(0.096)	(0.091)

Notes: This table presents the parameter estimates from the main model linking early skills earnings while controlling for marriage and fertility. All six columns present estimates from the earnings equation linking log earnings to the three unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. We report standard errors in parentheses.



(a) Males

(b) Females

Figure 2: Controlling for Marriage and Fertility: Figure 2 visualizes the impact of controlling for number of children and partnership status at age 33 at the earnings regression. It illustrates how the predicted weekly earnings in regression models vary after including the partnership and fertility controls, 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.

Table 8: Childhood Disadvantage

Outcome	Years of Education		Log Earnings	
	Low SES	High SES	Low SES	High SES
	[1]	[2]	[3]	[4]
Externalizing Factor	-0.218	-0.191	-0.002	0.145
	(0.170)	(0.096)	(0.065)	(0.031)
Internalizing Factor	0.014	-0.047	-0.034	-0.179
	(0.165)	(0.108)	(0.064)	(0.034)
Cognition	0.781	1.091	0.091	0.044
	(0.068)	(0.047)	(0.029)	(0.015)
Years of Education			0.089	0.070
			(0.013)	(0.004)
Female	-0.325	-0.533	-0.960	-0.850
	(0.121)	(0.061)	(0.050)	(0.021)
Constant	11.579	12.483	4.448	4.739
	(0.151)	(0.096)	(0.164)	(0.058)

Notes: This table presents the parameter estimates from the main model linking early skills to educational attainment and earnings separately for high and low SES children. Columns (1) to (2) contain parameters from the educational attainment equation linking years of education to the three unobserved factors. Columns (3) to (4) present estimates from the earnings equation linking log earnings to the three unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. We report standard errors in parentheses.

Table 9: Police Involvement

Outcome	Police Involvement			Log Earnings					
	All	Males	Females	All		Males		Females	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Externalizing Factor	0.051	0.066	0.021	0.119	0.118	0.098	0.091	0.064	0.068
	(0.011)	(0.018)	(0.007)	(0.027)	(0.027)	(0.025)	(0.023)	(0.032)	(0.031)
Internalizing Factor	-0.030	-0.035	-0.012	-0.130	-0.129	-0.146	-0.141	-0.051	-0.061
	(0.011)	(0.021)	(0.007)	(0.026)	(0.026)	(0.025)	(0.026)	(0.032)	(0.032)
Cognition	-0.041	-0.052	-0.017	0.056	0.057	0.034	0.036	0.117	0.116
	(0.006)	(0.012)	(0.005)	(0.015)	(0.015)	(0.013)	(0.014)	(0.021)	(0.021)
Years of Education				0.071	0.072	0.038	0.037	0.113	0.113
				(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)
Police Involvement					0.033		-0.039		0.124
					(0.045)		(0.027)		(0.134)
Missing Pol. Inv.					0.002		-0.009		0.010
					(0.018)		(0.016)		(0.034)
Female	-0.085			-0.865	-0.863				
	(0.009)			(0.019)	(0.020)				
Constant	0.137	0.156	0.038	4.718	4.712	5.148	5.169	3.341	3.330
	(0.016)	(0.031)	(0.030)	(0.055)	(0.056)	(0.049)	(0.050)	(0.096)	(0.096)

Notes: This table presents the parameter estimates from the main model linking early skills to police involvement and earnings. Columns (1) to (3) contain parameters from the police involvement equation linking a dummy for police involvement to the three unobserved factors. Columns (4) to (9) present estimates from the earnings equation linking log earnings to the three unobserved factors. The coefficients on the three factors have been standardized to represent a 1 standard deviation effect. We report standard errors in parentheses.