Childhood confidence, schooling, and the labor market: Evidence from the PSID

Lucy Page and Hannah Ruebeck*

MIT

June 2022

Abstract

We link over- and under-confidence in math at ages 8-11 to education and employment outcomes 22 years later among the children of PSID households. About twenty percent of children have markedly biased beliefs about their math ability, and beliefs are strongly gendered. Conditional on measured ability, childhood over- and under-confidence predict adolescent test scores, high school and college graduation, majoring or working in STEM, earnings, and unemployment. Across all metrics, higher confidence predicts better outcomes. These biased beliefs persist into adulthood and could continue to affect outcomes as respondents age, since intermediate outcomes do not fully explain these long-run correlations.

*Page: PhD candidate, MIT Economics, Cambridge, MA 02142, lucypage@mit.edu. Ruebeck: PhD candidate, MIT Economics, Cambridge, MA 02142, hruebeck@mit.edu (corresponding author). We are grateful to David Autor, Esther Duflo, Amy Finkelstein, Kartini Shastry, Rohini Pande, Frank Schilbach, and four anonymous referees and the editor for thoughtful and helpful comments. Both authors are supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1745302. This project is also supported by the George and Obie Shultz Fund at MIT.

The majority of data used in this paper is publicly available on the website of the Panel Study of Income Dynamics: https://psidonline.isr.umich.edu/. Several control and outcome variables used in the analysis rely on the restricted PSID dataset, which can be obtained via the process described at https://simba.isr.umich.edu/restricted/ProcessReq.aspx.

1 Introduction

Long-standing research in psychology finds that people have biased beliefs about their abilities in a range of domains.[1] Prior research has focused on "optimism bias," or over-confidence about one's performance, belief accuracy, or future outcomes (Moore and Healy, 2008; Sharot et al., 2011; Taylor and Brown, 1988). In contrast, psychologists also document "imposter syndrome," a form of systematic under-confidence in which people attribute their successes to luck or effort rather than skill (Langford and Clance, 1993; Sakulku, 2011). Recent lab-based work in behavioral economics has sought to microfound this empirical evidence of biased beliefs by documenting that people systematically under-weight or over-weight signals about the truth, especially in ego-relevant domains like intelligence and beauty (see Benjamin (2019) for a review).

Do these confidence gaps matter for economic decision-making in the real world? There are key reasons to expect that they might. For example, if adolescents or young adults perceive ability and educational investment to be complements, under-confident students might exert less effort in school or end their education earlier (Bénabou and Tirole, 2002). Later, under-confident adults may be less likely to complete costly and uncertain job applications, or may select away from jobs with higher returns to performance (Dohmen and Falk, 2011).[2] Individuals' beliefs about their own ability could also affect outcomes by shaping how others perceive them. If parents or teachers mistake confidence for aptitude and expect the returns of education to increase with ability, they may invest more in more confident children (Papageorge et al., 2018; Dizon-Ross, 2019). More confident applicants may appear more capable during job interviews, improving their employment prospects (Mobius and Rosenblat, 2006; Schwardmann and van der Weele, 2019).

As yet, there is limited evidence for how confidence affects economic outcomes in realistic settings and over the long term. In addition to the lab-based work on the short-term impli-

[1] We refer throughout the paper to *ability* and *beliefs about ability,* but we do not mean to imply that ability or beliefs are innate or fixed. Rather, we are referring to someone's ability or perceived ability to perform well in a certain domain or task at a particular time.

[2] Psychological theories of motivation, including Bandura's (1986) Social Cognitive Theory or Expectancy-Value Theory (see Wigfield and Eccles (2000)) also emphasize that individuals increase effort in domains in which they feel competent.

cations of confidence gaps cited above (Mobius and Rosenblat, 2006; Dohmen and Falk, 2011; Schwardmann and van der Weele, 2019), a small parallel literature in economics and sociology examines longer-term outcomes and finds that those with higher self-esteem get more education, are more likely to be employed, and earn higher wages (Murnane et al. 2001; Waddell 2006; Drago 2011; de Araujo and Lagos, 2013). However, this literature has struggled to demonstrate that these associations are not driven by omitted variables like unobserved ability. These papers typically control for IQ in an attempt to account for cognitive ability, but it is not feasible to control for subjects' "ability" across all domains that affect generalized self-esteem.

In this paper, we address the limitations of both prior literatures by examining the real-world and long-term implications of a dimension of confidence in which we can observe and control for demonstrated ability: childhood over- and under-confidence in math.[3] We use unique data from the Panel Study of Income Dynamics (PSID) to identify biased beliefs in math in a sample of 2,985 children in core PSID households; we then relate their childhood over- and under-confidence to educational and employment outcomes up to 22 years later, controlling for test scores, general confidence, and other key confounders.

The PSID is an ideal setting in which to examine long-term links with childhood confidence. Our sample is based on child-focused PSID supplements that measure children's performance on a standardized math test and their own reports of how "good" they are at math. We combine these measures to identify over-confident children as those who scored poorly on the math assessment and yet said they were good at math, and to identify under-confident children as those who scored well but said they were bad at math. The structure of the PSID also allows us to observe much of respondents' young adulthood: the child supplements and core survey followed our sample from 1997 through 2019, so we observe our oldest respondents from age 12 into their thirties.

Biased beliefs about math ability are prevalent in our sample: 5-20 percent of children are markedly over-confident and 7-16 percent are markedly under-confident (using several definitions

[3] We report all of the following analysis for parallel measures of reading over- and under-confidence in Appendix Tables A2-A5. We discuss our focus on math confidence in Section 2.

of biased beliefs, described in more detail below).[4] Over- and under-confidence in math are highly gendered: girls are 2.3 percentage points (pp) (17 percent) more likely to be under-confident and 2.7 pp (27 percent) less likely to be over-confident in math than boys. In contrast, girls are 30 percent *less* likely to be under-confident in reading than boys. This pattern is consistent with evidence that adults are more likely to be over-confident in stereotypically gender-congruent domains (Coffman, 2014; Coffman et al., 2019; Bordalo et al., 2019; Shastry et al., 2020).

One key concern with our measures of over- and under-confidence is that they may just capture children's private information about their own ability, driven by measurement error in the cognitive tests. We have several key pieces of evidence against this concern. First, the math assessment that we use has high test-retest reliability (Hicks and Bolen, 1996). Second, over- and under-confidence persist between waves of the child survey among the 60 percent of our sample with multiple measurements, so our measures seem to capture a stable psychological trait. Next, as we've noted, our measures show gender variation that is consistent with prior work on gendered patterns in belief updating, and which we would not expect to see in random testing error. Finally, our results largely persist when we use alternate measures of childhood over- and under-confidence that are less vulnerable to measurement error; we calculate these measures based on test scores and self-reported ability averaged over two waves of the PSID child supplement.[5]

Our main analysis is simple: we estimate the associations between biased beliefs about one's math ability in childhood and later educational and employment outcomes, controlling for childhood math and reading score deciles, working memory, general confidence, and a host of information on respondents' demographics and family backgrounds.

Children's biased beliefs in math strongly predict many of their medium- and long-term educational and employment outcomes. First, confidence has large associations with educational

[4]Using weights that adjust our sample to be nationally representative, these ranges are 6-30 percent and 6-15 percent, respectively.

[5]While these four pieces of evidence strongly suggest that our measures of over- and under-confidence capture more than *random* measurement error on the cognitive test, they do not negate the possibility that children have private information on a form of math ability that the test systematically excludes. We discuss this possibility in detail in Section 3.3.

achievement: over-confident children score higher than others with comparable prior scores on math assessments five years later, while under-confident children score lower. Biased beliefs in math also predict educational attainment: over-confident children are more likely to graduate from high school and under-confident children are less likely to graduate from college than others with comparable childhood scores. Under-confident children are also less likely to major in STEM during college and attend less selective colleges, though the latter result is imprecisely estimated. Finally, childhood math confidence predicts key employment outcomes at ages 26 and up. Under-confident children are less likely to work in STEM occupations as adults, and we find suggestive evidence that more confident children earn more and are less likely to be unemployed.

While we do not claim that these associations are causal, we do show that they are robust to several key potential confounders. First, children may form inaccurate beliefs about their ability in part because of how their parents or teachers perceive them, and these adult beliefs may themselves affect children's later success (Papageorge et al., 2018; Jussim and Harber, 2005; Wang et al., 2018). However, our main results are robust to controlling for parent and teacher expectations for children's later educational attainment, teacher perceptions of children's competence, and parent-reported measures of investment like often doing homework with their child. Second, children may assess their own ability relative to their school or classroom, while we evaluate their demonstrated ability relative to a national sample. We are limited in our ability to measure school quality, but the measures we do have – proxies for school income, investment, and average achievement – do not correlate with over- and under-confidence, conditional on our other controls. Controlling for these measures of school quality does not change our results. Finally, our results are also robust to controlling for childhood "Big-Five" personality traits, suggesting that over- and under-confidence in math are distinct from these more commonly-studied attributes.[6]

In addition to testing these confounders, we also show that our results hold when we use fourteen different formulations of over- and under-confidence – varying all of the key decision points

[6]The PSID child assessments do not include standard psychometric scales for the Big Five, so we construct proxies for these traits using parents' reports of children's behavior and personality. See Appendices B and E for details.

in constructing our main measures – as our key independent variables.

Two dynamic patterns could underlie the associations we estimate. First, children's over- and under-confidence could alter early patterns of educational investments by parents, teachers, or children themselves; these early investment patterns could then snowball forward into long-term gaps in education and employment. On the other hand, if children's biased beliefs persist, they may have direct psychological effects on choices and performance at each stage in a young adult's development, conditional on his or her performance up to that point. Our evidence suggests that this latter explanation may play a role in the associations we observe. Over- and under-confidence persist through adolescence and into young adulthood (ages 18-27), so biased beliefs could continue to directly affect young adults' decision-making as they age. Childhood confidence also continues to substantially predict later-life education when we hold fixed all intermediate outcomes.

Our results suggest that over- and under-confidence merit study as psychological traits with key economic implications. While our results are not causally identified, they are consistent with childhood confidence having important effects on later-life outcomes. Our evidence is also consistent with the idea that those with more confidence fare uniformly better: under-confident children have worse outcomes than their peers with comparable test scores, while over-confident children have better.[7] Our results leave ample room for future work: to experimentally test the impacts of childhood biased beliefs, to clarify the mechanisms underlying the associations we observe, and to design and test interventions that build confidence in childhood and later life.

Our paper contributes to three literatures in economics.[8] First, we add to a recent literature estimating the returns to psychological or social attributes in the labor market; we provide the first evidence on the returns to over- or under-confidence in the specific academic domain of math. In

[7] Since girls are more likely to be under-confident in math and less likely to be over-confident, these associations could help to explain key gender gaps in the labor market. Unfortunately, our results are too imprecise for us to conclude whether controlling for biased beliefs in math reduces the gender gaps in adolescent test scores, majoring or working in STEM, or earnings.

[8] Psychology research on academic confidence studies how these beliefs develop as children age (e.g. Eccles et al. 1984) and depend on social constructs like gender and race (e.g. Herbert and Stipek 2005; Usher and Pajares 2006). This work relies on self-reported psychometric scales and does not compare self-reported ability to a measure of objective ability, as we do in this paper.

addition to the work on general self-esteem and long-term outcomes that we cite above, parallel literatures examine the associations between economic outcomes and the Big-Five personality traits (Almlund et al., 2011; Heckman et al., 2019), competitiveness (Buser et al., 2021), and children's time, risk, and social preferences (List et al., 2021). While our data do not measure children's competitiveness or time and risk preferences, our results are robust to controlling for measures of the Big-Five traits in childhood. Together with this prior work, our paper suggests that future work should disentangle the economic importance of these various traits.

Second, we extend the literature on asymmetric belief updating in adults by documenting over- and under-confidence in a large sample of children in a real-world setting. This heterogeneity matches the lab-based economics literature, which has found mixed patterns of asymmetric updating (Benjamin, 2019; Zimmermann, 2020). As we've noted, the gender gaps in math confidence that we observe are consistent with lab-based evidence that people over-weight positive ability signals in stereotypically gender-congruent domains (e.g. Coffman et al., 2019).

Finally, the studies most relevant to our own examine the role of beliefs about ability in educational settings. Owen (2020) shows that male college students over-estimate their own ability in STEM and under-estimate the ability of others, while women are more likely to over-estimate others' ability; giving students information about their ability then shrinks gender gaps in beliefs and STEM credits. We find that even children have biased beliefs about their own abilities, with similar gendered patterns. Since children's beliefs may be more malleable than those of college students, our work suggests that interventions like Owen's may be fruitful at younger ages. Owen does not assess whether the de-biasing intervention has effects beyond the same semester, but our results suggest that longer-term effects could be substantial.

While Owen (2020) intervenes specifically to change students' beliefs about their ability, other interventions target self-perceptions more broadly. For example, several studies show that building children's generalized self-efficacy and grit can narrow gender gaps in both confidence and willingness to compete in math (Falco et al., 2010; Alan and Ertac, 2019). Similarly, Carlana et al. (2018) find that a multifaceted career-counseling intervention among high-achieving immigrant students in Italy increases self-efficacy and successfully closes native-immigrant gaps in pursuing a more academic high-school track. In contrast, we study math-specific confidence.[9] We also show that math confidence predicts long-term outcomes even when controlling for general confidence, so interventions to close math confidence gaps may be important complements to interventions that build general self-efficacy or grit.

Finally, Diamond and Persson (2017), the only related paper that considers both biased academic beliefs and long-term outcomes, show that receiving an undeservedly marked-up grade on a test at ages 14-16 leads to higher later test scores, more likely high school and college graduation, and higher earnings. Since marked-up scores in one subject raise later scores across all subjects, the authors argue that these effects arise in part by changing students' beliefs about their own ability. However, they do not actually observe students' beliefs about their own ability, as we do. Together, our papers strongly suggest that students' biased beliefs about ability matter for later educational and employment outcomes.

The paper proceeds as follows. Section 2 lays out a conceptual framework for how childhood confidence might affect economic outcomes, and Section 3 describes our sample and our measures of biased beliefs. Section 4 analyzes the prevalence and predictors of childhood over- and under-confidence in our sample, and Section 5 describes our strategy for estimating the links between biased beliefs in math and long-run economic outcomes. Section 6 presents results, Section 7 describes the stability of our results to potential confounders and alternate definitions of confidence, and Section 8 explores the dynamic patterns that these long-term associations might follow.

[9]Contemporaneous work by Anaya et al. (2021) uses the same data from the PSID and its child supplements to examine the relationship between majoring in STEM and early childhood achievement, self-assessed ability, and parent occupation, though they focus on including parent occupation as a novel explanatory variable in this regression. Like theirs, our main specifications include indicators for whether children's parents work in STEM, but adding these controls does not change our results. Anaya et al. also describe similar gender gaps in ability beliefs to those we document, but they do not specifically study over- and under-confidence or their relationships with long-term outcomes. In addition to this difference in our central research questions, we see our work as building on theirs in three ways: (1) We use a more comprehensive set of available data from the PSID and its child supplements; (2) We consider a larger set of outcomes observed over a much longer time frame; and (3) We define several new measures of over- and under-confidence to deal with complications with the raw data, an issue that Anaya et al. do not discuss.

2 How might childhood math confidence affect economic outcomes?

Ability or skill is a primary independent variable in almost every economic model of student and worker decision-making. These include settings where agents are investing in their own futures, like deciding to continue with schooling, choosing a college major or career, or searching for a job (e.g. Becker, 1964; Roy, 1951; McCall, 1970; Borjas, 1987; Kirkeboen et al., 2016), as well as settings where teachers or parents decide, for example, how to invest in or tailor their pedagogy to a child (Fryer, 2018; Dizon-Ross, 2019).

Over- and under-confidence would enter any of these models if ability is imperfectly observed: by parents, teachers, and even by the student or worker themself. Where ability and effort are complements, like college applications, over-confident agents may work harder. Consistent with these cases, psychological theories of motivation, including Bandura's (1986) Social Cognitive Theory or Expectancy-Value Theory (see Wigfield and Eccles (2000)), emphasize that individuals are more likely to attempt and succeed at tasks in which they feel competent. Where ability and effort are substitutes, like some school tests, over-confident agents may reduce their effort. Bénabou and Tirole (2002) model how over-confidence can persist in equilibrium in either setting.

Over- and under-confidence may also affect outcomes in any setting where teachers or parents decide how to invest time and resources into children based on their perceptions of each child's ability. If adults interpret more confident children as more skilled, they may over-invest in over-confident children and under-invest in under-confident children. Dizon-Ross (2019) shows that parents have inaccurate beliefs about their children's academic performance, and that correcting those beliefs causes them to adjust their investments. Similarly, Papageorge et al. (2018) show that having a teacher with higher expectations increases a student's chance of completing college. The same forces could operate in job applications, where potential employers are uncertain about applicants' skill: in lab experiments, Schwardmann and van der Weele (2019) show that interviewers rate more confident job applicants more favorably, and Mobius and Rosenblat (2006) show that employers offer higher wages to more confident workers.

Our focus on confidence gaps in math, not reading

Our data allows us to identify over- and under-confidence in both math and reading, but we focus the remainder of the paper on biased beliefs in math for several reasons. First, performance in math can be measured more objectively than performance in reading, so children's beliefs about their math ability may be more precise. Next, past work suggests that math ability during childhood and young adulthood more strongly predicts later achievement than does reading ability (e.g. Duncan et al., 2007; Castex and Kogan Dechter, 2014; Goodman, 2019). We find similar patterns in our data in Appendix Table A1, where we regress our main education and employment outcomes on childhood test scores and the set of controls that we will use throughout our main analysis. While both math and reading score percentiles predict later academic achievement and attainment, only math scores predict earnings, unemployment, and majoring in STEM. Thus, children's perceptions of their own ability in math may also link more strongly with later-life achievements than do their self-perceptions in reading. Finally, the Bureau of Labor Statistics predicts that employment in STEM occupations will continue to grow at faster rates than non-STEM occupations through 2030, so math ability may become an even more important predictor of success in the labor market.

That said, we conduct all of the subsequent analysis for reading confidence (Appendix Tables A2-A5). Reading confidence robustly predicts few educational or employment outcomes.

3 Measuring confidence and later-life outcomes in the PSID

3.1 Sample and survey design

We explore the links between biased childhood beliefs and outcomes in young adulthood using the rich data of the Panel Study of Income Dynamics (PSID). The PSID was first collected in 1968 among 5,000 nationally-representative households from two independent samples: a national sample of low-income families from the Survey of Economic Opportunity (the "SEO sample") and a national sample drawn by the Survey Research Center (the "SRC sample"). The PSID has since surveyed the descendant households of the original sample annually from 1968 to 1997 and biennially thereafter, adjusting the sample in 1997 to again make it nationally representative.

We combine the core PSID with two supplements that follow respondents from childhood into young adulthood: the Childhood Development Supplement (CDS) and the Transition into Adulthood Supplement (TAS). The CDS was introduced in 1997, sampling up to two children per PSID household who were then between the ages of 0 and 12 (3,563 children). The CDS collects detailed information from children themselves, from their primary caregivers, and from their elementary school teachers on areas including children's cognitive and emotional development, health, and exposure to parenting practices. The original CDS sample was re-interviewed in 2002-2003, then aged 5-17, and those still below age 18 were included in a third CDS wave in 2007.

In 2005, the PSID introduced the TAS as a bridge between the CDS and the main PSID survey for CDS respondents, the oldest of whom had reached ages 18 to 20 by that year. The TAS has been collected biennially since 2005, with younger CDS respondents aging into the TAS sample at 18. Individuals participate in the TAS until they become economically-independent heads of their own household, at which point they enter the adult PSID sample and are surveyed every two years. The TAS is designed to capture respondents' social and career development as they enter adulthood; we use its modules on education, employment, income, and personality.

The PSID-CDS-TAS data structure is uniquely suited to exploring the links between childhood confidence and long-term educational and employment outcomes. First, the CDS both administers a math test and asks children to evaluate their own math ability; we combine children's test scores and self-assessments to identify over- or under-confidence in math. Section 3.2 below details the CDS tests, self-assessments, and our confidence measures. Second, following CDS children into the TAS and then the PSID allows us to observe detailed data on educational and employment outcomes over 22 years, following our oldest respondents into their mid-thirties. Finally, the extensive data on parents' employment and income in the PSID and on parenting practices and other child characteristics in the CDS allows us to control for many covariates that could confound the relationship between biased beliefs and long-run outcomes.

For example, the detailed child module in the CDS allows us to control for other forms of ability and confidence that are distinct from skill and confidence in math, but which may correlate with them. We construct a measure of general confidence as the mean of standardized variables capturing whether children see themselves as broadly competent (see Appendix B for details); we have no measure of true ability by which to normalize this general confidence scale, so we use it as a control for unobserved abilities and other dimensions of confidence that may correlate with biased beliefs in math and also affect later-life outcomes. We also control for children's scores on the Digit Span subtest of the Wechsler Intelligence Scale for Children (Revised), a measure of short-term memory. Next, the CDS and core PSID collect detailed household information on total family income, household heads' education, primary caretakers' values and mental health, household structure, and financial characteristics like whether the household receives food stamps. Section 5 will detail the family and child variables that we control for in our main analysis.

Our final sample consists of the 2,985 CDS respondents with at least one year of math cognitive tests and self-assessments in the CDS, about 84 percent of all CDS respondents.[10] We report summary statistics for this sample in Appendix Table A6; all variables are observed in the same year in which we first observe childhood over- or under-confidence in math.

Our sample is non-randomly selected from the national population, both because the initial 1968 PSID sample oversampled low-income families and because there is unobserved selection in whether CDS participants report math test scores and self-assessments.[11] This selection appears in our sample statistics in notable ways. First, our sample is disproportionately Black: 45.8 percent are White, 41.7 percent are Black, and only 7.5 percent are Hispanic, while the U.S. Census Bureau reports that 69.1 percent of the US residents were White, 12.1 percent Black, and 12.5 percent

[10] In both the full CDS sample and our final analysis sample, 53 (38) percent of children are descended from the SRC (SEO) sample and 9 percent of children are from the immigrant sample added to the PSID in 1997.

[11] Most children who are missing test scores or self-assessments lack this data because they skipped the entire section of the CDS administered to the child, while completing the survey portions administered to the primary caregiver. These respondents largely have similar demographics to those for whom we observe confidence measures, but their mothers are less likely to have a high school degree, they have lower total family income, and they are about a year younger. Students who take the math cognitive assessments but do not give self-assessments (about 25 percent of the children who are missing test scores or self-assessments) score much lower on both the math assessment and the Digit Span memory test (Appendix Table A17).

Hispanic in 2000 (Greico and Cassidy, 2001). While the Census Bureau reports median household income in 1997 of $55,336, our sample's median taxable income is slightly lower, at $52,029 (both in 2016 USD). On the other hand, our study sample performs disproportionately well on the CDS standardized tests: we observe median CDS math and reading score percentiles of 60 and 54, respectively, relative to national norming samples.

While we do not weight our sample to be nationally representative in our main analysis, we include results that do so in Appendix Tables A7-A10. These weights are based on those published by the CDS, which capture the inverse probability of respondents' inclusion in the CDS sample; we then recalibrate these CDS weights via iterative proportional fitting, or raking, to ensure that our sample matches marginal distributions of percentile CDS math scores, race in 2000, and total household income in 1997. Our main results are less precisely estimated when we use weights, though they remain qualitatively similar.[12] We present all descriptive statistics both for the weighted and unweighted samples.

3.2 Measuring over- and under- confidence in math

Data on children's self-reported and demonstrated ability in math

The CDS assesses children's math skills using the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R), a test of academic achievement commonly used by school psychologists in the 1990s (Stinnett et al., 1994; Hicks and Bolen, 1996; Duffy and Sastry, 2014). The CDS administers the Applied Problems subtest of the WJ-R, comprising 60 word problems of increasing

[12] Since the error terms in our regressions are unrelated to the sampling criterion, conditional on our extensive controls for family income, race, and other characteristics, weighting may not improve the estimator's consistency and may reduce its precision (Solon et al., 2015). These recalibrated weights put less weight on children with high CDS math scores, in some cases leaving us underpowered to detect the correlations between under-confidence and long-term outcomes. A natural concern is that our unweighted regressions may estimate non-representative partial relationships between confidence and outcomes, if these associations are heterogeneous by race or income. However, weighted least squares estimates are not necessarily closer to the true population average partial relationship than ordinary least squares estimates (Solon et al., 2015). Instead, we directly estimate heterogeneity by characteristics related to the sampling scheme, like family income, race, age, and being in an SRC-sample family. These results are imprecise and show no robust patterns of heterogeneity (results available upon request).

difficulty that assess math reasoning and knowledge.[13] Each child completes only a subset of the test, beginning at a "basal" level, where they answer six consecutive questions correctly, and ending at a "ceiling" level, where they get six consecutive questions wrong. The CDS then reports each respondent's percentile rank relative to the nationally-representative WJ-R norming sample for their age group; we use these percentile ranks as our measure of each child's demonstrated ability in math. Panel A of Figure 1 shows the distribution of these scores in our sample.

In addition to collecting this measure of performance in math, the CDS also asks all respondents ages 8 or older to assess their own ability in math, asking them to answer "How good at math are you?" on a scale of 1 (not at all good) to 7 (very good). Children never receive their scores on the WJ-R math test, so these self-reports do not reflect feedback from the CDS. Panel B of Figure 1 shows the distribution of these self-assessments. Math self-perceptions are highly skewed towards positive responses, with over 89 percent of respondents ranking themselves as "Okay" or better at math. This skew may be partially explained by the distribution of percentile scores in Panel A, which skews heavily towards higher-performing children. While shifted upwards, children's self-reports do contain information about objective ability: in Panel C of Figure 1, average math test percentiles rise almost linearly with self-reported ability in math.

We measure children's over- and under- confidence in math in the first wave of the CDS in which they have non-missing cognitive test scores and self-assessments, leaving us with a sample of 2,985 children.[14] We first measure confidence for the median child before age 12, and we observe confidence by age 13 for 83 percent of children. Thus, we will interpret our measures as *childhood* over- and under-confidence in math. Throughout, our analysis will control for both birth year and the age at which we first observe confidence.

Defining binary measures of over- and under-confidence

We first identify over- and under-confidence in math using large mismatches between children's

[13] The 1997 CDS wave also included 58 WJ-R questions on calculation skills, and we use this test in the next section to assess the reliability of our over- and under-confidence measures.

[14] We first observe confidence from the 1997 CDS wave for 1,075 children, from the 2002 CDS wave for 1,347 children, and from the 2007 CDS wave for 563 children.

score percentiles and their self-assessments. In particular, we classify any respondent as under-confident in math if she scored above the 75th percentile nationally and ranked her own ability at 1 to 4, corresponding to the bottom 47 percent of the subjective-ability distribution in our sample, or if she scored above the 50th percentile nationally and ranked herself at 1 to 3, corresponding to the bottom 10 percent of the subjective-ability distribution. We define over-confidence among low-achievers using similar thresholds, but we account for the skewed self-assessment distribution by using stricter cut-offs to identify biased beliefs. In particular, we identify any respondent as over-confident in math if she scored below the 25th percentile nationally and rated her own ability at 6 or 7, corresponding to the top 39 percent of the subjective-ability distribution in our sample, or if she scored below the 50th percentile and rated herself at 7, corresponding to the top 22 percent of the subjective-ability distribution.

These measures of math over- and under-confidence have several key strengths: they are easy to define and observe, they refrain from putting too much stock in the cardinal value of children's self-assessed ability, and they account for the upward skew in self-assessments, which we consider to be a form of response bias separate from over- or under-confidence.[15]

However, our measures also have several limitations. First, we can only identify over-confidence among children scoring below the 50th percentile and under-confidence among those scoring above the 50th percentile; however, this strategy matches the existing literature, which typically documents under-confidence (imposter syndrome) among high-achievers (Sakulku, 2011). Another limitation is that these measures are not directly comparable to measures of over- and under-confidence from the lab-based literature, which can precisely measure respondents' beliefs about their quiz performance or rank relative to a group (e.g. Coffman et al., 2019; Möbius et al., 2014; Eil and Rao, 2011). Our measures of over- and under-confidence, in contrast, identify coarse cate-

[15]For example, upward response bias could arise based on children's interpretation of the qualitative labels on the scale ("Not at all good" at 1, "Okay" at 4, and "Very good" at 7) if, for example, they think that nearly everyone is at least "Okay" at math. Upward skew could also arise if self- or social-image concerns make children unwilling to tell a surveyor that they are worse than "Okay" at math. If, on the other hand, this upward skew does reflect true aggregate over-confidence, our estimates for long-term associations with over-confidence would simply reflect links with *particularly* over-confident beliefs in math.

gories of children with large gaps between their self-assessments and observed scores. Our second measure of biased beliefs, described below, aims to partially address these limitations.

Defining a more continuous measure of biased beliefs

Our second confidence measure identifies biased beliefs as the difference between children's self-reports and their observed performance on the CDS math test. To transform these objects to the same scale, we split the distribution of children's percentile scores uniformly into seven bins, where 1 includes the lowest 14 score percentiles and 7 includes the highest 14 score percentiles relative to the national norming sample. We then assume that students with full information about the national distribution of scores and their place in it would have self-reported their math ability as the bin from 1 to 7 in which their score percentile falls; we take the difference between their actual self-report and this bin as our measure of biased beliefs. This measure then takes on integer values from -6 to 6. For ease of interpretation, we standardize this variable to have mean 0 and standard deviation 1 throughout the rest of the paper.

This measure has three strengths relative to our main measure: it allows for more granularity in the extent of biased beliefs, aligns more closely with measurements of biased beliefs in the lab-based literature, and relies on fewer choices by the authors. However, by assuming that we can identify even small biases in beliefs about math ability, it is more likely to conflate actual biased beliefs with children's private information about their math ability (described in more detail in the next section). It may also be confounded by forms of reporting bias other than over-confidence that generate the overall upward skew in self-reports (see footnote 15).

We present results for all outcomes using both the binary and more continuous formulations of biased beliefs, and in general the results are extremely consistent. To ensure that our main results do not arise just from our particular choice of confidence measures, we show that they are robust to a range of alternate definitions of both our indicators for over- and under-confidence and this more continuous measure of biased beliefs. We describe these alternate measures in Section 7.

3.3 Biased beliefs or measurement error?

One key concern with our measures of biased beliefs in math is that they may conflate over- and under-confidence with children's private information about their math ability, perhaps driven by measurement error in the WJ-R assessment. Four key pieces of evidence support the claim that our measures truly capture biased beliefs in childhood.

First, prior work has shown that the WJ-R assessment is a reliable measure of children's math skills, with test-retest reliability for the applied math problems of about 0.85 in large samples (Hicks and Bolen, 1996).[16] We can also verify WJ-R reliability across math domains in our sample using the 1997 wave of the CDS, which administered both the Calculation and Applied Problems subtests of the WJ-R. For the 1,450 children who took both tests, the correlation in percentile ranks on the two sections is 0.69. Our binary designations of children as over-confident, under-confident, or neither are also highly consistent whether we measure objective math ability using children's percentile scores on the Calculation or Applied Problem subtest: 81 percent of children with both measures are classified in the same category regardless of which ability measure we use. Another ten (nine) percent switch from under-confident (over-confident) to neither or vice versa.[17]

Second, our measures of childhood math confidence persist over time. About 60 percent of the children in our sample appear in two waves of the CDS, allowing us to construct two measures of over- and under-confidence taken five years apart. Children appear in a second CDS wave at ages 13 to 19, so these second-wave measures capture biased beliefs in adolescence. Table 1 regresses our *adolescent* measures of biased beliefs on our *childhood* measure of the same variable, controlling for a set of demographics and parent characteristics that we will use throughout our later empirical analysis; we outline these specifications in detail in Section 5.[18] These regressions show

[16] Several studies find test-retest reliability of about 0.75 for certain ages, though these studies use small samples (Shull-Senn et al., 1995)

[17] We find similar reliability using our more continuous measure of degrees of confidence, which takes on integer values from -6 to 6. There, 32 percent of children are assigned the same value regardless of which math test we use as the measure of demonstrated ability, 62 percent are within one integer, and 83 percent are within two integers. See Appendix Figure A3 for the full joint distribution of the more continuous confidence measures based on the two math subtests.

[18] The regressions in Table 1 add controls for children's *adolescent* test score deciles in math and

substantial persistence: respondents who were over-confident in math as children are about 3 times as likely (12pp more likely) to be over-confident in math as adolescents, while under-confident children are about 1.7 times as likely (4pp more likely) to be under-confident as adolescents.[19] Similarly, a one-standard-deviation increase in the degree of biased beliefs in childhood predicts 0.18sd more biased beliefs in adolescence. If our confidence measures just captured random testing variability, we would not expect to see such substantial persistence.

Third, our main results are largely robust to using measures of over- and under-confidence that reduce potential measurement error by combining observations of children's test scores and self-reported ability across two waves of the CDS. We discuss these measures and results in more detail in Section 7. If measurement error is uncorrelated across tests taken 5 years apart, these average confidence measures will be less vulnerable to it than are our main measures.[20]

Finally, we describe in the next section that we observe substantial gender gaps in math over- and under-confidence, with girls more likely to be under-confident and less likely to be over-confident. This pattern is consistent with gender stereotypes about math ability, which may shape children's beliefs even at young ages, and mirrors results for adults in the lab (e.g. Coffman et al., 2019). Our measures of over- and under-confidence could only be entirely explained by measurement error if this error took a similar gendered pattern, beyond its correlation with WJ-R Applied

reading to our main specification. We add these controls to purge any correlations induced by the effects of childhood confidence on adolescent test scores, since childhood over- and under-confidence predict later test scores (see Section 6) and higher-scoring (lower-scoring) children are mechanically more likely to be classed as over-confident (under-confident).

[19]While our main measure of under-confidence persists only weakly into adolescence, several alternate definitions of under-confidence are strongly persistent (Appendix Figure A4). Our main measure's limited persistence may relate to the fact that adolescent test scores are much less upward-skewed than childhood test scores, so fewer respondents can be classified as under-confident in adolescence. The more persistent alternate definitions of under-confidence, in contrast, identify under-confidence among respondents with a wider set of test scores and thus are less affected by this distributional shift. Like our main measure, these alternate measures predict substantial gaps in long-run outcomes (Appendix Figures A5-A16).

[20]Despite this benefit, we do not use these averages as our preferred measures of confidence for three reasons: (i) over- and under-confidence at older ages may be more likely to be confounded by unobserved variables; (ii) we are interested in adolescent test scores and confidence measures as outcome variables; and (iii) only 60 percent of our sample has confidence measurements over multiple waves of the CDS.

Problems scores and with the many other controls we outline in Section 5.[21] We consider a few possible sources of non-random measurement error that could generate these patterns: skill in some dimension of math that the test does not cover, test-taking anxiety, and test-taking motivation.

First, the CDS data allow us to test for gender gaps in one central dimension of math skill that our main test scores do not directly capture: calculation skills. Using the 1997 CDS sample, when children took both the WJ-R Calculation subtest and the WJ-R Applied Problems subtest, we find no evidence that boys have better calculation skills conditional on the applied problems scores that we use in our main analysis.[22]

Next, differential measurement error in the CDS math tests could arise if boys or girls are more prone to testing anxiety that impairs performance. While past work finds that boys show higher physiological stress during test-taking (Weekes et al., 2006; Stroud et al., 2002), other research suggests that physiological stress only impairs performance when students psychologically appraise it as an indicator of potential failure (Jamieson et al., 2013; Mattarella-Micke et al., 2011). Girls tend to have higher psychological test anxiety and math anxiety, and most commentary suggests that it is these psychological manifestations of anxiety that pose first-order risks to test performance (Devine et al., 2012; Erturan and Jansen, 2015; Ballen et al., 2017). Thus, we would expect girls' test performance to differentially lag their true skill, producing gender gaps in confidence that would conflict with our empirical results.

Finally, we turn to test-taking motivation. Past work finds that girls are somewhat more moti-

[21] Differential random error by gender could not fully explain the gendered patterns of over- and under-confidence we observe, since the gender with more variable performance would be more likely to be both over- and under-confident. Nonetheless, comparing boys' and girls' performance on the Calculations and Applied Problems subtest in the 1997 CDS sample suggests that neither gender has differentially variable test performance. 81% of both boys and girls receive the same binary confidence designation when calculated using either the Calculations or Applied Problems percentile score as a measure of math skill, and the joint distributions of the more continuous measures are very similar for boys and girls (Appendix Figure A3).

[22] We estimate the following regression: $CALCpctile_i = \beta_0 + \beta_1 APpctile_i + \beta_2 Female_i + \beta_3 APpctile_i \times Female_i + \varepsilon_i$. Coefficient β_3 is not significantly distinguishable from zero, and β_2 is significant and positive. Thus, girls have stronger calculation skills than boys conditional on their Applied Problems scores, which would tend to make girls look more over-confident by our measures, the opposite of what we find.

vated than boys to exert effort on low-stakes tests, so boys' CDS math scores may be differentially low relative to their true skill in math (Segal, 2012; DeMars et al., 2013; Gneezy et al., 2019). Then, boys may appear more over-confident by our measures. While it is hard to fully eliminate this possible confounder in our setting, our results are robust to controlling for agreeableness and conscientiousness, two Big-5 personality traits that are positively correlated with unincentivized test effort (DeMars et al., 2013; Segal, 2012). (See Section 7 for more details.)

Together, most evidence from our empirical setting and from past work on test-taking strongly suggests that our confidence measures capture a meaningful psychological trait. However, we cannot fully eliminate the risk that these measures capture children's private information on some aspect of math ability that the test systematically excludes. Any such confounder could only explain our results if it is differentially weak among girls and affects outcomes beyond its correlation with demonstrated math ability, general confidence, digit span score, reading ability, and the many other controls we outline in Section 5.

4 Patterns of over- and under-confidence in the population

This section documents the prevalence and correlates of over- and under-confidence in our sample. Besides documenting biased beliefs in math in a real-world setting, these results are useful both to validate our measures of biased beliefs and to inform our strategy for estimating the links between childhood confidence and long-run outcomes, which we describe in Section 5.

4.1 Prevalence of biased beliefs

We find substantial over- and under-confidence among children in our sample: using our main binary measures, 8.5 percent of children are over-confident at their first measurement, while 12 percent are under-confident.[23] Since these measures identify large gaps between children's self-assessed and objective performance, these shares are strikingly high. Turning to our more continuous measure of biased beliefs, 21 percent of children report the same bin as their percentile score

[23] We find similar results when applying our raked weights to obtain nationally representative estimates: 9.2 percent of children are over-confident and 9.8 percent are under-confident.

would imply, 8.7 percent of children report ability levels that are at least 3 bins lower than that of their score, and 17 percent report ability levels that are at least 3 bins higher, where each bin spans 14 score percentiles. See Appendix Figure A1 for the full distribution of the continuous confidence measure. It is notable that over- and under-confidence are both prevalent in this large sample, given psychology's focus on over-confidence (Moore and Healy, 2008) and the mixed evidence from lab experiments on asymmetric belief updating (Benjamin, 2019).

Next, older children have more accurate beliefs. Panel A of Appendix Figure A2 plots the share of children who are over- or under-confident in math by age; Panel B plots the cumulative density function for the continuous confidence measure for three age groups, pooling respondents' observations across CDS waves. Both panels show that younger children are more likely to have incorrect beliefs about their math ability, and average belief accuracy increases almost monotonically as children age. We focus on the associations between confidence and later-life outcomes using first-observed confidence, so our confidence observations are drawn from young ages with more biased beliefs. We eliminate bias due to the timing of our confidence measurements by including fixed effects for the age at which confidence was measured in all regressions.

4.2 Biased beliefs and other child characteristics

Over- and under-confidence correlate with other child characteristics in largely expected ways (Table 2 and Appendix Table A11). Unsurprisingly, children with higher general confidence are more likely to be over-confident and less likely to be under-confident in math, and children with higher digit span scores are less likely to be under-confident. Math test score deciles strongly predict confidence gaps (though some of this correlation arises mechanically from how our measures are constructed), while reading test score deciles do not (Appendix Table A11). We will control for children's general confidence, digit span scores, and test score decile fixed effects in math and reading in all regressions of later-life outcomes on childhood biased beliefs.

Conditional on these measures of ability, children who have ever been in a gifted program are 8.7pp less likely to be under-confident in math and 2.6pp more likely to be over-confident. These correlations could reflect that schools and children share private information on children's ability conditional on CDS scores, that being in a gifted program alters children's confidence, or that children's confidence influences their treatment at school conditional on ability. To avoid controlling for mediators of the effects of confidence, our regressions will *not* control for this variable or other signals of ability from schools, like repeating a grade.[24]

Finally, gender is the strongest demographic predictor of math confidence. Girls are 2.3pp (20 percent) more likely to be under-confident and 2.7pp (27 percent) less likely to be over-confident in math than boys with the same score deciles, and on average, girls' biased beliefs are 0.1 standard deviations (sd) lower than the average boy's. Note that girls do not have more accurate beliefs, simply more negatively-biased ones.[25,26] This finding is consistent with prior literature showing that adults are more over-confident in gender-congruent domains (e.g. Coffman et al., 2019), but it is notable that we find it in children, the majority of whom have not yet entered puberty. These gender differences are present at almost all ages, but due to small sample sizes the patterns are imprecise (available upon request).[27]

Perhaps surprisingly, we find no significant links between children's math confidence and their parents' education or occupation, household income, or race, conditional on all other characteristics. However, noise in these estimates means we cannot reject potentially large correlations.

[24]Math over- and under-confidence also correlate with children's other attitudes towards math and school in reasonable ways (Appendix Table A18), suggesting that our measures isolate over- and under-confidence in the particular domain of math. See Appendix C for more discussion.

[25]In fact, there is no gender gap in the likelihood of having accurate or almost accurate beliefs (degrees of over- and under-confidence equal to zero, or between -1 and 1, respectively). Results are available upon request.

[26]Appendix Figure A17 shows that this gender gap is extremely robust to using alternate definitions of over- and under-confidence and alternate ways of calculating the more continuous degrees of confidence measure. This figure plots the coefficient on the female indicator when we exchange the dependent variables in Table 2 with these alternate measures (discussed further in Section 7).

[27]We also test whether childhood gender gaps in math under-confidence explain gender gaps in later education and employment outcomes: adolescent test scores, majoring in STEM, and earnings. Specifically, we estimate the change in the coefficient on gender when we estimate our preferred specification with and without the indicator for under-confidence (following Buser et al. (2021)). The results (available upon request) are quite noisy, so we leave it to future research to determine whether confidence gaps in math can help explain these and other gender gaps.

5 Confidence and long-term outcomes: Empirical strategy

Our empirical strategy is simple: we estimate the associations between biased ability beliefs in math and later education and work outcomes, *holding fixed measured childhood ability*. We use the PSID's rich data on childhood environment and family characteristics to control for extensive pre-determined confounders, but we refrain from interpreting our estimates as the causal effects of confidence. We estimate the following specification:

$$Y_{it} = \alpha + \beta_1 Over_{i0} + \beta_2 Under_{i0} + A'_{i0}\mu + X^{C\prime}_{i0}\delta_1 + X^{P\prime}_{i0}\delta_2 + \gamma_s + \omega_t + \varepsilon_{it}$$

where Y_{it} is individual i's outcome of interest in adolescence or adulthood, measured in wave t of the TAS or PSID, and $Over_{i0}$ and $Under_{i0}$ are indicators for being over- or under-confident in math as a child, respectively. All of our main tables also include regressions in which we replace $Over_{i0}$ and $Under_{i0}$ with the single $ZConf_{i0}$ variable, which captures the degree to which a child is over- or under-confident in standard deviations. Due to power limitations, we assume that $ZConf_{i0}$ has a linear relationship with our outcomes of interest.[28]

Next, all of our regressions include A_{i0}, a vector of controls for childhood ability. In particular, A_{i0} includes linear controls for childhood digit span score and general confidence, as well as fixed effects for test score deciles in both reading and math.[29] Our basic specification also includes state fixed effects γ_s, TAS or PSID wave fixed effects ω_t when the outcome is observed multiple times for each individual,[30] a set of child controls X^C_{i0}, and a set of parent controls X^P_{i0}. In our

[28]Appendix Figures A18, A19, and A20 show our main results when we relax this assumption; we plot the coefficients on indicators for each integer value of the variable underlying $ZConf_{i0}$: $Conf_{i0} = -6$, $Conf_{i0} = -5$, ..., $Conf_{i0} = 6$. While these results are noisy, the point estimates suggest that this linearity assumption is reasonable. We also show in Appendix D that we cannot generally reject the null hypothesis that over- and under-confidence predict economic outcomes in similar (opposite-signed) ways, further supporting this linearity assumption.

[29]One might worry that controlling for general confidence absorbs too much of the variation in math over- and under-confidence if over- and under-confidence in math are dimensions of confidence in general. While the economic impacts of general confidence are certainly of interest, we take the conservative approach of isolating math-specific over- and under-confidence as cleanly as possible by controlling for general confidence. That said, our results are remarkably similar with or without the control for general confidence (available upon request).

[30]For some outcome variables, like earnings and unemployment, we have multiple years of out-

first specification, X^C_{i0} and X^P_{i0} include only variables that are certainly unaffected by respondents' childhood math confidence: X^C_{i0} includes fixed effects for race, birth year, quarter of birth, gender, and age at which we observe confidence,[31] and X^P_{i0} includes family income, its square, and fixed effects for both parents' levels of education. All variables indexed at $t = 0$ are from the first CDS wave in which a child had WJ-R scores and an ability self-assessment. Since about two-thirds of the children in our sample have a sibling in the sample, we cluster standard errors by family. Our coefficients of interest are β_1 and β_2.

Our second specification takes advantage of the detailed caregiver interviews in the CDS to add additional controls for child and family characteristics that may correlate with both confidence and long-run outcomes. In addition to expanding the set of child controls, X^C_{i0}, with the primary caregiver's assessment of the child's general health, this specification supplements X^P_{i0} with additional parent and family controls: whether the family receives government transfers; whether the household includes the father or has two adults; parents' beliefs about gender norms and the qualities that are most important for success; and parent mental health (see Appendix B for details). Finally, we add four indicators for whether the child's parents work in STEM or another high-education occupation (based on Anaya et al. (2021); see footnote 9). We focus on this specification throughout the text, but results are generally consistent across these two specifications.

6 Confidence and long-term outcomes: Results

The following section presents our results, documenting strong associations between childhood under- and over-confidence in math and key later-life outcomes: adolescent test scores, graduation from high school and college, college major, career choice, earnings, and unemployment. We

comes across TAS and PSID waves for each respondent. In contrast, we observe our educational outcomes (e.g. whether respondents ever majored in STEM) only once per respondent; we do not include survey wave fixed effects in regressions linking childhood confidence to these outcomes. Note that we do not include respondent fixed effects even in regressions with multiple outcome observations per respondent, since we only measure childhood confidence once. We cluster standard errors by family in all regressions.

[31] Age at which confidence is first observed and birth year are not collinear. For example, children who had their confidence measured when they were 8 years old could have been born in 1989, 1995, or 1999 (and had their confidence measured in the 1997, 2003, or 2007 CDS, respectively).

present these results in Tables 3, 4 and 5.

6.1 Medium-term educational achievement

We first examine the links between childhood confidence and medium-term educational achievement, measured as adolescent scores on the CDS math assessments. We observe these scores at children's second CDS observation, about 5 years after we first observe their confidence in math.

Children's biased beliefs in math significantly predict adolescent math performance (Table 3, columns 1 and 2). Using our binary measures (Panel A), children who are over-confident in math score 2.7 percentiles (standard error = 1.5p) higher on the math assessment five years later than others with comparable baseline scores, while under-confident children score 5.9 percentiles (se = 1.5p) lower. Using our more continuous measure (Panel B), a child with 1 standard deviation (sd) higher math confidence in childhood scores 2.8 percentiles (se = 0.57p) higher on the math assessment 5 years later than others with comparable baseline scores. Children marked as over- or under-confident in our binary metrics differ from others by an average gap of 1.8sd and -1.6sd in continuous degrees of confidence, respectively, so our estimate magnitudes are remarkably consistent across the two panels. In contrast, there is no relationship between childhood math over- or under-confidence and adolescent reading scores using either measure of biased beliefs (Table 3, columns 3-4).

These associations are large relative to the links between raw math ability and later scores: increasing one's childhood math score by 10 percentiles is associated with scoring on average 5.3 percentiles higher in adolescence (Column 1 of Appendix Table A12).[32] Thus, being over- (under-) confident in math predicts as large a gap in adolescent test scores as does increasing (decreasing) one's childhood math test score by 5-11 percentiles.

[32]While we estimate the relationships between confidence and later outcomes in regressions with test score decile fixed effects, here we run an otherwise identical regression replacing these fixed effects with linear controls for test scores. We benchmark the links between biased beliefs against the coefficients on these linear score controls throughout Section 6.

6.2 Educational attainment

Biased beliefs in math during childhood also predict important gaps in high school and college graduation. Children who are over-confident in math are 6.2 percentage points (se = 2.6pp) more likely to graduate from high school, and children who are under-confident in math are 5.8pp (se = 2.8pp) less likely to graduate from college (Table 3, columns 5-8, Panel A). Since only 30 percent of our sample graduates from college, being under-confident predicts a 20 percent drop in the likelihood of college graduation. We find very similar results using our more continuous measure in Panel B: a child with 1sd higher math confidence in childhood is 1.8pp (se = 1.0pp) more likely to graduate from high school and 3.3pp (se = 1.1pp) more likely to graduate from college, though the first is only marginally significant. Again, the magnitudes of these results are similar regardless of which confidence measure we use.

These gaps are large relative to the associations between childhood math scores and educational attainment in our data: childhood math scores generally do not substantively predict high school graduation,[33] and increasing test scores by one decile is associated with being 2.9pp more likely to graduate from college on average (Appendix Table A12, columns 3-4).

6.3 College quality, college major choice, and graduate education

Next, we consider later-education outcomes among those who went to college: college quality, college major choice, and whether respondents complete a graduate degree. Since we restrict to college graduates, these regressions use much smaller samples than for our previous outcomes.

First, we find imprecise links between childhood math confidence and the quality of colleges that children later attend. We consider two quality measures: first, an index of general college quality, and second, colleges' 75th-percentile math SAT scores among incoming freshmen – a more specific measure of math quality.[34] We focus our discussion on colleges' 75th-percentile

[33]Math scores do significantly predict high school graduation, but the coefficient is precisely estimated and very small (increasing test scores by 10 percentiles is associated with being 0.8pp more likely to graduate from high school). The magnitude of this linear coefficient is half of the size of the coefficient on reading scores. Reading skills may be more important for high school graduation than math (e.g. since fewer years of math study are required to graduate).

[34]Using restricted data from the TAS, we link respondents with college quality data from the

math scores (Table 4, column 3 and 4), but our results are similar using the more general college quality index (columns 1 and 2). Under-confident children attend schools whose 75th-percentile math SAT scores are 11.3 points (se = 5.9 points) lower than others with the same childhood scores (p = 0.07); with 95 percent confidence, we can reject that under-confident children attend schools with math SAT scores that are over 0.3 points higher or 22.9 points lower. Over-confidence is not significantly associated with college quality among childhood low-scorers, but again we observe wide confidence intervals: we cannot reject that over confident children attend colleges that have 20 points lower to 26 points higher SAT scores than their peers. Using our more continuous measure of biased beliefs in Panel B yields consistent, but imprecise, results.

Next, we find that childhood under-confidence in math is starkly associated with major choice among those who go to college (Table 4, columns 5 and 6). Among those with a 4-year college degree, students who were under-confident in math are 16.2pp (se = 3.6pp) less likely to earn a STEM major[35] than their peers with comparable childhood scores, an 86 percent drop from the share of STEM majors across all college graduates in our sample. This large gap means that under-confident children who score above the 50th percentile on the CDS math test are only 1.3 times as likely to major in STEM, conditional on going to college, than the average child who *scores below the 50th percentile*; in contrast, other childhood high-scorers are 3.5 times as likely to major in STEM as low-scorers. We obtain very similar results using our more continuous measure of biased beliefs in Panel B: a 1sd increase in confidence is associated with a 7.8pp (se = 2.3pp) increase in the likelihood of majoring in STEM.

Finally, we find no significant relationships between biased beliefs and getting a graduate degree, though again our standard errors are large (Table 4, column 7 and 8).

National Center for Education Statistics (NCES) for the first college they attended in the first year they attended that college. Following Cohodes and Goodman (2014), we construct an index of college quality as the first component from a principal component analysis of colleges' 75th-percentile math SAT scores among incoming freshmen, graduation rates, and per-pupil instructional expenditures, separately by year. Details on variable construction are available in Appendix B. We then standardize this index to have mean 0 and standard deviation 1 in the full sample of four-year colleges in the US by year.

[35] We define STEM fields as engineering, math and computer sciences, and natural sciences. We find similar results if we also include health fields.

6.4 Employment outcomes

Next, we examine the links between childhood over- and under-confidence in math and employment outcomes in young adulthood: occupation type, earnings, and employment status. We follow respondents in the adult PSID when they age out of the TAS, so we observe our oldest CDS respondents through age 36 at the end of our sample period. Since respondents' employment outcomes in their early twenties may not yet be representative of their long-term career trajectories, we restrict the sample to observations in which respondents are older than 25; we observe about 70 percent of our sample above this threshold at least once.[36]

We first consider job choice. Under-confident children are about 4.9pp (se = 1.6pp) less likely to work in a STEM occupation[37] than their peers (Table 5, columns 1 and 2), a gap that is approximately equal to the baseline rate at which respondents later work in STEM in our sample. We find a similar result with our measure of the degrees of over- and under-confidence, where a 1sd increase in childhood confidence is associated with a 1.8pp (se = 0.6pp) increase in the likelihood that one works in STEM. These confidence gaps are large relative to the link between childhood math scores and later STEM employment, which is precisely estimated but very close to zero (Appendix Table A12, column 9).

On the other hand, there are no gaps in the likelihood that over- or under-confident children work in non-STEM high-education occupations[38] (Table 5, columns 3 and 4). These results are reassuring for our empirical design: the fact that math confidence matters for STEM employment, but not other high-education employment, helps to validate that we properly isolate long-term associations with children's biased beliefs *in math*, rather than picking up correlations with unobserved

[36]Appendix Table A19 replicates these results using one observation per child, where the dependent variable is calculated as the average outcome observed over ages 28-33. The results are meaningfully the same.

[37]We define STEM fields to include computer and mathematical occupations, architecture and engineering occupations, and life, physical, and social science occupations. We find similar results if we include healthcare occupations as STEM.

[38]We define non-STEM high-education occupations as management, business, and financial occupations, legal occupations, education, training, and library occupations, and occupations that focus on writing and communication (a subset of media, arts, and entertainment occupations).We exclude health fields, as they are STEM-adjacent.

self esteem or other abilities. Taking these point estimates at face value, about half of the under-confident children who do not pursue STEM careers switch into other high-education occupations, while the rest pursue other work. However, our 95-percent confidence intervals include estimates suggesting that under-confident children are up to 3.9pp less likely or up to 9.0pp more likely to work in other high-education occupations than their peers.

Next, we consider respondents' earnings. Our regression results are imprecisely estimated, but they broadly suggest that higher math confidence is associated with higher earnings later in life (Table 5, columns 5 and 6). While our binary measures of over- and under-confidence are not significantly associated with earnings (Panel A), a 1sd increase in the degree of childhood confidence is associated with 5.9 percent (se = 2.9 percent) higher earnings in adulthood. This gap is large relative to the association between childhood math scores and adult earnings: increasing test scores by one decile is associated with 7 percent higher earnings on average (Appendix Table A12, column 11).

Finally, we consider unemployment. Our regressions suggest that higher confidence may be associated with lower unemployment risk (Table 5, columns 7 and 8). Again, our binary indicators for over- and under-confidence are not significantly associated with unemployment (Panel A), but a 1sd increase in childhood confidence is associated with a 2.3pp (se = 0.9pp) lower likelihood of having been unemployed in the previous year. This gap is large relative to the association between childhood math scores and unemployment: increasing test scores by 10 percentiles is associated with 1.6pp lower unemployment risk on average (Appendix Table A12, column 12).

While most of our results are quite stable – both in magnitude and precision – to the many robustness tests we run in Section 7, our results for earnings and unemployment should be interpreted with caution. They are only statistically significant when using our more continuous measure of biased beliefs, which is more vulnerable to measurement error, and we show in Section 7 below that they are not robust to using measures of confidence that minimize measurement error by using data from two waves of the CDS. That said, they are suggestive and are consistent with our other findings on the long-term links between childhood confidence and later-life outcomes.

7 Robustness

In this section, we show that our main results are robust to controlling for a range of possible confounding variables and to many alternate definitions of our key measures of biased beliefs.

7.1 Key confounders: Personality, adult investment, and school quality

First, we show that math over- and under-confidence predict long-run outcomes beyond their correlation with (1) more commonly-studied personality traits, (2) parent and teacher beliefs and investment, and (3) and elementary school quality. We do not control for these variables in our main specifications because they are likely jointly determined with math confidence, but they may confound the links we estimate. See Appendix E for more details on data used in this section.

Section 1 of Appendix Table A13 adds controls for children's Big-Five personality traits: conscientiousness, agreeableness, neuroticism, openness, and extroversion. The CDS did not administer standard psychometric scales to identify the Big-Five traits among children, so we construct these measures from caregivers' reports of child behavior (see Appendices B and E for details.) These traits could confound the long-term associations that we observe: other work shows that Big-Five personality traits correlate with contemporaneous educational and employment outcomes (e.g. Almlund et al. 2011; Heckman et al. 2019), and we find some correlations between these common personality traits and measures of over- and under-confidence in our sample (Appendix Table A14). However, our estimates of the links between over- and under-confidence and long-run outcomes are broadly robust to controlling for them.[39]

Next, we add controls to our main specification for parent investments, like reading or doing homework with the child, teacher ratings of children's academic, social, and physical competence, and the educational attainment that parents and teachers predict for the child. Note that we only observe teacher perceptions for 20-34 percent of the sample. Teacher and parent beliefs and investments do correlate with children's beliefs in math in our sample (Appendix Table A15). If these adults' investments affect children's later-life success, they may drive the links between childhood

[39] Appendix Table A20 shows the coefficients on the personality measures in this regression; they correlate with long-run outcomes in expected ways (Almlund et al., 2011).

math confidence and later outcomes that we observe (Papageorge et al., 2018; Dizon-Ross, 2019). However, Section 2 of Appendix Table A13 shows that children's over- and under-confidence continue to predict long-run outcomes in similar ways when we add controls for adult perceptions and investment to our main regressions.

Finally, we show that our results are robust to controlling for the quality of school a child attended when we first observe their biased beliefs in math. If children assess their own ability relative to their peers, not the national distribution, school quality may shape children's self-assessments in math; over-confident children could just be those with low-performing peers, for example. However, these patterns would tend to bias our results *towards zero*, since later-life outcomes may be worse for children from lower-performing schools. We use restricted data from the CDS to match students with data on the percent of students at their school who qualified for free or reduced-price lunch (a proxy for income), the average student-teacher ratio at their school (a proxy for educational inputs), and levels and trends of their school's mean achievement levels in math and reading. Reassuringly, our results do not change meaningfully when we control for school quality (Appendix Table A13, Section 3).

7.2 Alternate definitions of biased beliefs

Next, we show that our results are robust to a range of alternate measures of biased childhood beliefs in math. Appendix F describes each of these alternate measures in more detail. None of these changes affects our main conclusions: that over- and under-confidence strongly and meaningfully predict long-term education and working in STEM.

Redefining over- and under- confidence: We first redefine our binary measures of over- and under-confidence by altering the CDS math score and self-report cutoffs on which they rely, making those designations more or less strict than our main measures. Second, we construct more data-driven measures of over- and under-confidence—what we refer to as the relative confidence measures—that identify over- and under-confident children as those in the tails of the distribution of math scores at each self-reported ability level. Finally, a third class of binary over- and under-confidence measures marks a child as over-confident if the degrees of confidence measure is greater than 2 and under-confident if it is less than -2.

Redefining degrees of confidence: Next, we also test robustness to the key design choice in our more continuous measure of confidence: how we map self-assessed ability and observed scores to the same scale. Our main measure assumes that children with accurate beliefs would report the numbered bin from 1 to 7 in which their CDS score falls when test score percentiles are uniformly distributed across 7 bins (i.e. each bin covers about 14 percentiles). We test robustness to two other transformations: the first assumes that children should have reported the bin from 1-7 in which their test score would fall if they had the CDS' empirical self-assessment distribution in mind, and the second instead differences children's percentiles of self-assessed ability and demonstrated ability. Each of these is converted to standard deviation units to facilitate comparisons.

Measurement error: To reduce the likelihood that our results are driven by measurement error, we also construct alternate confidence measures using testing and self-assessment data from two waves of the CDS for the 60 percent of children with multiple measures. We take two approaches to redefining our indicators for over- and under-confidence. In the first, we average children's test scores and self-reported ability over two waves and then apply our standard cutoff rules to these average scores and self-reports. In the second, we calculate indicators for being over- or under-confident separately in each of two waves and then average these indicators. We use the same logic in defining multi-wave versions of the more continuous confidence measure.

Results: Appendix Figures A4-A16 present specification charts showing results for each of our main outcomes of interest using these alternate measures of biased beliefs.[40] For simplicity, Appendix Table A16 presents a subset of these results: we iterate through alternate definitions of biased beliefs for each outcome, always using the control variables from our preferred specification. Panel A shows the results for over- and under-confidence, and Panel B shows the results for our more continuous measure of biased beliefs. Most coefficients that are statistically significant

[40]Besides testing alternate confidence measures, the specification charts also show that our main results are robust to dropping children in the lowest and highest math score deciles from our sample. Across all confidence measures, children at the upper (lower) tail of the score distribution are mechanically most likely to be identified as under-confident (over-confident).

in our main results are remarkably stable, leaving our conclusions unchanged. The only exceptions are our results for earnings and unemployment, which disappear when we use the more continuous measure of biased beliefs based on two waves of the CDS.

8 Snowballing investment or persistent over- and under-confidence?

Childhood over- and under-confidence in math are associated with gaps in key educational and employment outcomes down the line, from adolescent math performance to career choices in young adulthood. As we outlined in Section 2, these confidence gaps could arise if over- and under-confidence shape children's own investment decisions or those of parents, teachers, or potential employers. In this section, we explore the dynamic patterns through which these confidence gaps open up and persist. On one hand, math confidence could produce investment gaps in childhood that in turn snowball through children's later education and occupational choices. On the other hand, childhood over- and under-confidence in math may persist into adulthood and directly affect choices and performance at each stage of life, conditional on past achievement.

This section explores whether biased beliefs persist into adulthood, and whether gaps in later-life outcomes can be fully accounted for by the links between confidence and intermediate investments that we observe.

8.1 The persistence of childhood confidence in math

While Table 1, discussed in Section 3.2, shows that over- and under-confidence in math persist from childhood into adolescence, we also find that childhood biased beliefs persist even until we last observe respondents in the TAS at ages 18 through 27 (Table 6). This persistence is a necessary condition for children's biased beliefs to have direct behavioral effects on their educational and career choices as they age. We use the wealth of questions in the TAS to construct measures of young adults' confidence in math and reading, generalized academic confidence, career confidence, and general confidence.[41] See Appendix B for more detail each of these measures.

[41] Unlike our measures of biased beliefs from the CDS, these TAS confidence variables are not paired with measures of demonstrated ability in adulthood. However, the ideal regressions would test the links between childhood over- and under-confidence and *biases* in adult confidence, so as to avoid conflating the persistence of biased beliefs with the links between childhood confidence and adult achievement. We approximate this ideal by controlling for *adolescent* math and reading

First, we calculate an index of adult math confidence as the mean of standardized ratings of how good respondents think they would be in a job requiring math or technology. By this metric, childhood math confidence strongly persists into adulthood. Respondents who were over-confident in math as children score about 0.26sd (se = 0.06sd) higher in math confidence as adults than others with comparable childhood test scores, while under-confident children score about 0.25sd (se = 0.05sd) lower (Table 6, columns 1 and 2, Panel A). Likewise, a 1sd increase in our more continuous measure of childhood math confidence predicts 0.17sd (se = 0.02sd) higher math confidence as an adult (Panel B). In contrast, children who were under-confident in math score about 0.17sd (se = 0.05sd) higher in adult reading confidence—measured by standardizing subjects' ratings of how good they would be in a job requiring them to read and write a lot—than others with comparable childhood test scores (Table 6, columns 3 and 4). This pattern may arise because under-confident children are less likely to work in STEM occupations, making them more likely to have a job requiring reading and writing.

Next, childhood over- and under-confidence in math predict gaps in general academic confidence and career confidence in adulthood (Table 6, columns 5-8). Generalized academic confidence captures respondents' beliefs in their skill at solving problems, thinking logically, listening, and teaching others, and career confidence captures respondents' belief that they can attain and succeed in their dream job. Children who are over-confident in math score about 0.08sd (se = 0.05sd) higher in adult academic confidence and 0.11sd (se = 0.05sd) higher in adult career confidence than peers with comparable childhood test scores. Similarly, a 1sd increase in childhood math confidence predicts a 0.04sd (se = 0.02) increase in adult academic confidence and a 0.05sd (se = 0.02) increase in adult career confidence. While it is unsurprising that adult math, academic, and career confidence are correlated, it is reassuring that the links between continuous childhood and adult math confidence are 3-4 times as large as those with these other forms of adult confidence.

However, there are no significant relationships between childhood math confidence and a measure of adult general confidence (Table 6, columns 9-10), which captures respondents' conviction

scores, digit span scores, and general confidence as proxies for adult ability.

in their ability to lead and supervise, their independence and decisiveness, and their life's direction. Since these regressions control for childhood and adolescent general confidence, they suggest that while general confidence correlates with math confidence in childhood, childhood math confidence is not significantly linked with the *evolution* of general confidence as respondents age.[42]

In sum, childhood over- and under-confidence in math persist through childhood and into young adulthood as confidence gaps across academic domains and in one's career. If these biased beliefs directly affect respondents' educational or employment success in adulthood, this persistence may be a key factor in the long-term economic associations that we observe.

8.2 Gaps in intermediate outcomes do not fully explain results

Despite the persistence of childhood confidence, the links we observe between childhood biased beliefs and later-life outcomes could still be fully explained by gaps in intermediate educational investments. In Figure 2, we explore the role of past investment by estimating the marginal relationships between childhood biased beliefs and later-life outcomes, conditional on all intermediate, observable outcomes along the chronological chain of education and entry into the labor market. We then compare these results to those from our baseline specification. If childhood biased beliefs continue to predict long-run gaps conditional on intermediate outcomes, these remaining gaps may be related to contemporaneous adult confidence. Of course, this analysis is imperfect, especially since we cannot control for all intermediate investments.

Figure 2 reproduces our baseline estimates (Tables 3, 4, and 5, even-numbered columns) for math over- and under-confidence in darker blue, while the lighter blue points present our estimates with controls for all outcomes that precede the outcome of interest. In particular, we re-examine educational outcomes through college holding fixed adolescent math and reading test scores, re-

[42] As additional evidence that our results capture links with math confidence, not general self esteem or ability, we consider a set of placebo outcomes: individuals' relationship status, general mental health, social anxiety, alcohol consumption, and dangerous behavior as young adults (all from the TAS). We expect each of these outcomes to be affected by general self-esteem, but not by math over- and under-confidence specifically. Reassuringly, we generally find no relationships between biased beliefs in math and any of these placebo outcomes, except that math over-confidence predicts a lower likelihood of being in a romantic relationship (Appendix Table A21).

examine having a graduate degree and occupation choice holding fixed all previously-observed educational outcomes, and re-examine log earnings and unemployment history with controls for all educational outcomes and past occupation choices.

Many of the large confidence gaps we've observed in educational and employment outcomes persist when we condition on observable intermediate outcomes. Controlling for adolescent academic achievement does not change the relationship between childhood biased beliefs and any of our educational outcomes, and under-confidence remains half as predictive of working in STEM when we control for all educational outcomes, including whether respondents majored in STEM. Gaps in respondents' earnings fall by up to 60 percent when we condition on intermediate outcomes, though our standard errors remain large. The unemployment coefficients are largely unaffected when we add intermediate outcomes as controls.

Together with the persistence of math confidence into adulthood, these results suggest that over- and under-confidence may continue to directly affect economic outcomes as respondents age.

9 Conclusion

In this paper, we identify over- and under-confidence in math among a large sample of children. In doing so, we are the first to show that even children have markedly biased beliefs about their own math ability. These beliefs are distinct from Big-Five personality traits and general confidence. Girls are less confident in math than boys with the same test scores and general confidence, so gender stereotypes about math may shape ability perceptions even at young ages.

We then estimate striking associations between respondents' childhood over- and under-confidence in math and their educational and employment outcomes up to 22 years later, including comprehensive controls for children's demonstrated ability and family backgrounds. In the near term, under-confident children perform worse on the CDS math tests five years later, while over-confident children score higher. In the longer term, childhood math confidence significantly predicts key aspects of later education and work trajectories: whether respondents graduate from high-school and college, their college major and occupation choices, their earnings, and whether they experience unemployment. We do not observe similar associations with long-run outcomes for childhood confidence in reading, a puzzle that we leave for future work.

Our results suggest that biased beliefs about math ability in childhood may predict later-life outcomes both through accumulated differences in educational investments and by continuing to affect economic outcomes as respondents age. Childhood over- and under-confidence persist into adolescence and adulthood, and childhood confidence continues to broadly predict later-life outcomes, particularly in education, when we control for all observable educational and career investments along the chronological chain of education and labor-market entry.

While our results are not causal, they suggest that confidence in math may crucially shape the education we achieve and jobs we get, with effects possibly taking root as early as childhood. Our results provide key early evidence on the importance of math confidence, but they leave substantial room for future exploration. Besides re-examining the associations we estimate for math over- and under-confidence in an experimental setting, research should explore the mechanisms by which childhood math confidence affects later-life outcomes. For example, do less confident children perform worse later because they get less encouragement from teachers, or do they simply choose to exert less effort at school? Next, we've seen that high-achievers with low confidence are less likely to work in STEM jobs; do they fare worse in job interviews for those positions, or do they simply not apply? Finally, if future research verifies that confidence causally affects later-life outcomes, what interventions can close those gaps?

References

Alan, Sule and Seda Ertac, "Mitigating the gender gap in the willingness to compete: Evidence from a randomized field experiment," *Journal of the European Economic Association*, 2019, *17* (4), 1147–1185.

Almlund, Mathilde, Angela Lee Duckworth, James J Heckman, and Tim D Kautz, "Personality psychology and economics," *NBER Working Paper No. 16822*, February 2011.

Anaya, Lina, Frank Stafford, and Gema Zamarro, "Gender gaps in math performance, perceived mathematical ability and college STEM education: the role of parental occupation," *Education Economics*, September 2021, *0* (0), 1–16.

Ballen, Cissy J., Shima Salehi, and Sehoya Cotner, "Exams disadvantage women in introductory biology," *PLoS One*, 2017, *12* (10), e0186419. Publisher: Public Library of Science San Francisco, CA USA.

Bandura, Albert, *Social foundations of thought and action: A social cognitive theory.*, Englewood Cliffs, NJ: Prentice Hall., 1986.

Becker, Gary S., *Human capital: A theoretical and empirical analysis with special reference to education*, The University of Chicago Press, 1964.

Benjamin, Daniel J., "Errors in probabilistic reasoning and judgment biases," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 69–186.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Beliefs about gender," *American Economic Review*, 2019, *109* (3), 739–73.

Borjas, George J., "Immigrants, minorities, and labor market competition," *ILR Review*, 1987, *40* (3), 382–392.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, "Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures," *NBER Working Paper No. 28916*, 2021.

Bénabou, Roland and Jean Tirole, "Self-confidence and personal motivation," *The Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti, "Goals and gaps: Educational careers of immigrant children," *CReAM Discussion Paper Series 1812*, 2018, *Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.*

Castex, Gonzalo and Evgenia Kogan Dechter, "The changing roles of education and ability in wage determination," *Journal of Labor Economics*, 2014, *32* (4), 685–710.

Coffman, Katherine B., Manuela Collis, and Leena Kulkarni, "Stereotypes and belief updating," *Harvard Business School Working Paper No. 19-068*, January 2019.

Coffman, Katherine Baldiga, "Evidence on self-stereotyping and the contribution of ideas," *The Quarterly Journal of Economics*, November 2014, *129* (4), 1625–1660.

Cohodes, Sarah R. and Joshua S. Goodman, "Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy," *American Economic Journal: Applied Economics*, 2014, *6* (4), 251–85.

de Araujo, Pedro and Stephen Lagos, "Self-esteem, education, and wages revisited," *Journal of Economic Psychology*, 2013, *34*, 120–132.

DeMars, Christine E., Bozhidar M. Bashkov, and Alan B. Socha, "The role of gender in test-taking motivation under low-stakes conditions," *Research & Practice in Assessment*, 2013, *8*, 69–82.

Devine, Amy, Kayleigh Fawcett, Dénes Szucs, and Ann Dowker, "Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety," *Behavioral and Brain Functions*, 2012, *8* (1), 1–9.

Diamond, Rebecca and Petra Persson, "The long-term consequences of teacher discretion in grading of high-stakes tests," *NBER Working Paper No. 22207*, 2017.

Dizon-Ross, Rebecca, "Parents' beliefs about their children's academic ability: Implications for educational investments," *American Economic Review*, August 2019, *109* (8), 2728–2765.

Dohmen, Thomas and Armin Falk, "Performance pay and multidimensional sorting: Productivity, preferences, and gender," *American Economic Review*, 2011, *101* (2), 556–590.

Drago, Francesco, "Self-esteem and earnings," *Journal of Economic Psychology*, June 2011, *32* (3), 480–488.

Duffy, Denise and Narayan Sastry, "Achievement tests in the Panel Study of Income Dynamics Child Development Supplement," *PSID Technical Series Paper #14-02*, 2014, p. 29.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, and Jeanne Brooks-Gunn, "School readiness and later achievement.," *Developmental psychology*, 2007, *43* (6), 1428.

Eccles, Jacquelynne, C. Midgley, and T. Adler, "Grade-related changes in the school environment: Effects on achievement motivation," in "Advances in Motivation and Achievement: The Development of Achievement Motivation," Vol. 3 January 1984, pp. 282–331.

Edmonds, Grant W., Lewis R. Goldberg, Sarah E. Hampson, and Maureen Barckley, "Personality stability from childhood to midlife: Relating teachers' assessments in elementary school to observer-and self-ratings 40 years later," *Journal of Research in Personality*, 2013, *47* (5), 505–513.

Eil, David and Justin M. Rao, "The good news-bad news effect: asymmetric processing of objective information about yourself," *American Economic Journal: Microeconomics*, 2011, *3* (2), 114–38.

Erturan, Selin and Brenda Jansen, "An investigation of boys' and girls' emotional experience of math, their math performance, and the relation between these variables," *European Journal of Psychology of Education*, 2015, *30* (4), 421–435.

Fahle, Erin M., Belen Chavez, Demitra Kalogrides, Benjamin R. Shear, Sean F. Reardon, and Andrew D. Ho, "Stanford Education Data Archive: Technical Documentation (Version 4.1).," Technical Report 2021. Retrieved from http://purl.stanford.edu/db586ns4974.

Falco, Lia D., Jessica J. Summers, and Sheri Bauman, "Encouraging mathematics participation through improved self-efficacy: A school counseling outcomes study," *Educational Research and Evaluation*, 2010, *16* (6), 529–549.

Fryer, Roland G., "The "pupil" Factory: Specialization and the production of human capital in schools," *American Economic Review*, March 2018, *108* (3), 616–656.

Gneezy, Uri, John A. List, Jeffrey A. Livingston, Xiangdong Qin, Sally Sadoff, and Yang Xu, "Measuring success in education: the role of effort on the test itself," *American Economic Review: Insights*, 2019, *1* (3), 291–308.

Goodman, Joshua, "The labor of division: Returns to compulsory high school math coursework," *Journal of Labor Economics*, 2019, *37* (4), 1141–1182.

Greico, Elizabeth M. and Rachel C. Cassidy, "Overview of race and Hispanic origin: Census 2000 brief," Technical Report C2KBR/01-1, US Census Bureau, US Department of Commerce 2001.

Hampson, Sarah E. and Lewis R. Goldberg, "A first large cohort study of personality trait stability over the 40 years between elementary school and midlife.," *Journal of Personality and Social Psychology*, 2006, *91* (4), 763.

Heckman, James J., Tomáš Jagelka, and Timothy D. Kautz, "Some contributions of economics to the study of personality," *NBER Working Paper No. 26459*, 2019.

Herbert, J. and D. Stipek, "The emergence of gender differences in children's perceptions of their academic competence," *Journal of Applied Developmental Psychology*, 2005, *26* (3), 276–295.

Hicks, Peggy and Larry M. Bolen, "Review of the Woodcock-Johnson Psycho-Educational Battery-Revised," *Journal of School Psychology*, 1996, *34* (1), 93–102.

Jamieson, Jeremy P., Wendy Berry Mendes, and Matthew K. Nock, "Improving acute stress responses: The power of reappraisal," *Current Directions in Psychological Science*, 2013, *22* (1), 51–56.

Jussim, Lee and Kent D Harber, "Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies," *Personality and Social Psychology Review*, 2005, *9* (2), 131–155.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Langford, Joe and Pauline Rose Clance, "The imposter phenomenon: Recent research findings regarding dynamics, personality and family patterns and their implications for treatment," *Psychotherapy: Theory, research, practice, training*, 1993, *30* (3), 495.

List, John A., Ragan Petrie, and Anya Samek, "How experiments with children inform economics," *NBER Working Paper No. 28825*, 2021.

Mattarella-Micke, Andrew, Jill Mateo, Megan N. Kozak, Katherine Foster, and Sian L. Beilock, "Choke or thrive? The relation between salivary cortisol and math performance depends on individual differences in working memory and math-anxiety.," *Emotion*, 2011, *11* (4), 1000.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

Mobius, Markus M. and Tanya S. Rosenblat, "Why beauty matters," *American Economic Review*, March 2006, *96* (1), 222–235.

Moore, Don A. and Paul J. Healy, "The trouble with overconfidence.," *Psychological Review*, 2008, *115* (2), 502.

Murnane, Richard J., John B. Willett, M. Jay Braatz, and Yves Duhaldeborde, "Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY," *Economics of Education Review*, 2001, *20* (4), 311–320.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat, "Managing self-confidence," *Working paper*, 2014.

Owen, Stephanie, "College field specialization and beliefs about relative performance: An experimental intervention to understand gender gaps in STEM," *Working Paper*, 2020.

Papageorge, Nicholas W., Seth Gershenson, and Kyung Min Kang, "Teacher expectations matter," *NBER Working Paper No. 25255*, 2018.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Sakulku, Jaruwan, "The impostor phenomenon," *The Journal of Behavioral Science*, 2011, *6* (1), 75–97.

Schwardmann, Peter and Joel van der Weele, "Deception and self-deception," *Nature Human Behaviour*, 2019, *3* (10), 1055–1061.

Segal, Carmit, “Working when no one is watching: Motivation, test scores, and economic success,” *Management Science*, 2012, *58* (8), 1438–1457.

Sharot, Tali, Christoph W. Korn, and Raymond J. Dolan, “How unrealistic optimism is maintained in the face of reality,” *Nature Neuroscience*, 2011, *14* (11), 1475–1479.

Shastry, Gauri Kartini, Olga Shurchkov, and Lingjun Lotus Xia, “Luck or skill: How women and men react to noisy feedback,” *Journal of Behavioral and Experimental Economics*, October 2020, *88*, 101592.

Shull-Senn, Shannon, Michael Weatherly, Sandra Kanouse Morgan, and Sharon Bradley-Johnson, “Stability reliability for elementary-age students on the Woodcock-Johnson Psychoeducational Battery—Revised (Achievement section) and the Kaufman Test of Educational Achievement,” *Psychology in the Schools*, 1995, *32* (2), 86–92.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge, “What are we weighting for?,” *Journal of Human Resources*, March 2015, *50* (2), 301–316.

Soto, Christopher J., “The Little Six personality dimensions from early childhood to early adulthood: Mean-level age and gender differences in parents’ reports,” *Journal of Personality*, August 2016, *84* (4), 409–422.

Stinnett, Terry A., J. Michael Havey, and Judy Oehler-Stinnett, “Current test usage by practicing school psychologists: A national survey,” *Journal of Psychoeducational Assessment*, 1994, *12* (4), 331–350.

Stroud, Laura R., Peter Salovey, and Elissa S. Epel, “Sex differences in stress responses: social rejection versus achievement stress,” *Biological psychiatry*, 2002, *52* (4), 318–327.

Taylor, Shelley E and Jonathon D Brown, “Illusion and well-being: A social psychological perspective on mental health,” *Psychological Bulletin*, 1988, *103* (2), 193–210.

Usher, Ellen L. and Frank Pajares, "Sources of academic and self-regulatory efficacy beliefs of entering middle school students," *Contemporary Educational Psychology*, April 2006, *31* (2), 125–141.

van den Akker, Alithe, Maja Deković, J.J. Asscher, and Peter Prinzie, "Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting," *Journal of Personality and Social Psychology*, August 2014, *107.*

Waddell, Glen R., "Labor-market consequences of poor attitude and low self-esteem in youth," *Economic Inquiry*, 2006, *44* (1), 69–97.

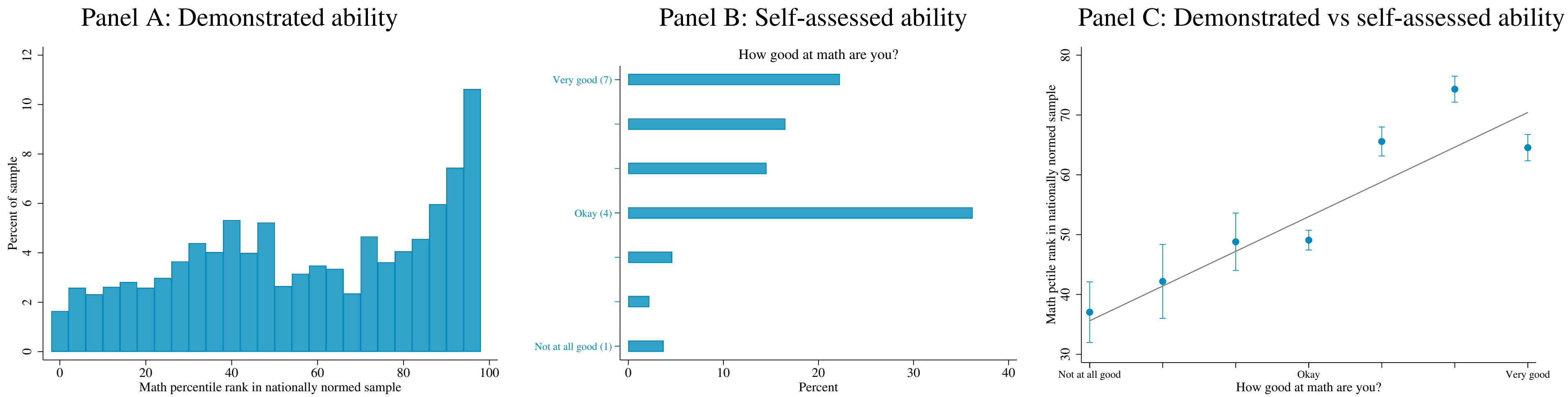
Wang, Shengnan, Christine M. Rubie-Davies, and Kane Meissel, "A systematic review of the teacher expectation literature over the past 30 years," *Educational Research and Evaluation*, 2018, *24* (3-5), 124–179.

Weekes, Nicole, Richard Lewis, Falgooni Patel, Jared Garrison-Jakel, Dale E. Berger, and Sonia J. Lupien, "Examination stress as an ecological inducer of cortisol and psychological responses to stress in undergraduate students," *Stress*, 2006, *9* (4), 199–206.

Wigfield, Allan and Jacquelynne S. Eccles, "Expectancy–value theory of achievement motivation," *Contemporary Educational Psychology*, January 2000, *25* (1), 68–81.

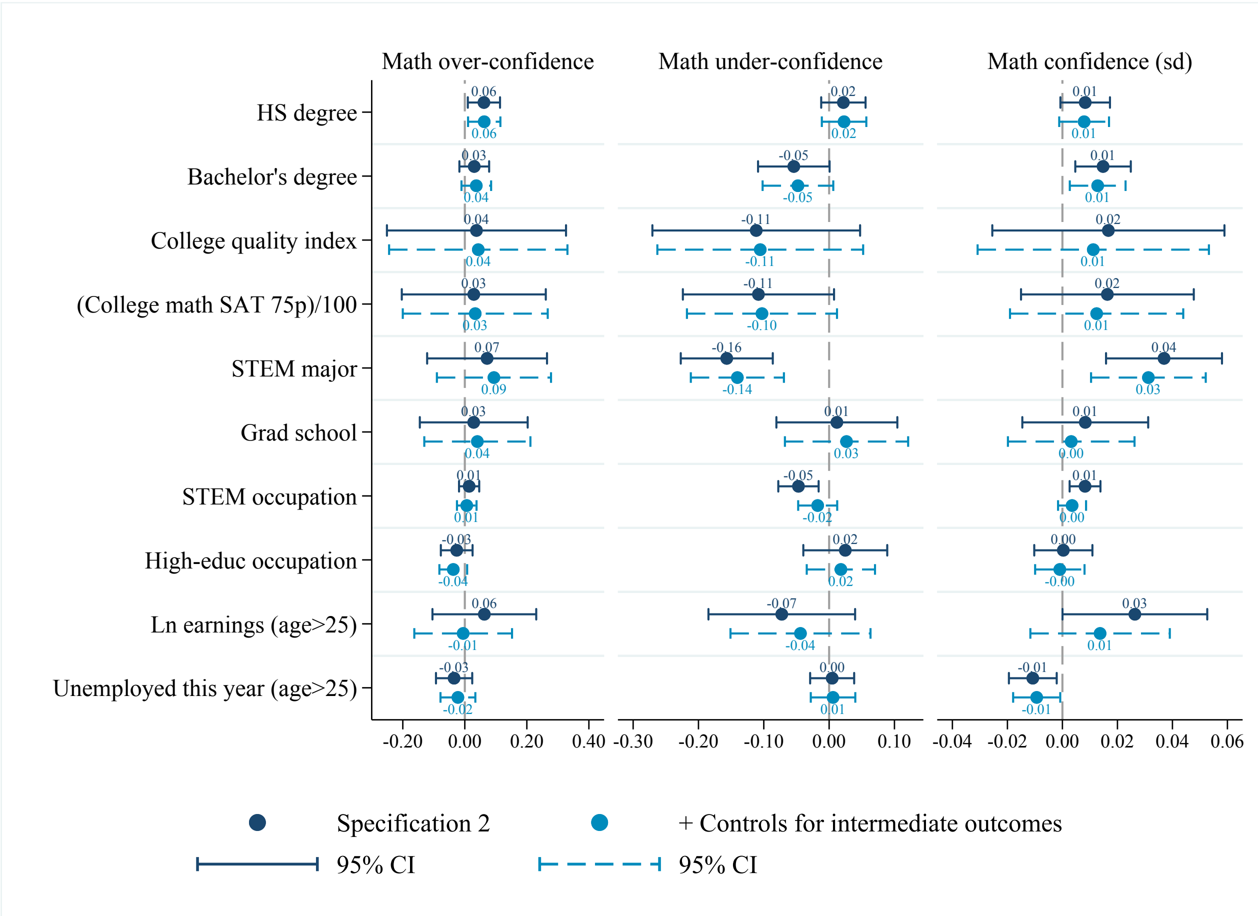
Zimmermann, Florian, "The dynamics of motivated beliefs," *American Economic Review*, 2020, *110* (2), 337–361.

Figure 1: Distributions of self-assessed and demonstrated ability



Note: We plot first-observed math test scores and self-assessments for the 2985 CDS respondents with at least one year of both measurements. We measure respondents' ability and self-beliefs in math at ages ranging from 8 to 19, though we observe the median child at 11 and more than 90% of children by age 13. Panel A plots the distribution of respondents' percentile ranks (calculated relative to a nationally-representative norming sample) on a portion of the Woodcock-Johnson Psycho-Educational Battery Revised (WJ-R) testing math reasoning and knowledge. Panel B plots the distribution of children's responses when asked to answer "How good at math are you?" on a scale from 1 (not at all good) to 7 (very good). Finally, Panel C plots the average math percentile rank within each category from 1 to 7 of children's self-reported ability in math.

Figure 2: Controlling for intermediate outcomes



Note: This figure plots the coefficient on over- or under-confidence in our baseline specification (2) and the same coefficient when we add controls for mediating factors. When the outcome is high school or college graduation, majoring in STEM, or college quality, we add controls for adolescent math and reading test scores. When the outcome is earning a graduate degree, we add controls for all previously-observed education outcomes. When the outcome is occupation choice, we add controls for all observed education outcomes: math and reading scores in adolescence, whether the respondent graduated from high school, college, or graduate school, the 75th percentile of the math SAT score distribution of the college he or she attended, and whether he or she majored in STEM. When the outcome is earnings or unemployment, we add controls for all observed educational outcomes and occupational choice.

Table 1: The persistence of math over- and under-confidence

	(1)	(2)
Panel A: Math over-confidence	0.118^{***}	0.116^{***}
	(0.031)	(0.032)
N	1747	
Sample mean	0.041	
Panel B: Math under-confidence	0.042^{*}	0.042^{*}
	(0.024)	(0.025)
N	1747	
Sample mean	0.063	
Panel C: Math confidence (SD units)	0.182^{***}	0.185^{***}
	(0.025)	(0.025)
N	1747	
Sample mean	0.000	
Basic controls:	✓	✓
Added background controls:		✓

Notes: This table regresses adolescent confidence outcomes on various definitions of childhood math confidence with various controls. Adolescent confidence is measured five years after the childhood measurement. In each row, the dependent variable is the adolescent measurement of the independent variable described. The measures of over- and under-confidence are our main binary measures. Our secondary measure of degrees of confidence takes on values from -6 to 6 and persistence of that variable is shown in the third row. The fourth row standardizes the degrees of confidence measure to have mean 0 and standard deviation 1, to facilitate ease of interpretation. All controls that are time-variant are observed in the same year as the confidence measures. Basic controls include child gender, race, decile fixed effects for math and reading test percentile scores, digit span test scores, a general confidence index, family taxable income and its square, parent education, quarter-of-birth fixed effects, year-of-birth fixed effects, age at which confidence was measured fixed effects, and state fixed effects. We also include fixed effects for adolescent test score deciles in math and reading. Added background controls are parents' rating of child health, indicators for receiving government transfers, household structure, parenting practices, parent occupation, and parent mental health and confidence measures. All controls are recoded to zero if missing and we include a missing indicator. Standard errors are clustered by family, and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table 2: Demographic predictors of over- and under-confidence

	Over-confidence	Under-confidence	Confidence (sd)
Demographic Characteristics			
Female	-0.027***	0.023**	-0.097***
	(0.01)	(0.01)	(0.03)
Black	0.014	0.015	0.031
	(0.02)	(0.02)	(0.04)
Hispanic	-0.038*	0.019	-0.037
	(0.02)	(0.03)	(0.07)
Asian or Native American	-0.021	0.024	-0.042
	(0.02)	(0.03)	(0.06)
Only child	0.005	0.016	0.008
	(0.02)	(0.02)	(0.05)
First child	0.015	0.038**	-0.012
	(0.02)	(0.02)	(0.04)
Second child	0.029*	0.004	0.055
	(0.02)	(0.02)	(0.04)
Father graduated high school	-0.011	-0.039	0.072
	(0.02)	(0.02)	(0.06)
Father has bachelors	0.014	-0.009	0.024
	(0.01)	(0.02)	(0.04)
Mother graduated high school	-0.019	-0.010	-0.013
	(0.02)	(0.02)	(0.05)
Mother has bachelors	-0.007	-0.025	-0.024
	(0.01)	(0.03)	(0.05)
Father works in STEM	0.004	-0.036	0.044
	(0.02)	(0.03)	(0.05)
Mother works in STEM	-0.010	-0.006	-0.009
	(0.01)	(0.02)	(0.04)
Father works in non-STEM high-educ	0.000	-0.008	0.022
	(0.01)	(0.03)	(0.05)
Mother works in non-STEM high-educ	-0.017	0.019	-0.062*
	(0.01)	(0.02)	(0.04)
Family taxable income (thous 2016 USD)	0.000	0.000	0.000
	(0.00)	(0.00)	(0.00)

Table 2: Demographic predictors of over- and under-confidence (continued)

	Over-confidence	Under-confidence	Confidence (sd)
Other Child Characteristics			
Child ever in gifted prog	0.026**	-0.087***	0.146***
	(0.01)	(0.02)	(0.03)
Child ever in special ed prog	0.007	-0.008	0.071
	(0.02)	(0.02)	(0.05)
Child has repeated grade	-0.016	-0.012	0.007
	(0.02)	(0.01)	(0.05)
Parent's rating of child health	-0.001	-0.013**	0.022
	(0.01)	(0.01)	(0.01)
School Quality Measures			
Percent FRPL	-0.027	0.058*	-0.074
	(0.03)	(0.03)	(0.08)
Student-teacher ratio	0.000	-0.000	0.002***
	(0.00)	(0.00)	(0.00)
Average math and reading achievement	-0.003	0.003	-0.015
	(0.00)	(0.00)	(0.01)
Difference btwn math and reading achievement	0.010	-0.004	0.030
	(0.01)	(0.02)	(0.04)
Cohort slope of average achievement	-0.036	0.079	-0.228
	(0.06)	(0.06)	(0.15)
Unable to link to NCES id	0.049*	0.013	0.109
	(0.03)	(0.03)	(0.07)
Other Child Ability Measures			
Digit span score	-0.000	-0.004**	0.009**
	(0.00)	(0.00)	(0.00)
General confidence	0.038***	-0.054***	0.211***
	(0.01)	(0.01)	(0.02)
Mean of dependent variable	0.085	0.121	0.000
N	2985	2985	2985
R-squared	0.21	0.21	0.57

Notes: Each column regresses a measure of childhood biased beliefs in math on child characteristics. In columns 1 and 2, the dependent variable is our main indicator for over-confidence or under-confidence, respectively. In column 3, the dependent variable is a linear measure of biased beliefs that ranges from -6 to 6, where negative values represent under-confidence, *which has been standardized to have mean 0 and standard deviation one in our sample*. All variables are taken from the first year in which we observe the child's confidence in math. Additional controls include fixed effects for math and reading test score deciles, birth year, birth quarter, state, and age at which confidence was measured fixed effects. The coefficients on the ability deciles are shown in Appendix Table A11. All controls are recoded to be zero if missing and the regressions include missing indicators for each variable (not shown). All variables are either continuous or binary indicators, except for child race and birth order. The omitted category for race is non-Hispanic whites, and the omitted category for birth order is any birth order higher than two. Standard errors are clustered by family. *, **, and *** indicate significance at the 10, 5, and 1 percent level, respectively.

Table 3: Childhood math confidence and medium-term educational achievement and attainment

Dependent variable:	Adolescent math scores		Adolescent reading scores		High school degree		College degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	2.637*	2.666*	-0.362	-0.286	0.057**	0.062**	0.027	0.031
	(1.468)	(1.496)	(1.381)	(1.385)	(0.026)	(0.026)	(0.024)	(0.024)
Under-confidence	-5.705***	-5.860***	0.353	0.162	0.015	0.022	-0.057**	-0.058**
	(1.482)	(1.497)	(1.439)	(1.452)	(0.017)	(0.017)	(0.028)	(0.028)
N	1747	1747	1745	1745	2714	2714	2725	2725
OC = -1*UC? *p-value*:	0.147	0.138	0.997	0.951	0.022	0.008	0.413	0.457
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	2.806***	2.827***	0.111	0.128	0.019*	0.018*	0.032***	0.033***
	(0.566)	(0.569)	(0.587)	(0.580)	(0.010)	(0.010)	(0.011)	(0.011)
N	1747	1747	1745	1745	2714	2714	2725	2725
Sample mean of dep. var.	50.808		48.231		0.876		0.297	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses educational achievement and attainment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table 4: Childhood math confidence and college quality, college major choice, and post-college schooling

Dependent variable:	College quality index		College's 75th pctile math SAT score		STEM Major		Graduate degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.067	0.037	6.244	3.133	0.067	0.076	0.035	0.032
	(0.145)	(0.148)	(12.112)	(11.829)	(0.096)	(0.097)	(0.083)	(0.087)
Under-confidence	-0.095	-0.127	-9.312	-11.312*	-0.158***	-0.162***	0.014	0.006
	(0.081)	(0.082)	(5.958)	(5.925)	(0.036)	(0.036)	(0.046)	(0.048)
N	1107	1107	1117	1117	736	736	810	810
OC = -1*UC? *p-value*:	0.866	0.601	0.819	0.537	0.365	0.405	0.607	0.704
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.044	0.041	4.198	3.631	0.077***	0.078***	0.023	0.022
	(0.047)	(0.046)	(3.460)	(3.417)	(0.023)	(0.023)	(0.025)	(0.025)
N	1107	1107	1117	1117	736	736	810	810
Sample mean of dep. var.	0.053		594.172		0.189		0.200	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses college outcomes outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table 5: Childhood math confidence and employment outcomes

Dependent variable:	Works in STEM		Non-STEM high-educ occ.		Ln(Earnings)		Unemployed this year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.011	0.014	-0.019	-0.025	0.045	0.064	-0.034	-0.035
	(0.016)	(0.017)	(0.025)	(0.026)	(0.085)	(0.085)	(0.030)	(0.030)
Under-confidence	-0.049***	-0.049***	0.029	0.026	-0.067	-0.075	0.006	0.005
	(0.016)	(0.016)	(0.033)	(0.033)	(0.056)	(0.057)	(0.017)	(0.017)
N	4592	4592	4592	4592	4423	4423	4975	4975
OC = -1*UC? *p-value*:	0.096	0.127	0.822	0.987	0.833	0.917	0.437	0.395
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.018***	0.018***	-0.001	0.001	0.049*	0.059**	-0.023**	-0.023**
	(0.006)	(0.006)	(0.011)	(0.012)	(0.028)	(0.029)	(0.009)	(0.009)
N	4592	4592	4592	4592	4423	4423	4975	4975
Sample mean of dep. var.	0.046		0.163		10.185		0.167	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses employment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Basic controls also include year fixed effects when the outcome is observed in a panel. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table 6: Childhood math confidence and young adult confidence outcomes

Dependent variable:	Math confidence		Reading confidence		Academic confidence		Career confidence		General confidence	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Independent variables are binary measures of over- and under-confidence										
Over-confidence	0.264***	0.260***	0.002	-0.002	0.090*	0.086*	0.103**	0.106**	0.056	0.056
	(0.058)	(0.058)	(0.067)	(0.067)	(0.047)	(0.048)	(0.050)	(0.050)	(0.043)	(0.043)
Under-confidence	-0.250***	-0.251***	0.159***	0.163***	-0.038	-0.041	-0.058	-0.055	0.002	0.000
	(0.047)	(0.048)	(0.053)	(0.054)	(0.033)	(0.033)	(0.038)	(0.038)	(0.031)	(0.031)
N	6632	6632	6634	6634	8096	8096	6265	6265	8050	8050
OC = -1*UC? *p-value*:	0.850	0.904	0.064	0.062	0.362	0.441	0.487	0.418	0.268	0.289
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units										
Confidence	0.167***	0.168***	-0.046*	-0.048*	0.037**	0.038**	0.055***	0.055***	0.020	0.022
	(0.021)	(0.021)	(0.026)	(0.026)	(0.017)	(0.017)	(0.018)	(0.018)	(0.016)	(0.016)
N	6632	6632	6634	6634	8096	8096	6265	6265	8050	8050
Sample mean of dep. var.	-0.000		-0.000		-0.000		0.009		0.000	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓		✓

Notes: This table regresses young adult confidence outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Basic controls also include year fixed effects when the outcome is observed in a panel. In this table, we add controls for adolescent test score deciles in math and reading, as well as adolescent general confidence and digit span scores in all specifications. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Online Appendix: Childhood Confidence, Schooling, and the Labor Market: Evidence from the PSID

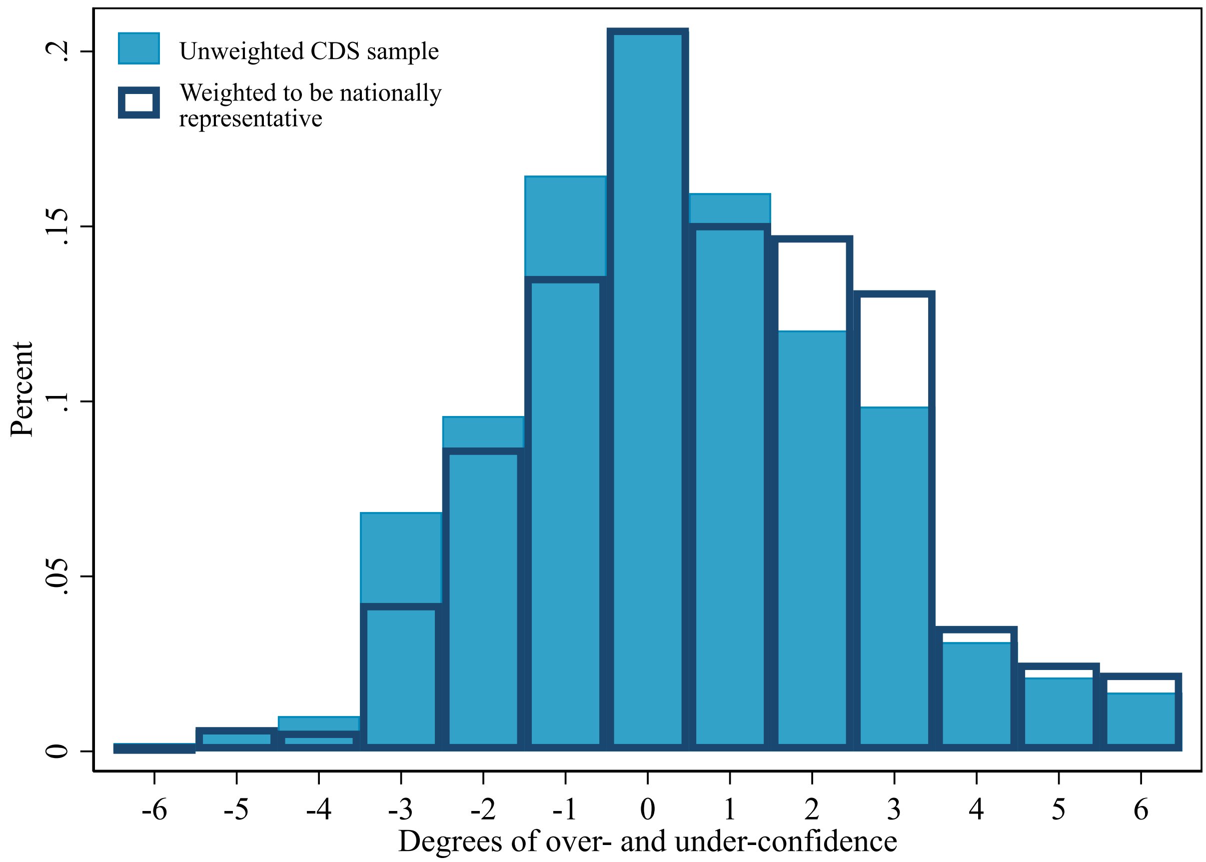
Lucy Page and Hannah Ruebeck

June 2021

Appendix A contains supplementary tables and figures. Appendix B describes the variables that make up each index used as an outcome or control variable in our main analysis and details index construction. Appendix C describes how our measures of childhood over- and under-confidence in math correlate with a range of children's attitudes towards math and school. Appendix D compares our results for over- versus under-confidence, and Appendix E provides more detail on our measures of childhood personality, teacher and parent beliefs and investment, and elementary/middle school quality. Finally, Appendix F outlines the alternate definitions of biased beliefs in math that we use in our robustness checks.

A Supplementary Figures and Tables

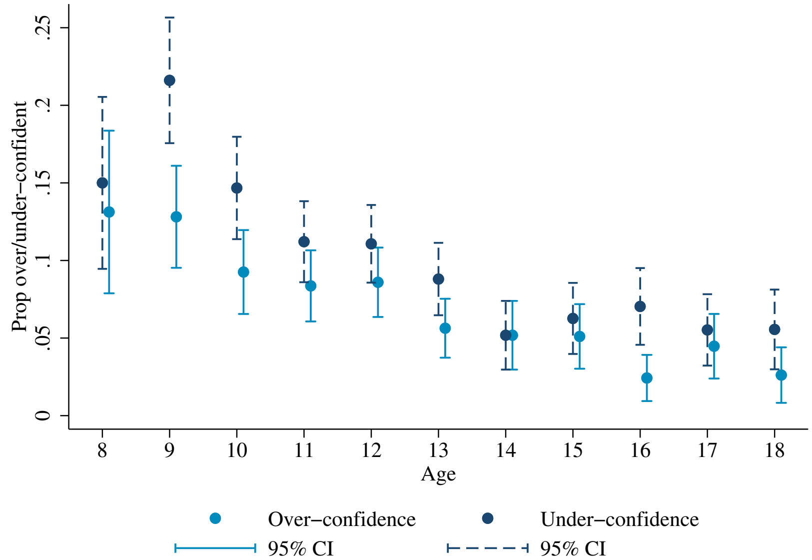
Figure A1: Distribution of the degrees of over- and under-confidence measure



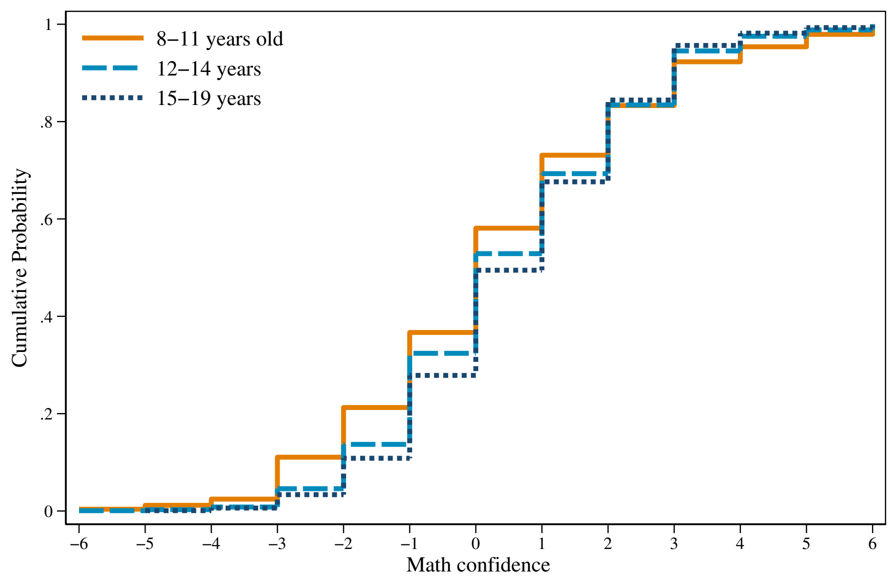
Note: This figure plots the distributions of our more continuous measure of confidence for our sample and when we use weights that make the sample nationally representative. The measure takes on value of integers from -6 (under-confident) to 6 (over-confident) and is calculated as the difference between children's self-assessed ability from 1-7 and the bin in which they should have placed themselves if they knew their score and the (uniform) national distribution of test scores. Weights are calculated using iterative proportional fitting (raking) on the original weights provided by the CDS so that our sample matches population shares in quintiles of income, in race categories, and in deciles of nationally-normed WJ-R math percentile scores.

Figure A2: Patterns in over- and under-confidence by age

Panel A: Proportion over- and under-confident (binary measure)

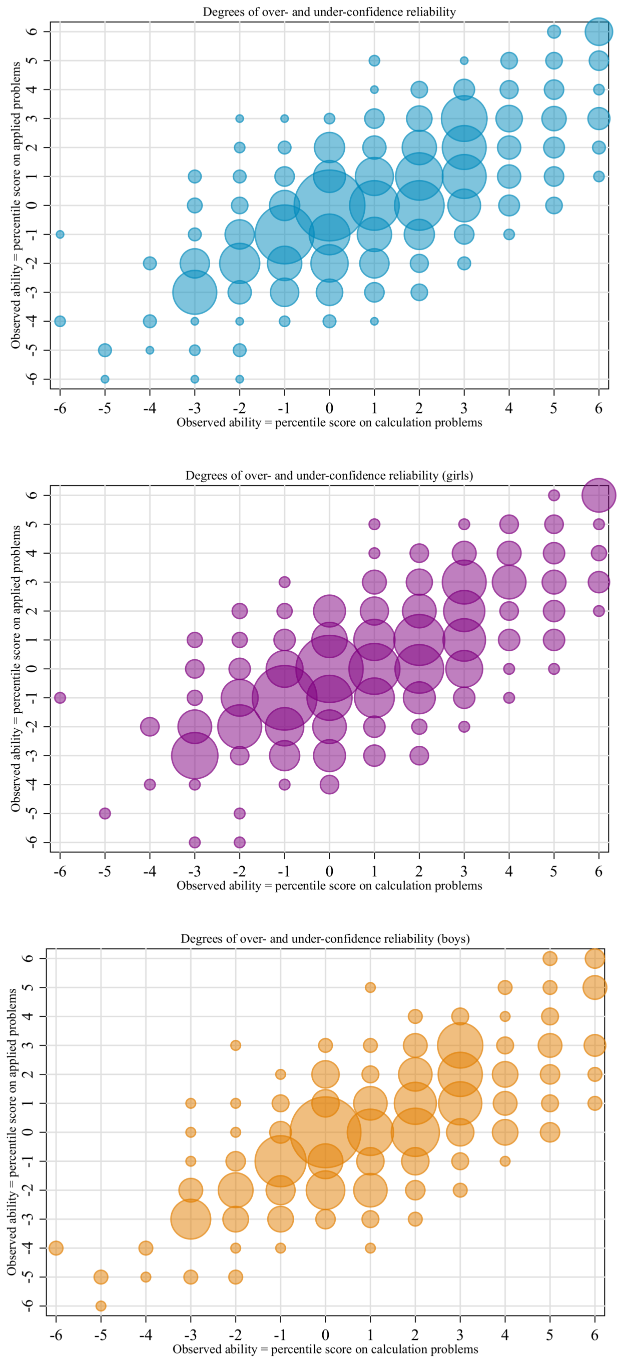


Panel B: CDF of degrees of over- and under-confidence



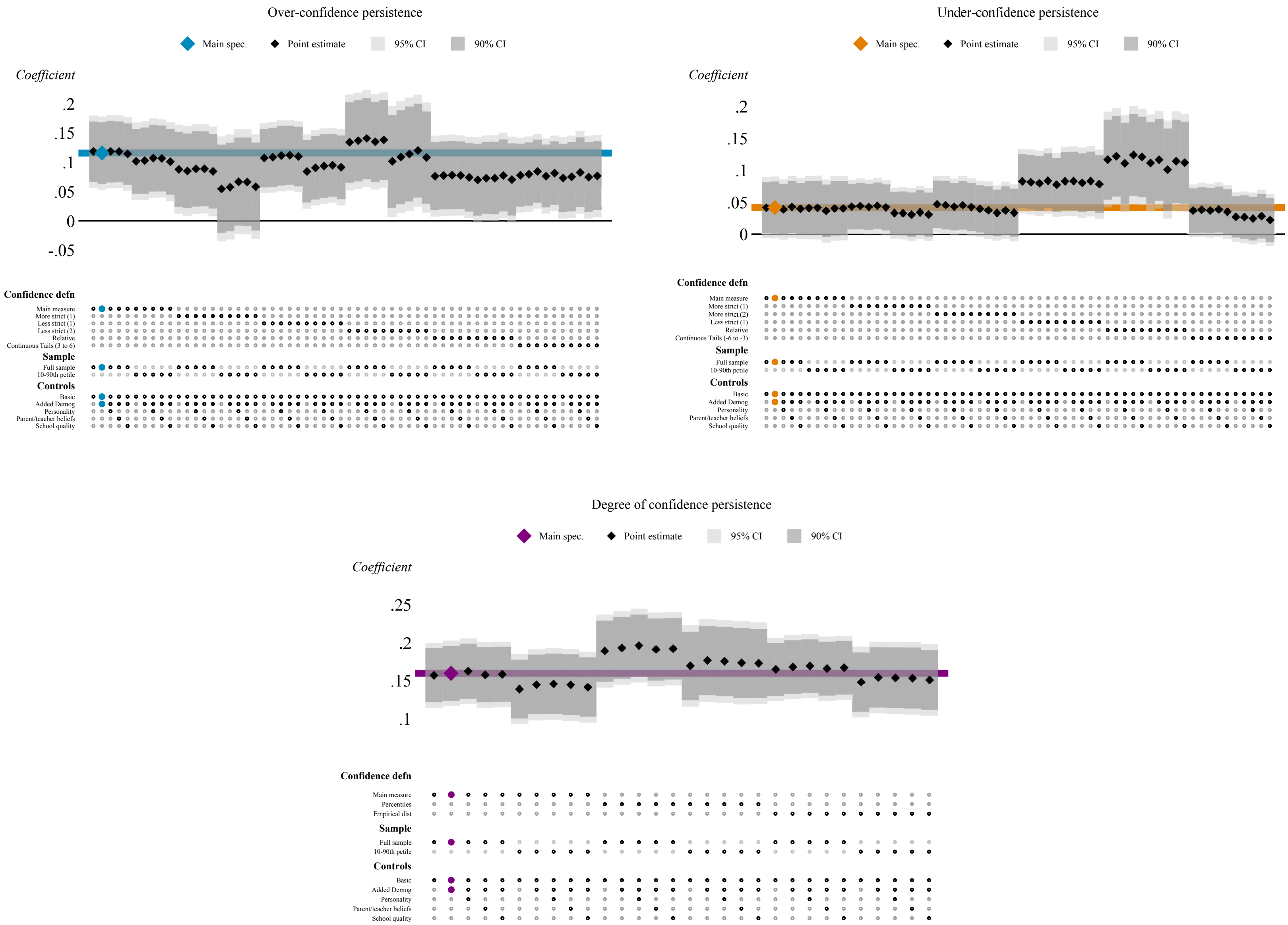
Note: Panel A plots the proportion of respondents that are over- and under-confident by age. Panel B plots the cumulative density function for the degrees of confidence measure, which takes on values from -6 to 6, separately for children in three age groups. We estimate these proportions using all observations of self-reported ability and test scores in math for our 2985 respondents, including two reports for the 60% of our sample with confidence measures in two CDS waves. In Panel A, we identify over- and under-confidence in math using gaps between children's self-reported math ability and their performance on the WJ-R math test administered in the CDS. In particular, we classify a respondent as under-confident if she scored above the 75th percentile on the WJ-R math assessments and ranked herself at 1-4 on the 7-point scale of math ability, or if she scored above the 50th percentile and ranked herself at 1-3. Similarly, we identify any respondent as over-confident in math if she scored below the 25th percentile and rated herself at 6 or 7 on the response scale, or if she scored below the 50th percentile and rated herself at 7. In Panel B, we measure biased beliefs as the difference between children's self-assessed ability (between 1 and 7) and the bin of the ability distribution in which they should have placed themselves if they had full information about the national distribution of scores and their place in it.

Figure A3: Differences in over- or under-confidence classification using 2 subtests of the WJ-R



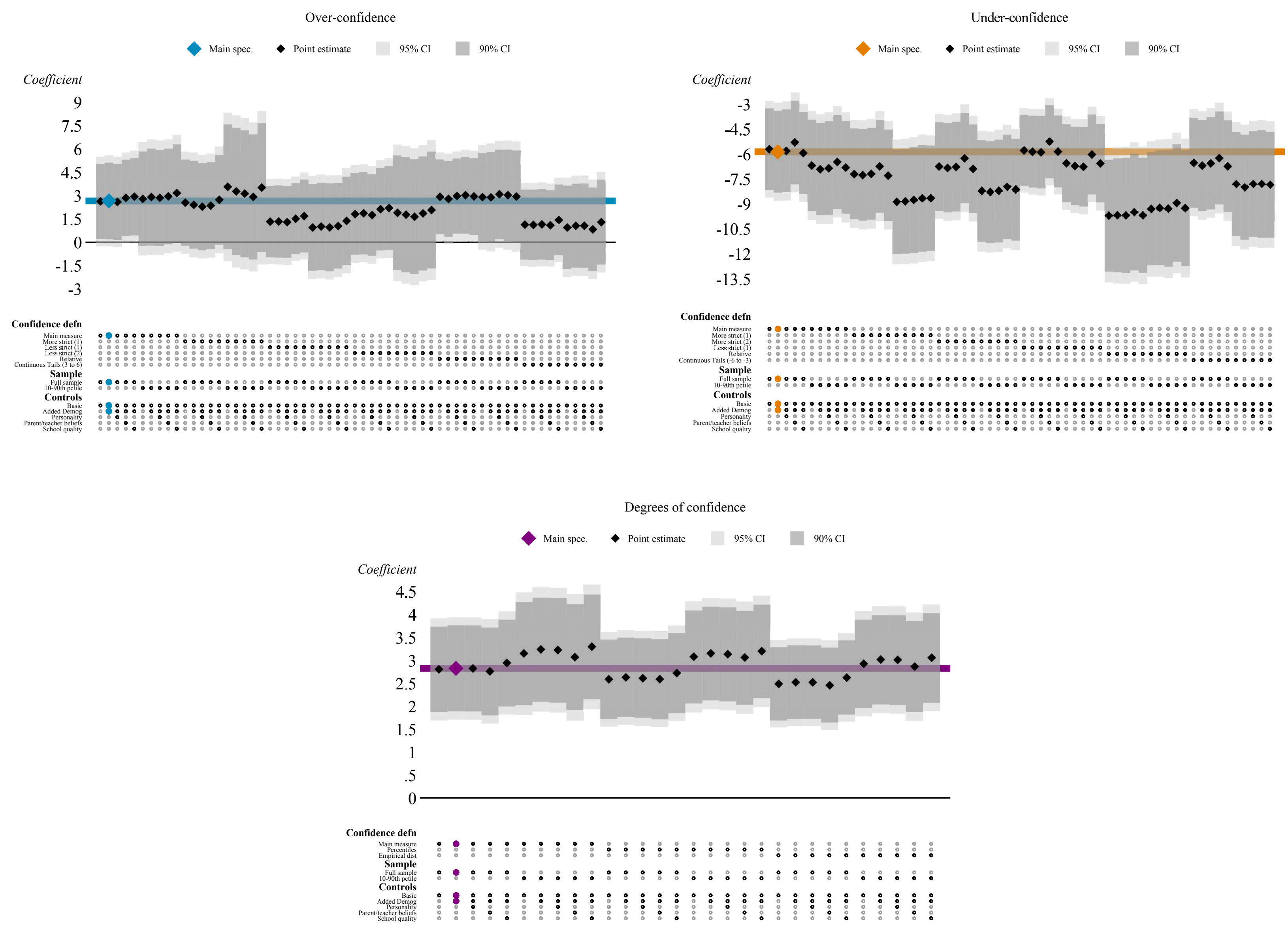
Note: This figure plots the joint distribution of children's degree of over- and under-confidence, which takes on values of integers from -6 (under-confident) to 6 (over-confident), when we use two different measures of demonstrated ability: percentile scores on the applied reasoning section of the WJ-R test (our main measure) and percentile scores on the calculation section of the WJ-R test (only administered in 1997).

Figure A4: Specification chart for persistence into adolescence



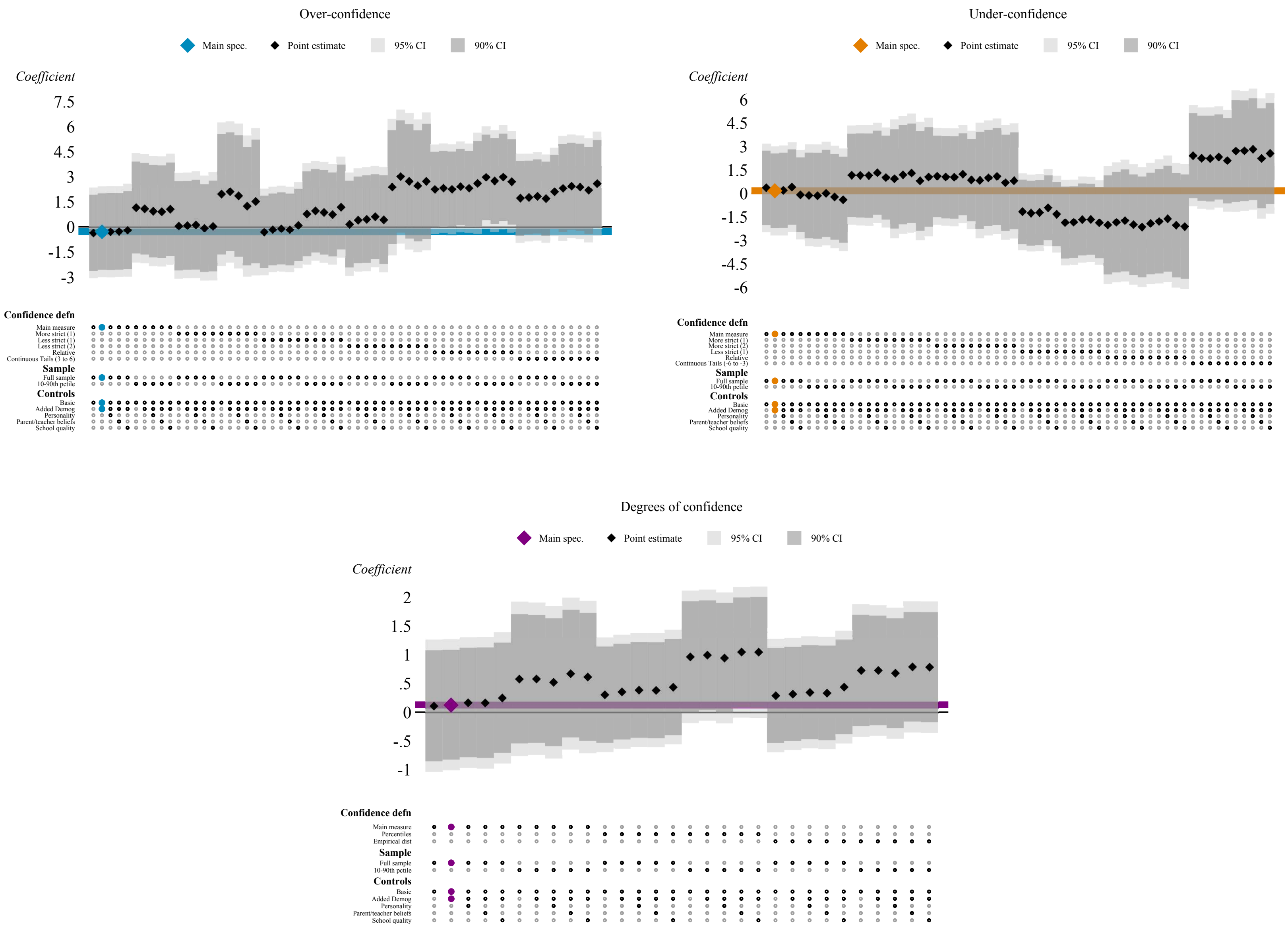
Note: This figure plots the coefficient of interest on either over- or under-confidence or our degrees of confidence measure for a large number of specification tests. The outcome is the same confidence measure as the main independent variable, observed 5 years later. We test the relationship between childhood biased beliefs and adolescent biased beliefs when we (a) change our definitions of over- and under-confidence and the degrees of confidence measure, (b) drop the bottom and top ten percent of the ability distribution, since those children are mechanically the most likely to be over- or under-confident, respectively, and (c) iterate through each of five sets of control variables. Appendix F describes each alternate confidence definition in detail.

Figure A5: Specification chart for adolescent math scores



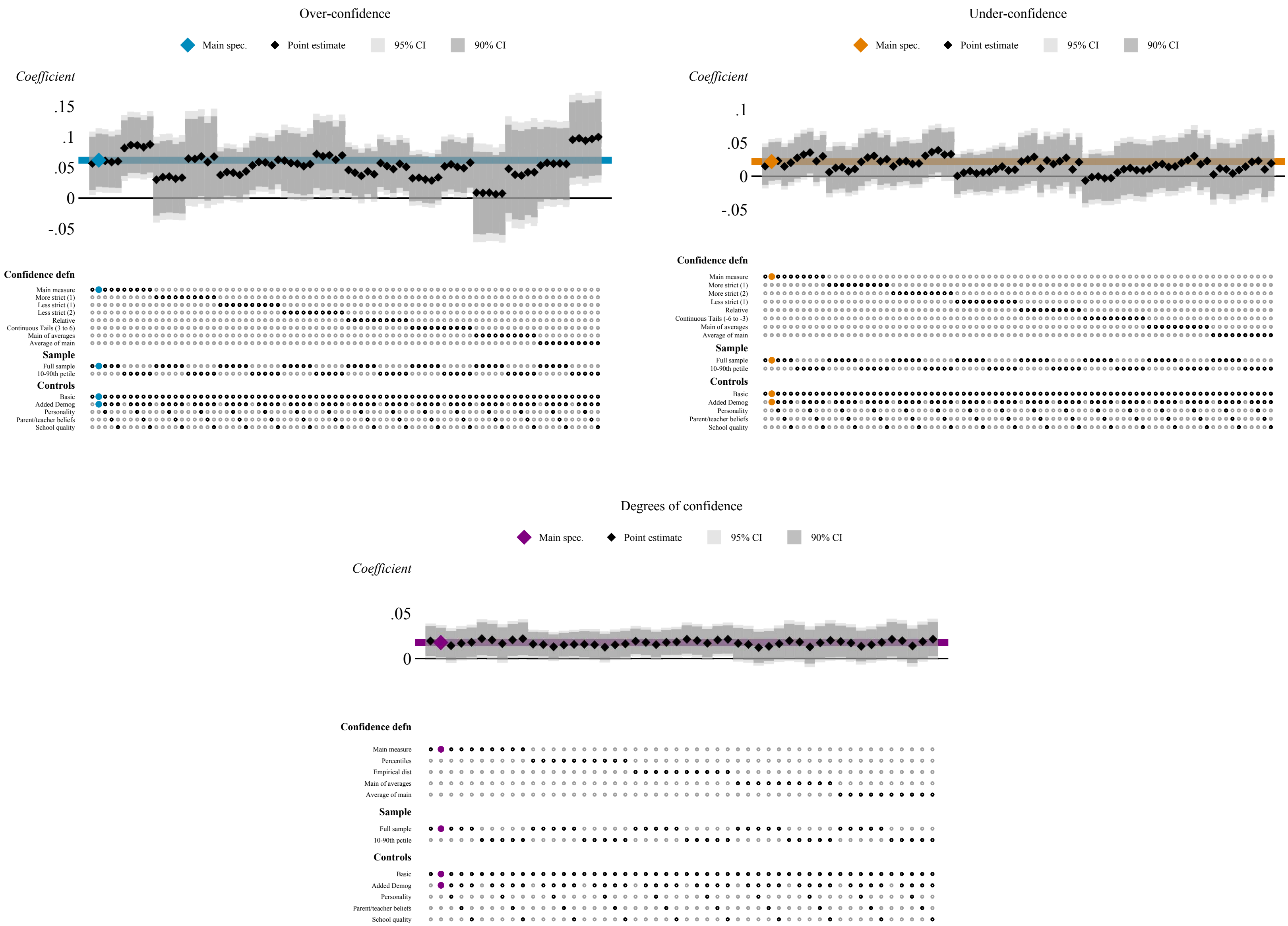
Note: This figure is analogous to Appendix Figure A4, but the outcome is adolescent math test scores.

Figure A6: Specification chart for adolescent reading scores



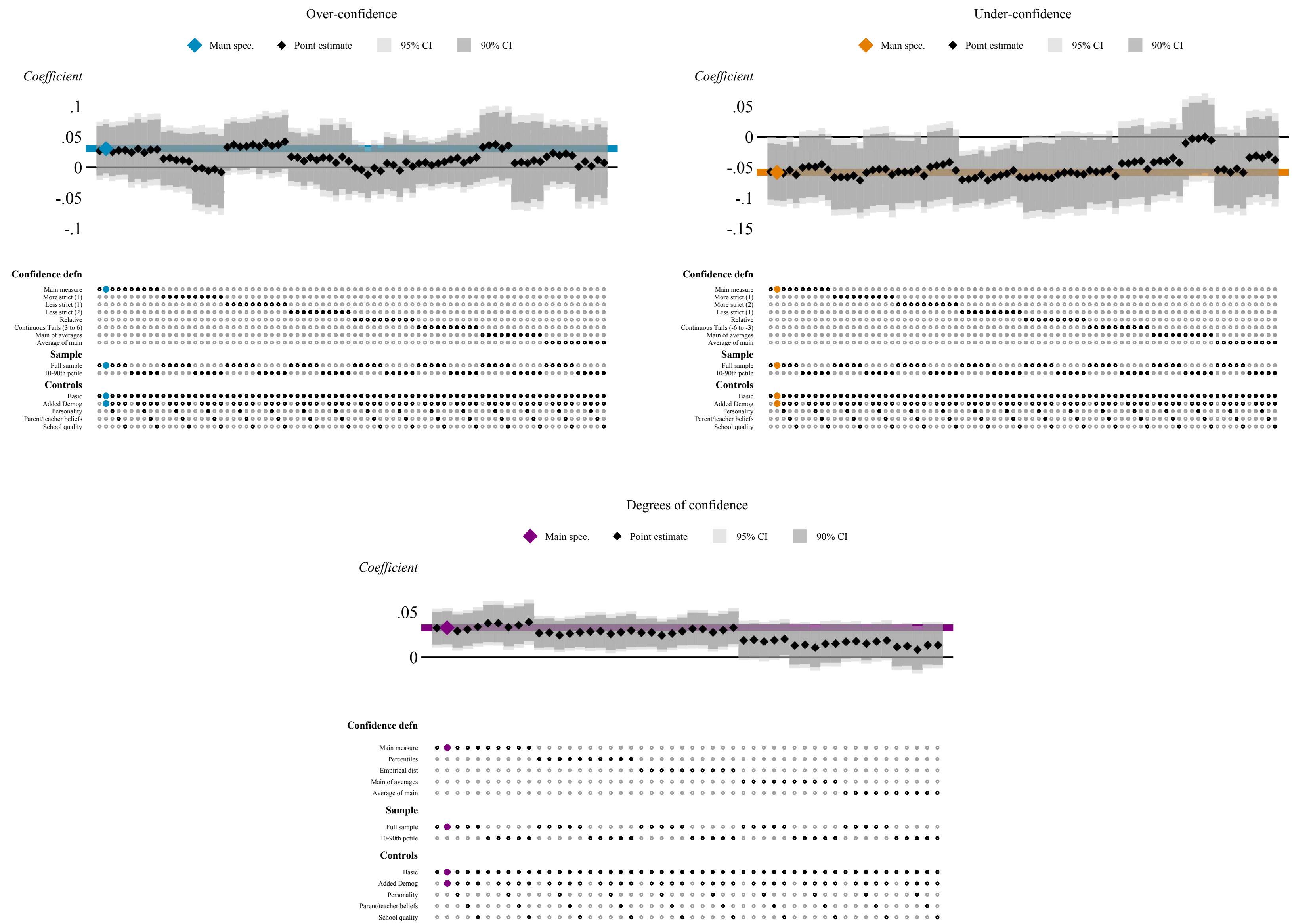
Note: This figure is analogous to Appendix Figure A4, but the outcome is adolescent reading test scores.

Figure A7: Specification chart for graduating from high school



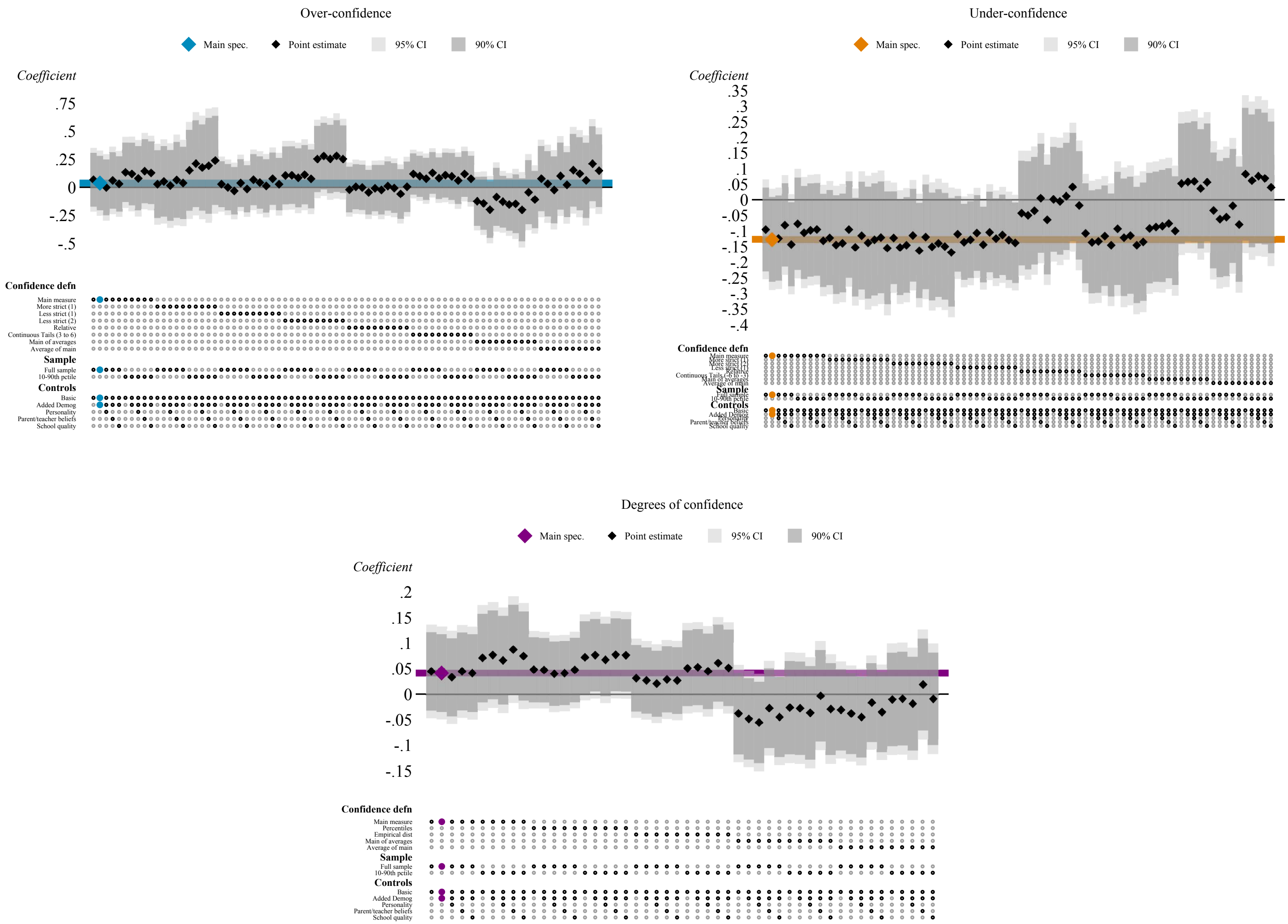
Note: This figure plots the coefficient of interest on either over- or under-confidence or our degrees of confidence measure for a large number of specification tests. The outcome is an indicator for graduating from high school. We test the relationship between childhood biased beliefs and high-school graduation when we (a) change our definitions of over- and under-confidence and the degrees of confidence measure, (b) drop the bottom and top ten percent of the ability distribution, since those children are the mechanically most likely to be over- or under-confident, respectively, and (c) iterate through each of five sets of control variables. Here, our alternate definitions of biased beliefs include measures of confidence that replicate the main measure but are based on information from multiple waves of the CDS. Appendix F describes each alternate confidence definition in detail.

Figure A8: Specification chart for graduating from college



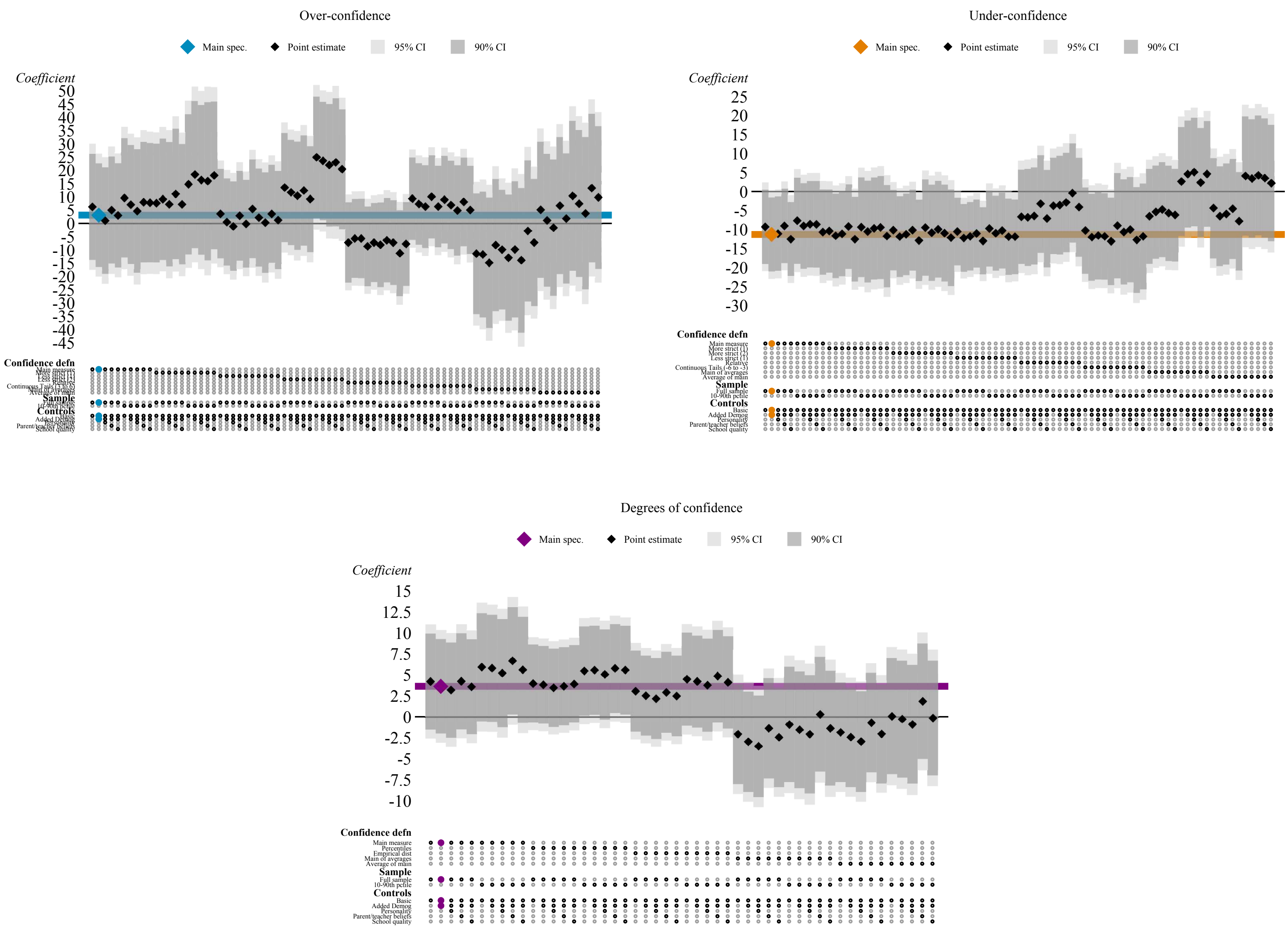
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for graduating from college.

Figure A9: Specification chart for college quality index



Note: This figure is analogous to Appendix Figure A7, but the outcome is the index of college quality for a student's first college attended (conditional on going to college).

Figure A10: Specification chart for college's 75th percentile math SAT score

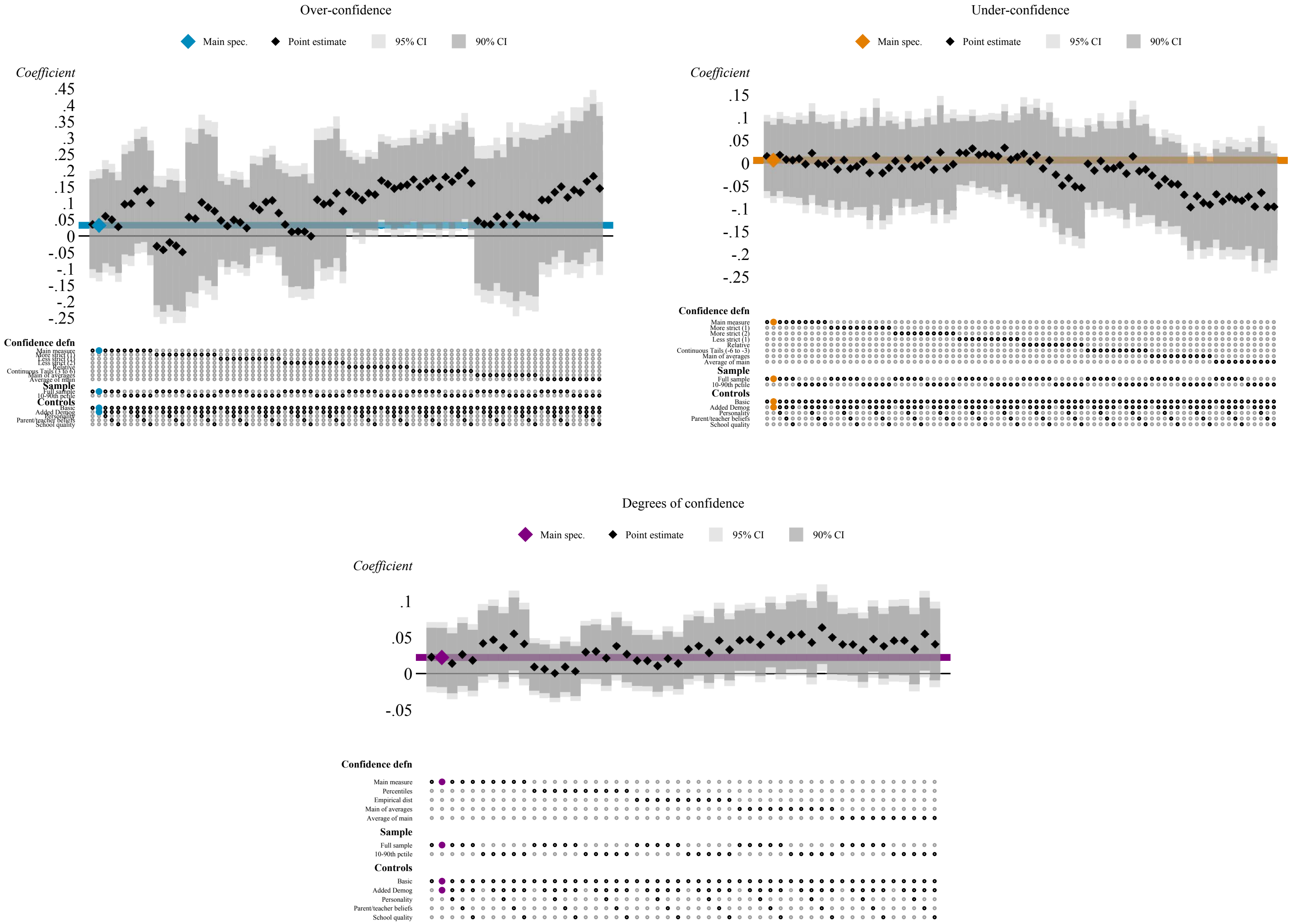


Note: This figure is analogous to Appendix Figure A7, but the outcome is the 75th percentile math SAT score at a student's first college attended (conditional on going to college).

Figure A11: Specification chart for majoring in STEM

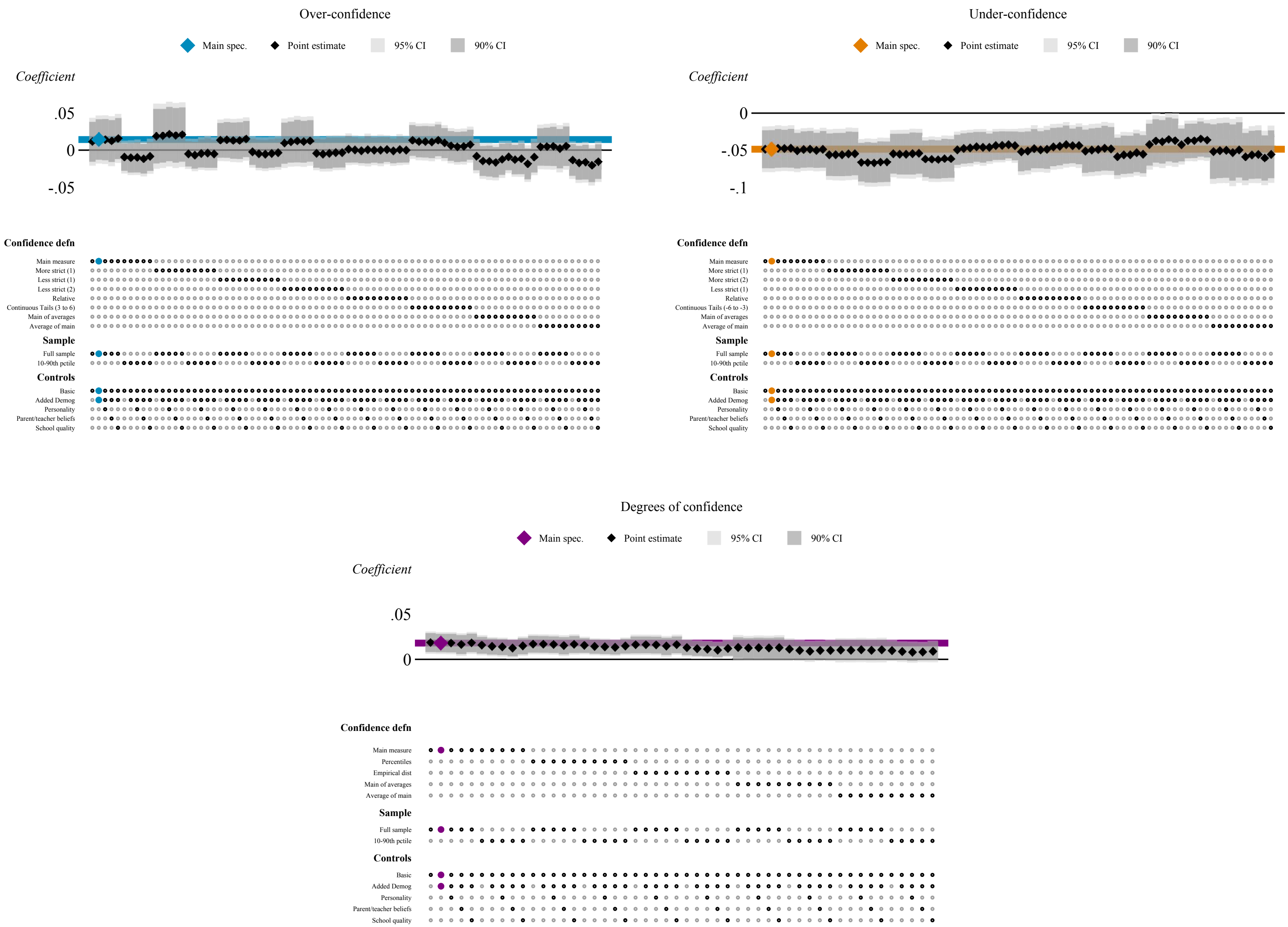
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for majoring in STEM (conditional on going to college).

Figure A12: Specification chart for having a graduate school degree



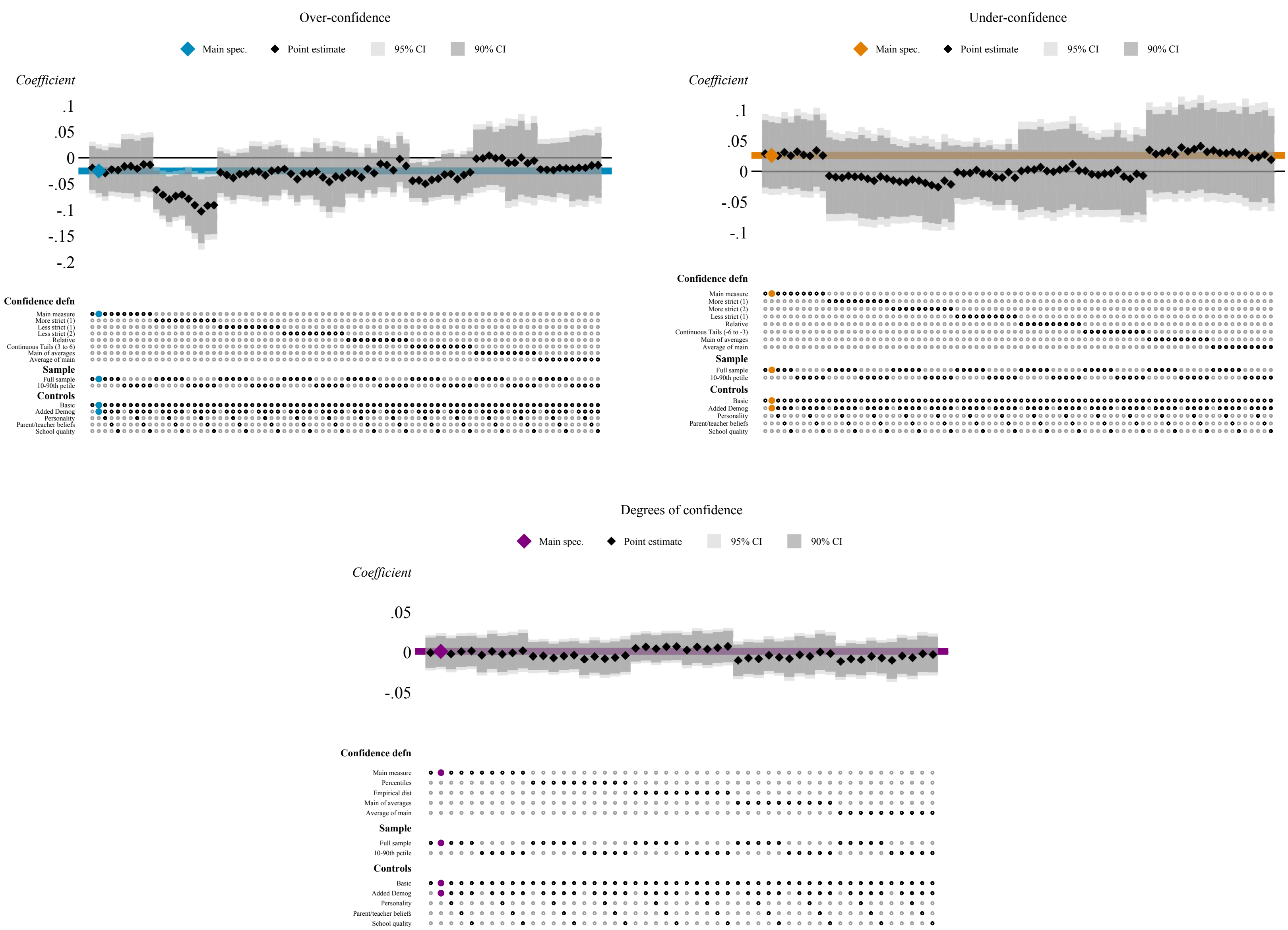
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for having a graduate degree (conditional on going to college).

Figure A13: Specification chart for working in STEM



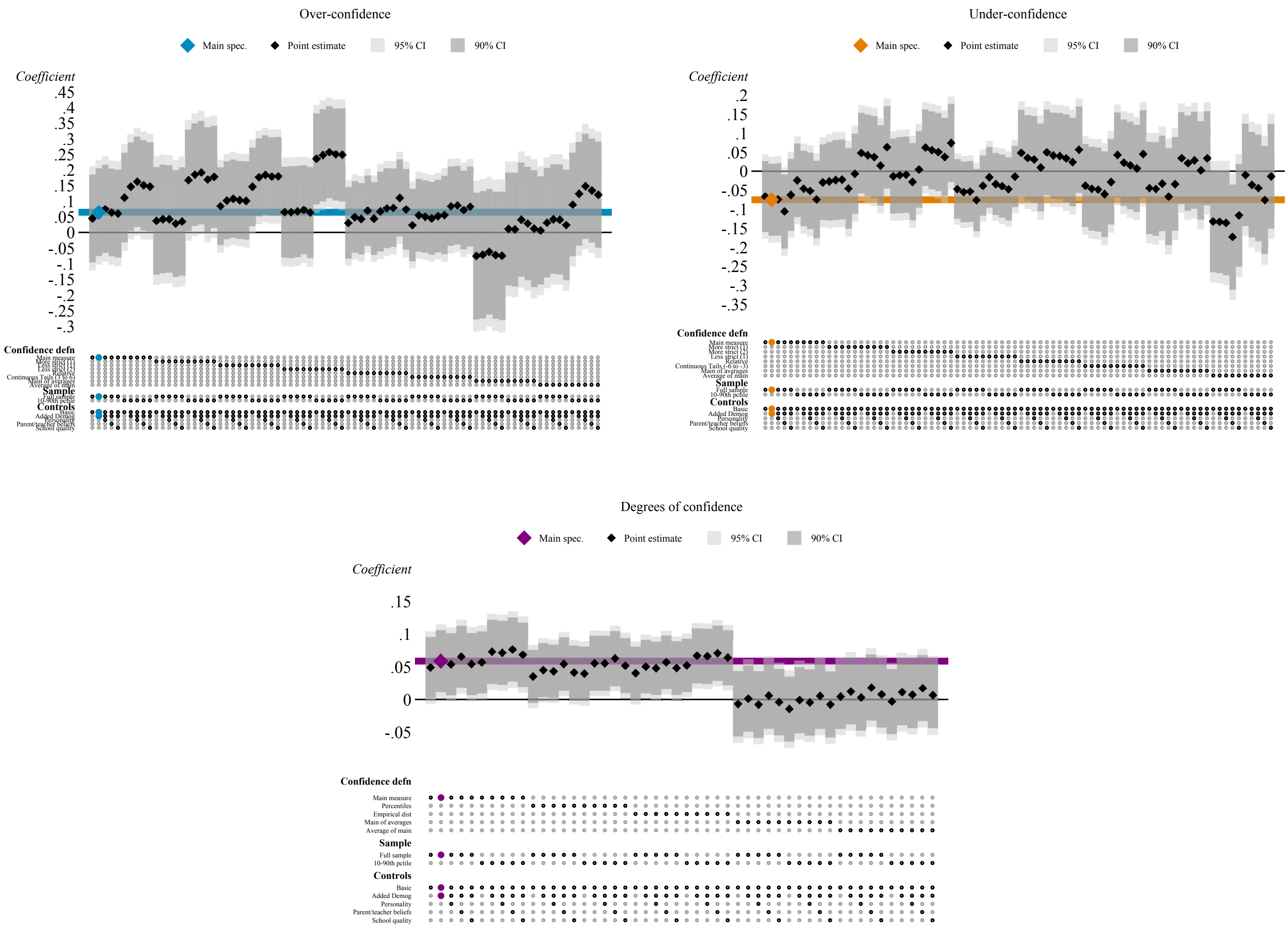
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for working in STEM.

Figure A14: Specification chart for working in non-STEM high-education occupation



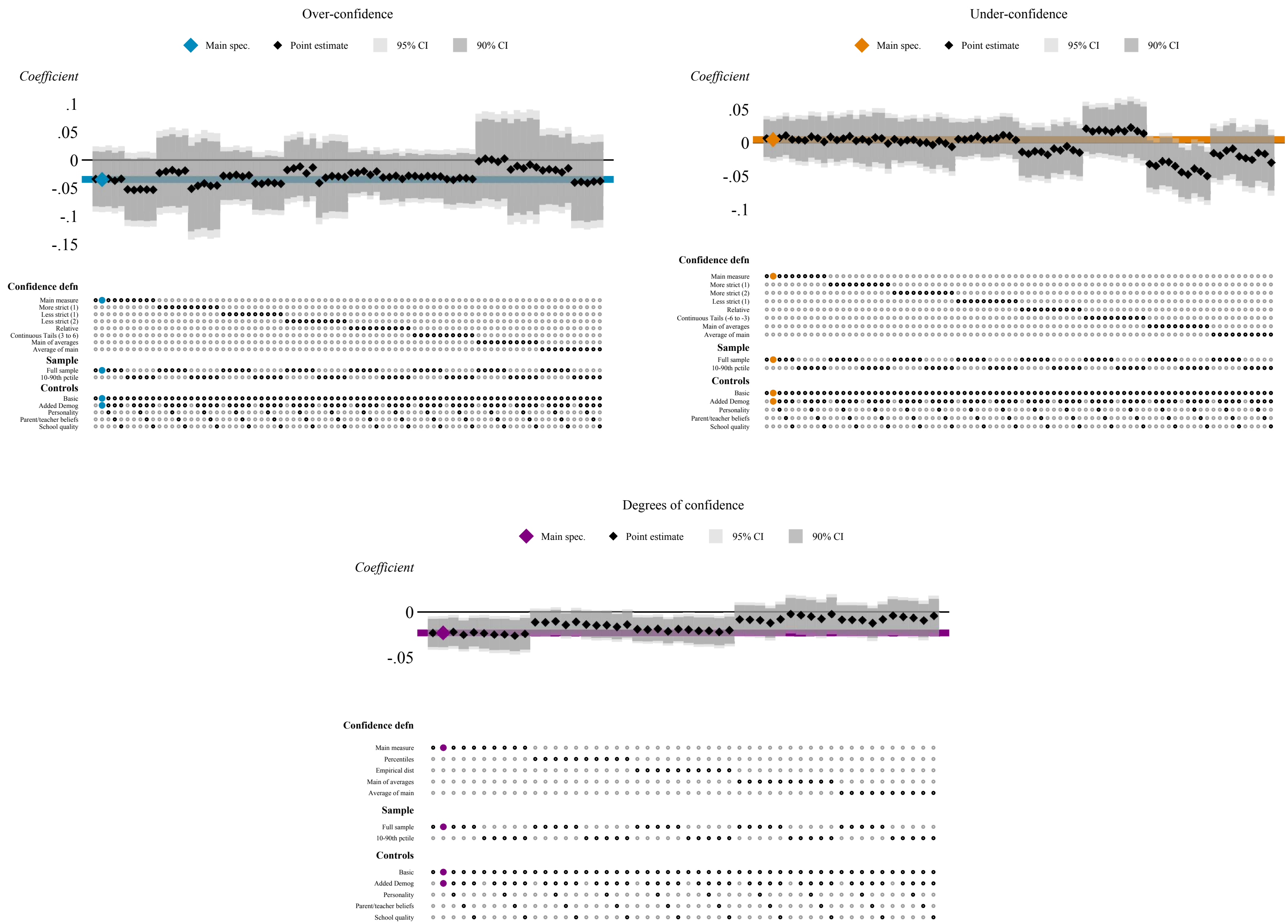
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for working in a non-STEM high-education job.

Figure A15: Specification chart for ln(earnings)



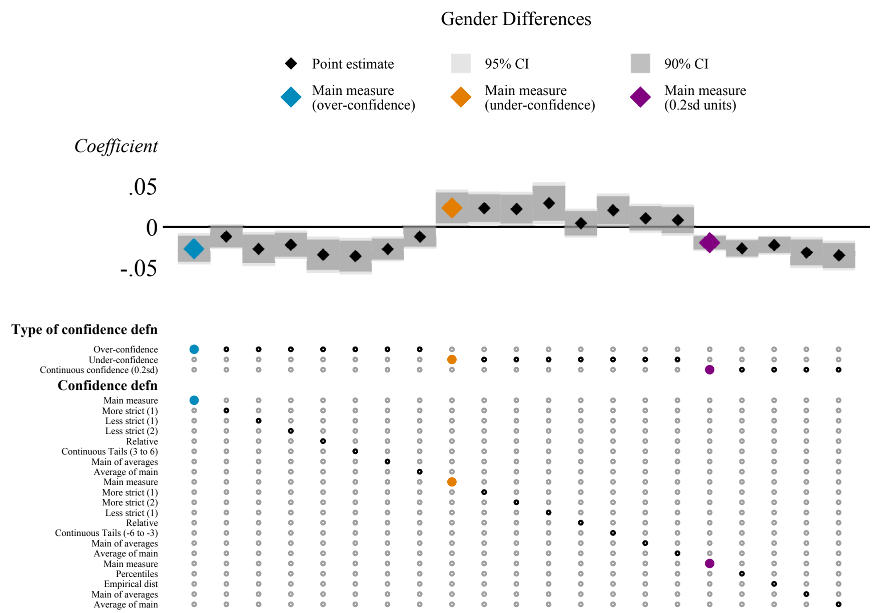
Note: This figure is analogous to Appendix Figure A7, but the outcome is log earnings.

Figure A16: Specification chart for unemployment



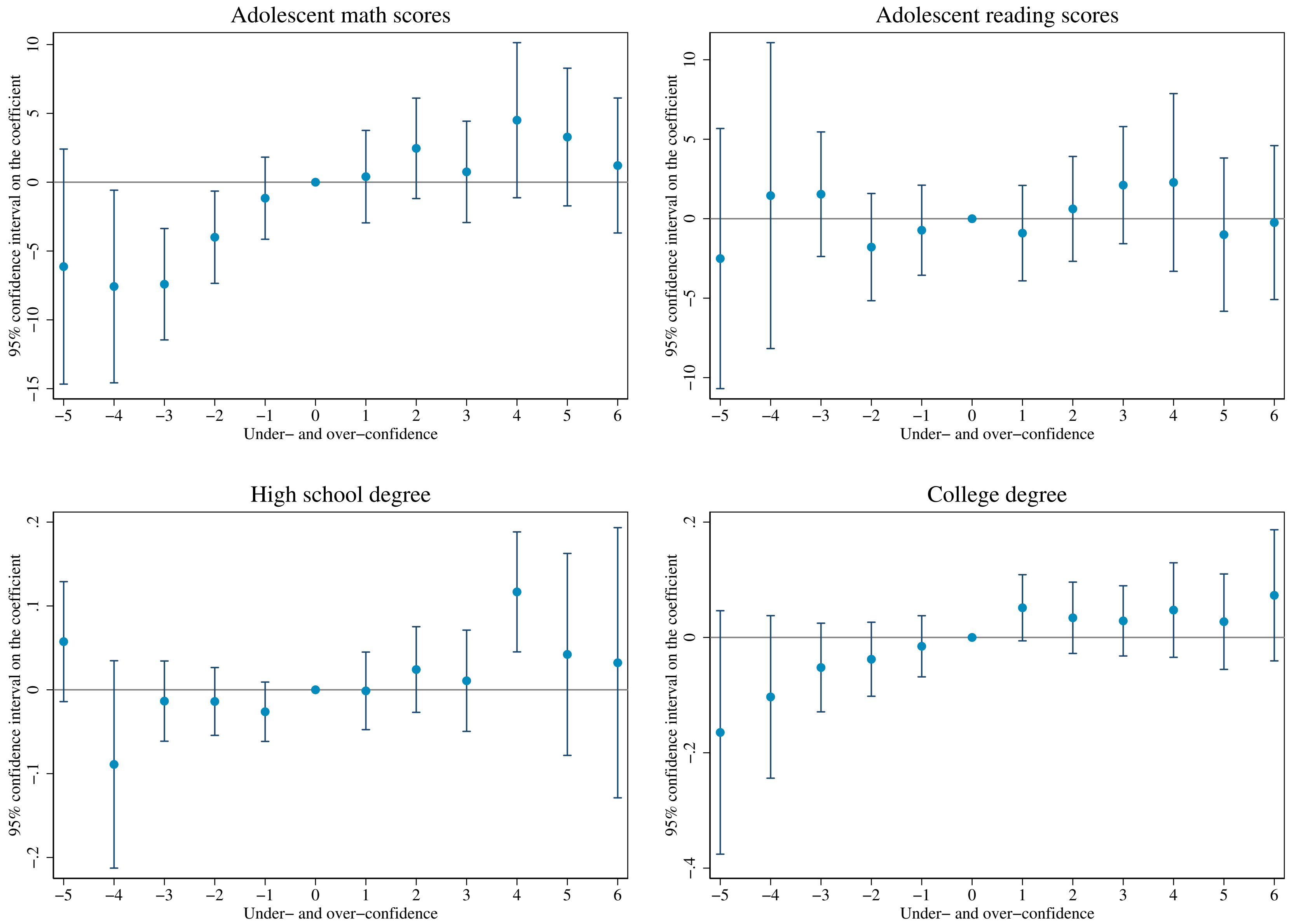
Note: This figure is analogous to Appendix Figure A7, but the outcome is an indicator for being ever unemployed in that year.

Figure A17: Specification chart for gender differences



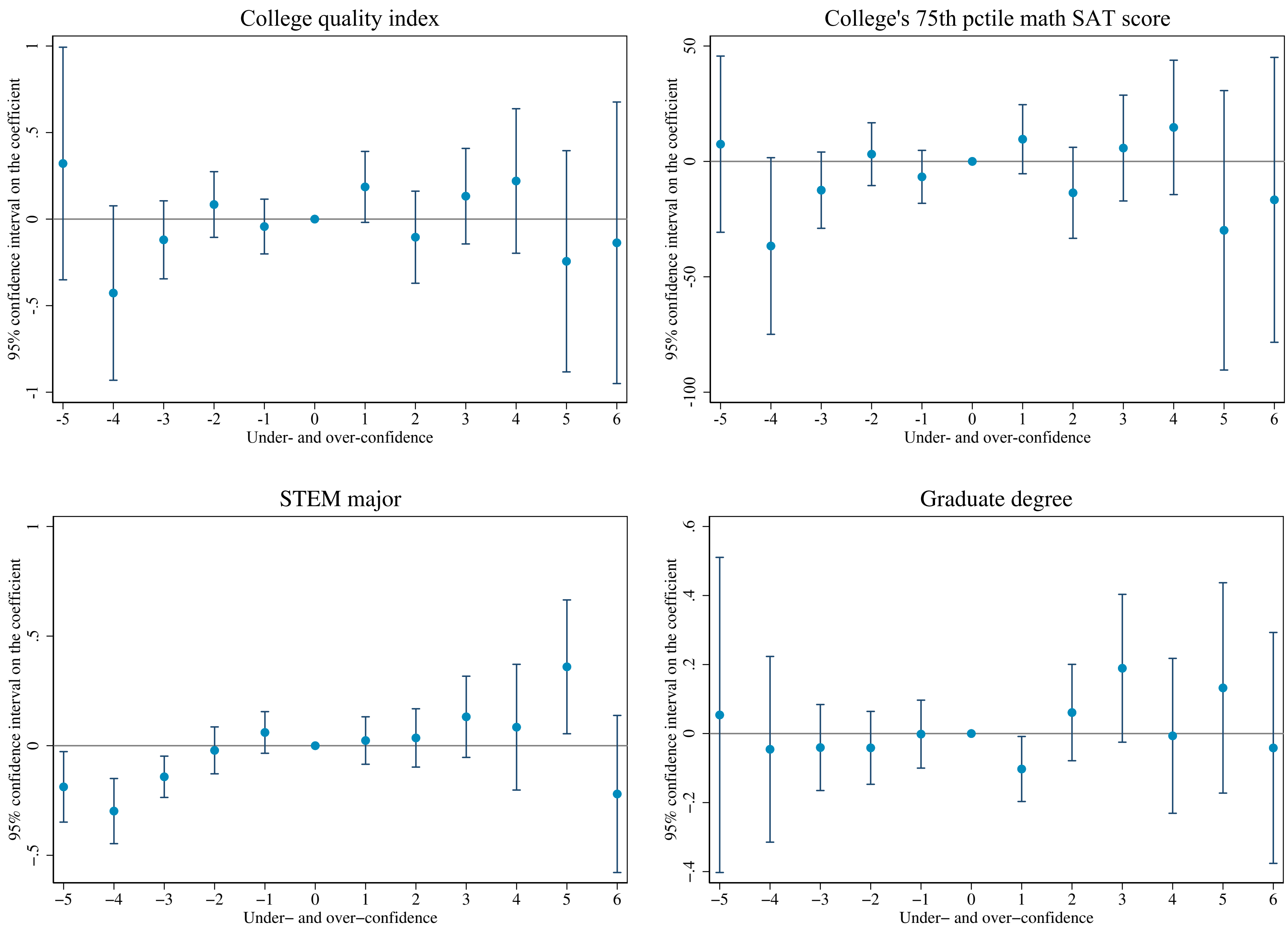
Note: This figure presents the gender gap in confidence for every measure of confidence we consider. For all measures, there is a robust gender gap: girls are less likely to be over-confident, more likely to be under-confident, and have lower degrees of confidence. Each point plots the coefficient on the female indicator when we replace the dependent variables in Table 2 with each of our alternate measures. Note that the more continuous measures of degrees of over- and under-confidence are all divided by 5 so that the resulting coefficient is on a similar scale as the coefficients when the outcome is an indicator for over- and under-confidence. This chart aims to communicate the stability of these coefficients, but one can obtain the gender gap in standard deviations by multiplying the coefficient for the more continuous measures by 5.

Figure A18: Coefficients on each confidence level fixed effect (medium-term educational achievement and attainment)



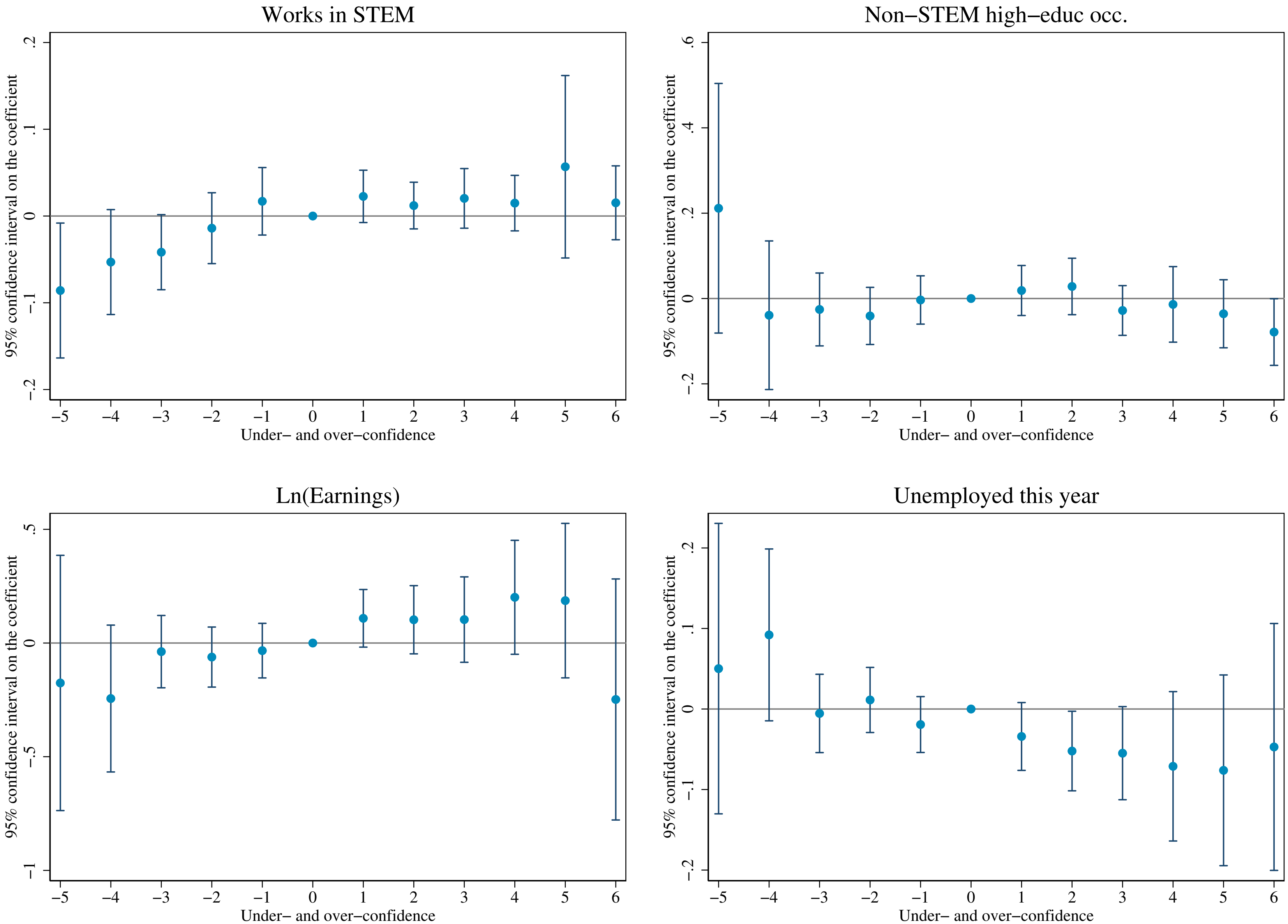
Notes: This figure plots results from a version of our main specification where we include fixed effects on each integer value of the degrees of confidence measure, which takes on integer values from -6 to 6. It is measured as the difference between a child's self-assessed ability (from 1-7) and the bin from 1-7 in which they should have placed themselves if they knew the national score distribution and their place in it. The outcomes are the same as in Table 3, where the degrees of confidence measure enters linearly into the specification. These figures support that linearity assumption.

Figure A19: Coefficients on each confidence level fixed effect (college quality, college major, and post-college schooling)



Notes: This figure parallels figure A18, but for the outcomes presented in Table 4.

Figure A20: Coefficients on each confidence level fixed effect (employment outcomes)



Notes: This figure parallels figure A18, but for the outcomes presented in Table 5.

Table A1: WJ-R Applied Problems section scores predict long-run outcomes

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	STEM major (5)	STEM occup (6)	High-educ occup (7)	ln(Earnings) (8)	Unempl (9)
Math score percentile/10	4.953***	1.316***	0.008***	0.026***	0.023***	0.006***	0.008**	0.058***	-0.011***
	(0.231)	(0.241)	(0.003)	(0.004)	(0.008)	(0.002)	(0.004)	(0.010)	(0.003)
Reading score percentile/10	1.503***	6.371***	0.015***	0.021***	0.007	-0.001	0.014***	0.014	-0.002
	(0.242)	(0.242)	(0.003)	(0.004)	(0.008)	(0.002)	(0.004)	(0.010)	(0.003)
N	1747	1745	2714	2725	736	4592	4592	4423	4975
Sample mean	50.808	48.231	0.876	0.297	0.189	0.046	0.163	10.185	0.167
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table regresses educational and employment outcomes on CDS math and reading scores. We drop our fixed effect controls for math and reading score deciles, replacing them with linear controls for math and reading score percentiles. These regressions include all controls included in Table 1 except for childhood over- and under-confidence in math. Basic controls also include year fixed effects when the outcome is observed in a panel. All controls that are indices are normalized relative to the weighted distribution. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A2: The persistence of reading over- and under-confidence

	(1)	(2)
Panel A: Reading over-confidence	0.194***	0.198***
	(0.037)	(0.037)
N	1732	
Sample mean	0.153	
Panel B: Reading under-confidence	0.098***	0.102***
	(0.034)	(0.034)
N	1732	
Sample mean	0.061	
Panel C: Reading confidence (SD units)	0.165***	0.166***
	(0.024)	(0.024)
N	1732	
Sample mean	-0.002	
Basic controls:	✓	✓
Added background controls:		✓

Notes: This table regresses adolescent confidence outcomes on childhood reading confidence with various controls. All controls are the same as described in Table 1. Standard errors are clustered by family, and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A3: Childhood reading confidence and medium-term educational achievement and attainment

Dependent variable:	Adolescent math scores		Adolescent reading scores		High school degree		College degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	-3.630***	-3.745***	1.668	1.564	-0.033	-0.032	0.008	0.008
	(1.265)	(1.291)	(1.320)	(1.314)	(0.025)	(0.025)	(0.020)	(0.020)
Under-confidence	2.075	2.002	-5.720***	-5.811***	-0.004	-0.008	-0.040	-0.043
	(1.835)	(1.830)	(1.707)	(1.749)	(0.024)	(0.023)	(0.040)	(0.040)
N	1734	1734	1732	1732	2698	2698	2709	2709
OC = -1*UC? *p-value*:	0.487	0.436	0.057	0.050	0.288	0.244	0.467	0.423
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	-1.773***	-1.750***	1.989***	2.008***	-0.014	-0.013	-0.006	-0.004
	(0.609)	(0.623)	(0.565)	(0.577)	(0.010)	(0.010)	(0.010)	(0.010)
N	1734	1734	1732	1732	2698	2698	2709	2709
Sample mean of dep. var.	50.949		48.421		0.875		0.297	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses educational achievement and attainment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A4: Childhood reading confidence and college quality, college major choice, and post-college schooling

Dependent variable:	College quality index		College's 75th pctile math SAT score		STEM Major		Graduate degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	-0.011	-0.005	-6.274	-5.052	0.057	0.065	0.009	0.016
	(0.112)	(0.110)	(9.595)	(9.452)	(0.070)	(0.070)	(0.076)	(0.074)
Under-confidence	-0.221*	-0.193*	-13.598*	-10.917	0.103*	0.079	0.060	0.054
	(0.117)	(0.114)	(8.098)	(7.932)	(0.062)	(0.064)	(0.062)	(0.062)
N	1103	1103	1112	1112	732	732	804	804
OC = -1*UC? *p-value*:	0.154	0.212	0.117	0.198	0.081	0.126	0.467	0.461
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.103**	0.090**	7.223**	6.193*	-0.045*	-0.041	-0.010	-0.012
	(0.042)	(0.042)	(3.259)	(3.238)	(0.025)	(0.027)	(0.026)	(0.026)
N	1103	1103	1112	1112	732	732	804	804
Sample mean of dep. var.	0.051		594.052		0.189		0.199	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses college outcomes outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A5: Childhood reading confidence and employment outcomes

Dependent variable:	Works in STEM		Non-STEM high-educ occ.		Ln(Earnings)		Unemployed this year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.003	0.005	0.012	0.005	-0.048	-0.063	-0.007	-0.003
	(0.012)	(0.012)	(0.022)	(0.022)	(0.071)	(0.071)	(0.022)	(0.022)
Under-confidence	0.056*	0.055*	-0.006	-0.012	0.073	0.054	-0.037	-0.034
	(0.033)	(0.033)	(0.045)	(0.045)	(0.071)	(0.070)	(0.023)	(0.022)
N	4564	4564	4564	4564	4395	4395	4943	4943
OC = -1*UC? *p-value*:	0.092	0.079	0.897	0.878	0.801	0.925	0.164	0.235
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	-0.005	-0.005	0.003	0.001	-0.054*	-0.051*	0.006	0.006
	(0.007)	(0.007)	(0.012)	(0.012)	(0.029)	(0.029)	(0.009)	(0.009)
N	4564	4564	4564	4564	4395	4395	4943	4943
Sample mean of dep. var.	0.045		0.163		10.185		0.168	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses employment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Basic controls also include year fixed effects when the outcome is observed in a panel. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A6: Summary statistics

	Mean	SD	Med.	Share Mi.
Panel A: Child Demographics				
Child is female	0.497	0.500	0	0.000
Child is white	0.458	0.498	0	0.000
Child is black	0.417	0.493	0	0.000
Child is hispanic	0.075	0.264	0	0.000
Child’s birth year	1990.020	3.748	1990	0.000
Panel B: Parent and Family Demographics				
Father at least graduated high school	0.835	0.371	1	0.370
Father at least has bachelors	0.257	0.437	0	0.369
Mother at least graduated high school	0.817	0.387	1	0.110
Mother at least has bachelors	0.134	0.341	0	0.113
Mother works in STEM	0.018	0.133	0	0.157
Father works in STEM	0.072	0.259	0	0.321
Mother works in other high-educ field	0.169	0.375	0	0.157
Father works in other high-educ field	0.109	0.311	0	0.321
Total family taxable income (thous 2016 USD)	69.777	80.333	52.03	0.000
HH receives govt transfers	0.478	0.500	0	0.000
# Siblings in the HH	1.415	1.081	1	0.000
Panel C: Parenting Practices and Beliefs				
Father figure in HH	0.727	0.445	1	0.008
Two adults in HH	0.645	0.479	1	0.000
Parent says key thing for success is:				
to obey	0.278	0.448	0	0.023
to think for one’s self	0.711	0.453	1	0.023
to work hard	0.284	0.451	0	0.023
to help others in need	0.174	0.380	0	0.023
At least once/week, parent:				
reads with child	0.386	0.487	0	0.004
does art with child	0.072	0.258	0	0.004
plays sports with child	0.157	0.364	0	0.005
does homework with child	0.634	0.482	1	0.005
plays board games with child	0.145	0.352	0	0.004
shows phys. affection to child	0.912	0.283	1	0.642
says *I love you* to child	0.894	0.308	1	0.005
Parent’s traditional gender norms (index)	0.019	0.561	0	0.147
Parent’s poor mental health (index)	-0.001	0.674	-0	0.116
Parent’s self esteem (index)	0.011	0.992	0	0.146
Parent’s self efficacy (index)	0.012	0.994	-0	0.144
Aggravation in parenting (index)	0.006	1.008	-0	0.143
Parent expectations for educ. attainment:				
Graduate degree	0.125	0.331	0	0.010
Bachelors’ degree	0.493	0.500	0	0.010
High school degree	0.376	0.484	0	0.010
High school dropout	0.006	0.078	0	0.010

Table A6: Summary statistics (continued)

	Mean	SD	Med.	Share Mi.
Panel D: Other Child Characteristics				
Child ever in gifted prog	0.243	0.429	0	0.031
Child ever in special ed prog	0.127	0.333	0	0.032
Child has repeated grade	0.122	0.327	0	0.021
Child qualifies for FRP lunch	0.598	0.490	1	0.256
Parent's rating of child health	0.014	1.000	1	0.005
Big 5 personality scores (indices)				
Conscientiousness	-0.003	0.684	0	0.007
Extroversion	0.005	0.618	-0	0.005
Neuroticism	-0.006	0.622	-0	0.009
Agreeableness	-0.003	0.644	0	0.008
Openness to experiences	-0.001	0.500	0	0.010
Panel E: Teacher Beliefs				
Perceptions of competence (stdized):				
Academic competence	0.003	0.992	-0	0.806
Social competence	-0.011	1.002	-0	0.805
Physical competence	0.003	1.006	-0	0.818
Teacher expectations for educ. attainment:				
Graduate degree	0.155	0.362	0	0.660
Bachelors' degree	0.358	0.480	0	0.660
High school degree	0.425	0.495	0	0.660
High school dropout	0.062	0.241	0	0.660
Panel F: School Quality				
Percent FRPL	0.000	0.000	0	0.000
Student-teacher ratio	0.000	0.000	0	0.000
Average math and reading achievement	0.000	0.000	0	0.000
Difference btwn math and reading achievement	0.000	0.000	0	0.000
Cohort slope of average achievement	0.000	0.000	0	0.000
Panel G: Child Ability Measures				
Math score percentile	58.477	29.214	60	0.000
Reading score percentile	55.386	28.907	54	0.004
Digit span score	14.246	3.718	14	0.041

Notes: All variables marked as indices are standardized to mean 0 and a standard deviation of 1 by year. All variables are taken from the first year in which we observe the child's over-confidence in reading or math. Except for the indicator that the child lives in a two-adult household, all variables in Panel C are reported by the child's primary caregiver. We identify two-parent households by whether a family has both a *head* and a *wife* in the main PSID.

Table A7: The persistence of math over- and under-confidence (weighted)

	(1)	(2)
Panel A: Math over-confidence	0.179***	0.177***
	(0.042)	(0.041)
N	1747	
Sample mean	0.036	
Panel B: Math under-confidence	0.046	0.041
	(0.033)	(0.033)
N	1747	
Sample mean	0.065	
Panel C: Math confidence (SD units)	0.260***	0.256***
	(0.039)	(0.040)
N	1747	
Sample mean	-0.008	
Basic controls:	✓	✓
Added background controls:		✓

Notes: This table regresses adolescent confidence outcomes on childhood math confidence with various controls. All controls are the same as described in Table 1. Observations are weighted so that the analysis sample matches the racial makeup of the US population in the 1990 census and so that the distribution of math percentile scores is uniform by decile, and the distribution of income is uniform by quartile. All controls that are indices are normalized relative to the weighted distribution. Standard errors are clustered by family, and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A8: Childhood math confidence and medium-term educational achievement and attainment (weighted)

Dependent variable:	Adolescent math scores		Adolescent reading scores		High school degree		College degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	5.103**	5.039**	1.553	1.872	0.104***	0.110***	0.053	0.067*
	(2.349)	(2.318)	(2.082)	(2.028)	(0.039)	(0.039)	(0.035)	(0.035)
Under-confidence	-4.565**	-4.108**	1.046	0.901	0.042*	0.042*	-0.011	-0.007
	(1.995)	(2.008)	(1.891)	(1.922)	(0.023)	(0.023)	(0.043)	(0.041)
N	1747	1747	1745	1745	2714	2714	2725	2725
OC = -1*UC? *p-value*:	0.861	0.759	0.355	0.302	0.002	0.001	0.446	0.264
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	2.871***	2.684***	-0.204	-0.020	0.027	0.026	0.017	0.016
	(0.826)	(0.841)	(0.832)	(0.834)	(0.017)	(0.017)	(0.015)	(0.015)
N	1747	1747	1745	1745	2714	2714	2725	2725
Sample mean of dep. var.	46.836		46.068		0.868		0.270	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses educational achievement and attainment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Observations are weighted so that our sample matches population shares in quintiles of income, in race categories, and in deciles of nationally-normed WJ-R math percentile scores. All controls that are indices are normalized relative to the weighted distribution. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A9: Childhood math confidence and college quality, college major choice, and post-college schooling (weighted)

Dependent variable:	College quality index		College's 75th pctile math SAT score		STEM major		Graduate degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.035	0.083	11.390	10.273	0.167	0.211*	-0.030	-0.029
	(0.165)	(0.167)	(11.115)	(11.569)	(0.122)	(0.128)	(0.060)	(0.070)
Under-confidence	-0.118	-0.132	-12.123*	-12.976*	-0.121**	-0.113**	-0.046	-0.071
	(0.106)	(0.106)	(7.178)	(6.990)	(0.054)	(0.050)	(0.059)	(0.061)
N	1107	1107	1117	1117	736	736	810	810
OC = -1*UC? *p-value*:	0.689	0.813	0.957	0.845	0.732	0.472	0.359	0.281
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.145**	0.137**	11.005***	10.256***	0.086**	0.074**	0.027	0.029
	(0.066)	(0.064)	(3.997)	(3.889)	(0.035)	(0.035)	(0.032)	(0.030)
N	1107	1107	1117	1117	736	736	810	810
Sample mean of dep. var.	0.091		599.785		0.186		0.180	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses college outcomes outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Observations are weighted so that our sample matches population shares in quintiles of income, in race categories, and in deciles of nationally-normed WJ-R math percentile scores. All controls that are indices are normalized relative to the weighted distribution. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A10: Childhood math confidence and employment outcomes (weighted)

Dependent variable:	Works in STEM		Non-STEM high-educ occ.		Ln(Earnings)		Unemployed this year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.067	0.064	-0.020	-0.020	-0.015	-0.006	-0.012	-0.009
	(0.054)	(0.046)	(0.035)	(0.034)	(0.135)	(0.136)	(0.036)	(0.036)
Under-confidence	-0.063**	-0.060***	0.018	0.021	-0.013	-0.018	-0.019	-0.024
	(0.025)	(0.023)	(0.040)	(0.038)	(0.074)	(0.072)	(0.020)	(0.021)
N	4592	4592	4592	4592	4423	4423	4975	4975
OC = -1*UC? *p-value*:	0.949	0.949	0.965	0.990	0.859	0.873	0.449	0.436
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.042***	0.040***	-0.008	-0.006	-0.009	-0.001	-0.014	-0.013
	(0.016)	(0.014)	(0.016)	(0.015)	(0.047)	(0.046)	(0.013)	(0.012)
N	4592	4592	4592	4592	4423	4423	4975	4975
Sample mean of dep. var.	0.049		0.151		10.180		0.136	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses employment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Basic controls also include year fixed effects when the outcome is observed in a panel. Observations are weighted so that our sample matches population shares in quintiles of income, in race categories, and in deciles of nationally-normed WJ-R math percentile scores. All controls that are indices are normalized relative to the weighted distribution. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A11: Demographic predictors of over- and under-confidence (decile coefficients)

	Over-confidence	Under-confidence	Confidence (sd)
Math score deciles			
Decile 1	0.147***	-0.257***	1.918***
	(0.04)	(0.03)	(0.09)
Decile 2	0.212***	-0.257***	1.767***
	(0.03)	(0.03)	(0.07)
Decile 3	0.204***	-0.246***	1.564***
	(0.03)	(0.03)	(0.07)
Decile 4	0.140***	-0.250***	1.231***
	(0.02)	(0.02)	(0.06)
Decile 5	0.189***	-0.250***	1.075***
	(0.02)	(0.02)	(0.05)
Decile 6	-0.001	-0.163***	0.945***
	(0.01)	(0.03)	(0.06)
Decile 7	-0.003	-0.161***	0.675***
	(0.01)	(0.03)	(0.06)
Decile 8	-0.004	0.028	0.250***
	(0.01)	(0.03)	(0.05)
Decile 9	0.004	0.049	0.064
	(0.01)	(0.03)	(0.05)
Decile 10	0.000	0.000	0.000
	(.)	(.)	(.)
Reading score deciles			
Decile 1	0.160**	-0.026	0.281
	(0.07)	(0.08)	(0.21)
Decile 2	0.074	-0.021	0.114
	(0.07)	(0.08)	(0.21)
Decile 3	0.001	-0.030	-0.006
	(0.07)	(0.08)	(0.20)
Decile 4	0.021	-0.034	0.043
	(0.07)	(0.08)	(0.20)
Decile 5	0.001	-0.001	-0.098
	(0.07)	(0.08)	(0.20)
Decile 6	-0.012	0.006	-0.157
	(0.07)	(0.08)	(0.20)
Decile 7	0.002	0.010	-0.197
	(0.07)	(0.08)	(0.20)
Decile 8	-0.000	0.041	-0.240
	(0.07)	(0.08)	(0.20)
Decile 9	-0.018	0.051	-0.277
	(0.07)	(0.08)	(0.20)
Decile 10	-0.019	0.093	-0.301
	(0.07)	(0.08)	(0.20)
Mean of dependent variable	0.085	0.121	0.000
N	2985	2985	2985
R-squared	0.21	0.21	0.57

Notes: This table shows the coefficients on math and reading test score decile fixed effects that are not included in Table 2 due to space constraints.

Table A12: Benchmarking the relationships between confidence and long-run outcomes and test scores

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality ind. (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earnings) (11)	Unempl (12)
Panel A: Independent variables are binary measures of over- and under-confidence												
Math over-confidence	3.618**	0.544	0.054**	0.038	0.070	5.079	0.117	-0.007	0.023	-0.030	0.033	-0.029
	(1.464)	(1.319)	(0.026)	(0.024)	(0.135)	(11.253)	(0.095)	(0.083)	(0.017)	(0.027)	(0.083)	(0.029)
Math under-confidence	-6.482***	0.797	0.021	-0.056**	-0.136*	-11.503**	-0.149***	0.005	-0.049***	0.023	-0.089	0.007
	(1.475)	(1.407)	(0.017)	(0.027)	(0.080)	(5.706)	(0.035)	(0.046)	(0.015)	(0.032)	(0.057)	(0.016)
Math pctile/10	5.291***	1.289***	0.008***	0.029***	0.056***	4.068***	0.030***	0.015*	0.008***	0.005	0.060***	-0.012***
	(0.235)	(0.250)	(0.003)	(0.004)	(0.014)	(1.107)	(0.009)	(0.008)	(0.003)	(0.004)	(0.011)	(0.003)
Reading pctile/10	1.597***	6.387***	0.015***	0.022***	0.048***	3.850***	0.007	0.013	0.000	0.013***	0.014	-0.002
	(0.240)	(0.241)	(0.003)	(0.004)	(0.016)	(1.082)	(0.008)	(0.008)	(0.002)	(0.004)	(0.010)	(0.003)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Math confidence	2.963***	0.223	0.017*	0.037***	0.060	4.795	0.076***	0.015	0.018***	0.004	0.058**	-0.021**
	(0.563)	(0.560)	(0.010)	(0.011)	(0.044)	(3.237)	(0.022)	(0.025)	(0.006)	(0.011)	(0.028)	(0.009)
Math pctile/10	5.583***	1.357***	0.011***	0.034***	0.062***	4.563***	0.038***	0.019**	0.009***	0.008	0.068***	-0.016***
	(0.260)	(0.274)	(0.004)	(0.005)	(0.016)	(1.260)	(0.009)	(0.009)	(0.003)	(0.005)	(0.012)	(0.004)
Reading pctile/10	1.655***	6.399***	0.016***	0.023***	0.050***	3.962***	0.009	0.013	0.000	0.013***	0.015	-0.002
	(0.240)	(0.242)	(0.003)	(0.004)	(0.016)	(1.089)	(0.008)	(0.008)	(0.002)	(0.004)	(0.010)	(0.003)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975

Notes: This table presents the same regressions as the even-numbered columns of Tables 3, 4, and 5, but replaces the math and reading test score decile fixed effects with linear terms for math and reading percentile scores divided by 10 (so that the coefficients can be interpreted in terms of increasing test scores by one decile). We use this table to benchmarch the relationships between math confidence versus math test scores and long-term outcomes. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A13: Robustness to potential confounders

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality ind. (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earnings) (11)	Unempl (12)
Section 1: Controlling for childhood Big 5 personality traits												
Panel A: Independent variables are binary measures of over- and under-confidence												
Over-confidence	2.600*	-0.286	0.061**	0.025	-0.001	1.075	0.068	0.060	0.014	-0.030	0.075	-0.033
	(1.483)	(1.393)	(0.026)	(0.025)	(0.150)	(12.165)	(0.104)	(0.087)	(0.017)	(0.027)	(0.086)	(0.030)
Under-confidence	-5.808***	0.187	0.023	-0.060**	-0.123	-11.110*	-0.160***	0.017	-0.048***	0.025	-0.074	0.007
	(1.513)	(1.464)	(0.017)	(0.028)	(0.083)	(5.961)	(0.037)	(0.048)	(0.016)	(0.032)	(0.058)	(0.017)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Confidence	2.824***	0.168	0.014	0.029***	0.033	3.194	0.077***	0.014	0.018***	-0.002	0.054*	-0.022**
	(0.569)	(0.579)	(0.010)	(0.011)	(0.046)	(3.442)	(0.024)	(0.025)	(0.006)	(0.012)	(0.029)	(0.010)
N		1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Section 2: Controlling for parent and teacher expectations and investments												
Panel A: Independent variables are binary measures of over- and under-confidence												
Over-confidence	2.861*	-0.274	0.059**	0.028	0.061	5.025	0.089	0.050	0.012	-0.023	0.064	-0.037
	(1.487)	(1.385)	(0.025)	(0.024)	(0.153)	(12.230)	(0.099)	(0.087)	(0.017)	(0.026)	(0.085)	(0.030)
Under-confidence	-5.295***	0.398	0.015	-0.055**	-0.081	-9.065	-0.150***	0.007	-0.048***	0.031	-0.106*	0.011
	(1.520)	(1.460)	(0.017)	(0.028)	(0.081)	(5.946)	(0.038)	(0.049)	(0.015)	(0.033)	(0.059)	(0.018)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Confidence	2.761***	0.165	0.017*	0.031***	0.044	4.228	0.078***	0.026	0.016***	0.001	0.065**	-0.025**
	(0.582)	(0.583)	(0.010)	(0.011)	(0.046)	(3.461)	(0.024)	(0.026)	(0.006)	(0.012)	(0.029)	(0.010)
N		1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Section 3: Controlling for elementary school quality												
Panel A: Independent variables are binary measures of over- and under-confidence												
Over-confidence	2.938*	-0.195	0.060**	0.028	0.030	3.039	0.076	0.028	0.016	-0.024	0.060	-0.033
	(1.511)	(1.408)	(0.026)	(0.025)	(0.150)	(11.908)	(0.097)	(0.086)	(0.017)	(0.026)	(0.086)	(0.030)
Under-confidence	-5.931***	-0.097	0.020	-0.062**	-0.143*	-12.553**	-0.159***	0.006	-0.047***	0.026	-0.062	0.005
	(1.499)	(1.460)	(0.018)	(0.028)	(0.082)	(5.892)	(0.037)	(0.049)	(0.016)	(0.033)	(0.058)	(0.017)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Confidence	2.944***	0.251	0.018*	0.034***	0.041	3.577	0.073***	0.018	0.018***	0.001	0.054*	-0.022**
	(0.574)	(0.582)	(0.010)	(0.011)	(0.046)	(3.426)	(0.023)	(0.025)	(0.006)	(0.012)	(0.029)	(0.010)
N		1745	2714	2725	1107	1117	736	810	4592	4592	4423	4975

Notes: This table presents the robustness of our main results to adding controls for potential confounders. In the first section, we add controls for measurements of children's big 5 personality traits taken at the same time as the confidence measurements (conscientiousness, openness, extroversion, agreeableness, and neuroticism). In the second section, we add controls for parent and teacher expectations and investment: how teachers rate the child's social, physical, and academic competency; whether parents report reading, playing sports, doing homework, playing games, expressing physical affection, and saying *I love you* more than once per week; and separate indicators for whether parents and teachers think the child will get a high school or bachelors degree. Finally, the third section adds controls for elementary school quality at the time of confidence measurement: the student-teacher ratio, the percent of students qualifying for free or reduced-price lunch, and three measures of school achievement *from 2009-2018*: the average math and reading score, the difference between math and reading scores, and the cohort slope on the average math and reading score. Each of these sets of controls is individually added to our main specification in the even-numbered columns of Tables 3, 4, and 5 – the sections are not cumulative. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A14: Correlations between childhood confidence and personality measures

	Math over-confidence		Math under-confidence		Math confidence (sd)	
	(1)	(2)	(3)	(4)	(5)	(6)
Conscientiousness	-0.023**	-0.024**	0.025**	0.028**	-0.226***	-0.232***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.038)	(0.038)
Extroversion	0.000	-0.002	-0.001	0.000	0.027	0.022
	(0.010)	(0.009)	(0.010)	(0.010)	(0.032)	(0.032)
Neuroticism	0.014	0.013	0.005	0.004	0.046	0.046
	(0.012)	(0.012)	(0.013)	(0.013)	(0.042)	(0.042)
Agreeableness	-0.006	-0.009	-0.011	-0.007	0.075	0.068
	(0.015)	(0.015)	(0.014)	(0.014)	(0.048)	(0.049)
Openness	-0.017	-0.018	0.010	0.011	-0.160***	-0.163***
	(0.013)	(0.013)	(0.011)	(0.011)	(0.041)	(0.041)
General Confidence	—	0.020**	—	-0.033***	—	0.070**
		(0.008)		(0.009)		(0.028)
R-squared	0.009	0.023	0.005	0.013	0.026	0.029
N	2985	2985	2985	2985	2985	2985

Note: This table shows the relationship between math confidence and measures of childhood personality and general confidence. In columns 1 and 2 the outcome is our main binary measure of over- or under-confidence, respectively. In column 3, the outcome is our measure of the degrees of confidence that takes on values of -6 to 6, *standardized to have mean zero and standard deviation one in our sample*. All independent variables are recoded to zero if missing and we include a missing indicator (coefficient not shown). Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A15: Parent and teacher predictors of math over- and under-confidence

	Math over-confidence			Math under-confidence			Confidence (sd)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Teacher Perceptions									
Academic competence	0.015	0.017	0.016	-0.019	-0.019	-0.020	-0.006	-0.002	-0.008
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Social competence	-0.019	-0.016	-0.015	0.017	0.016	0.015	-0.029	-0.024	-0.020
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Physical competence	-0.015	-0.019	-0.018	-0.002	-0.000	-0.001	-0.026	-0.038	-0.037
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Teacher Expectations									
Expects grad degree	0.002	0.003	0.002	-0.073*	-0.073*	-0.052	0.153**	0.157**	0.134*
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)
Expects bachelors' degree	0.010	0.009	0.012	0.041	0.040	0.048*	0.047	0.049	0.051
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.06)
Parent Investment									
Reads to child	0.008	0.008	0.006	0.024*	0.024*	0.025*	-0.016	-0.014	-0.016
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Art with child	-0.035*	-0.028	-0.031	-0.015	-0.020	-0.019	-0.060	-0.043	-0.048
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
Sports with child	0.006	0.004	0.006	-0.030*	-0.026	-0.027*	0.077**	0.068*	0.076**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Homework with child	0.007	0.005	0.006	0.025*	0.026*	0.026*	-0.022	-0.028	-0.029
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Games with child	0.005	0.004	0.003	-0.030*	-0.032**	-0.033**	0.070*	0.071*	0.068*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Physical affection to child	0.028	0.029	0.027	-0.039	-0.039	-0.037	0.074	0.083	0.078
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.09)	(0.09)	(0.09)
Says *I love you* to child	-0.016	-0.011	-0.012	0.010	0.009	0.012	-0.008	0.010	0.009
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.05)	(0.05)
Parent Expectations									
Expects grad degree	-0.016	-0.015	-0.015	-0.043**	-0.044**	-0.020	0.036	0.036	0.019
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.05)
Expects bachelors' degree	-0.003	-0.001	0.000	0.003	0.003	0.014	0.024	0.032	0.029
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Digit span score	0.000	0.000	-0.000	-0.004*	-0.003*	-0.003*	0.009**	0.008**	0.009**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
General confidence	0.037***	0.037***	0.038***	-0.054***	-0.056***	-0.054***	0.208***	0.212***	0.210***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Added demographic controls		✓	✓		✓	✓		✓	✓
Added all other Table 2 controls			✓			✓			✓
N	2985	2985	2985	2985	2985	2985	2985	2985	2985
R-squared	0.20	0.21	0.22	0.20	0.21	0.23	0.57	0.57	0.58

Notes: All variables are taken from the first year in which we observe the child's confidence in math. Teacher and parent expectations are indicators for each adult's expected educational attainment for each child, and the omitted category is expecting a child to obtain a high school degree or less. Parent investment controls are indicators for doing each activity more than once per week. Teacher perceptions of competence in each domain are standardized to have mean zero and standard deviation one based on teacher reports of whether a child is extremely competent to not at all competent on a four-point scale. Additional controls in all columns include math and reading test score decile fixed effects, birth year, birth quarter, state, and age at which confidence was measured fixed effects. All controls are recoded to be zero if missing and the regressions include missing indicators for each variable (not shown). Standard errors are clustered by family.

Table A16: Robustness to definitions of confidence

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earn) (11)	Unempl (12)
Panel A: Independent variables are binary measures of over- and under-confidence												
Main measure:												
Over-confidence	2.666*	-0.286	0.062**	0.031	0.037	3.133	0.076	0.032	0.014	-0.025	0.064	-0.035
	(1.496)	(1.385)	(0.026)	(0.024)	(0.148)	(11.829)	(0.097)	(0.087)	(0.017)	(0.026)	(0.085)	(0.030)
Under-confidence	-5.860***	0.162	0.022	-0.058**	-0.127	-11.312*	-0.162***	0.006	-0.049***	0.026	-0.075	0.005
	(1.497)	(1.452)	(0.017)	(0.028)	(0.082)	(5.925)	(0.036)	(0.048)	(0.016)	(0.033)	(0.057)	(0.017)
More strict (1):												
Over-confidence	2.412	0.091	0.034	0.016	0.052	9.116	-0.154	-0.042	0.019	-0.071***	0.042	-0.020
	(1.717)	(1.609)	(0.036)	(0.027)	(0.190)	(13.872)	(0.121)	(0.115)	(0.022)	(0.024)	(0.105)	(0.038)
Under-confidence	-7.246***	1.150	0.012	-0.066**	-0.144	-11.660*	-0.183***	-0.014	-0.056***	-0.009	-0.023	0.004
	(1.668)	(1.629)	(0.020)	(0.032)	(0.096)	(6.749)	(0.039)	(0.054)	(0.018)	(0.037)	(0.064)	(0.020)
More strict (2):												
Over-confidence	2.631*	-0.290	0.062**	0.030	0.038	3.173	0.075	0.032	0.014	-0.025	0.063	-0.035
	(1.495)	(1.384)	(0.026)	(0.024)	(0.148)	(11.827)	(0.097)	(0.087)	(0.017)	(0.026)	(0.085)	(0.030)
Under-confidence	-6.820***	1.046	0.021	-0.058*	-0.152	-11.870*	-0.165***	-0.012	-0.055***	-0.016	-0.010	0.001
	(1.622)	(1.595)	(0.019)	(0.031)	(0.093)	(6.464)	(0.039)	(0.052)	(0.017)	(0.035)	(0.061)	(0.019)
Less strict (1):												
Over-confidence	1.344	-0.161	0.043*	0.037	-0.000	0.537	0.066	0.030	0.014	-0.033	0.102	-0.028
	(1.406)	(1.300)	(0.024)	(0.023)	(0.126)	(9.982)	(0.073)	(0.076)	(0.015)	(0.025)	(0.076)	(0.026)
Under-confidence	-5.859***	-1.266	0.005	-0.069***	-0.135*	-12.247**	-0.128***	0.022	-0.048***	-0.003	-0.055	0.004
	(1.311)	(1.255)	(0.017)	(0.024)	(0.074)	(5.400)	(0.035)	(0.043)	(0.012)	(0.026)	(0.049)	(0.016)
Less strict (2):												
Over-confidence	1.873	0.411	0.057*	0.016	0.106	11.714	-0.038	0.012	0.011	-0.031	0.064	-0.015
	(1.607)	(1.570)	(0.030)	(0.024)	(0.156)	(12.051)	(0.081)	(0.102)	(0.017)	(0.025)	(0.089)	(0.031)
Under-confidence	-5.895***	-1.265	0.004	-0.071***	-0.136*	-12.345**	-0.128***	0.021	-0.048***	-0.002	-0.057	0.005
	(1.309)	(1.253)	(0.017)	(0.024)	(0.074)	(5.408)	(0.035)	(0.043)	(0.012)	(0.026)	(0.048)	(0.015)
Relative:												
Over-confidence	2.795*	2.308*	0.041*	-0.004	0.004	-5.672	0.066	0.121*	0.000	-0.030	0.049	-0.023
	(1.480)	(1.361)	(0.023)	(0.023)	(0.122)	(9.029)	(0.062)	(0.067)	(0.010)	(0.024)	(0.070)	(0.022)
Under-confidence	-9.674***	-1.848	0.025	-0.066*	-0.049	-6.820	-0.160***	0.003	-0.051***	0.003	0.035	-0.017
	(2.081)	(1.956)	(0.022)	(0.034)	(0.105)	(7.755)	(0.049)	(0.064)	(0.016)	(0.040)	(0.069)	(0.022)

Table A16: Robustness to definitions of confidence (continued)

	Math Score	Reading Score	HS grad	College grad	College quality	College math SAT 75p	STEM major	Grad degree	STEM occup	High-educ occup	ln(Earn)	Unempl
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Continuous tails:												
Over-confidence	1.127	1.751	0.033	0.008	0.098	7.333	0.084	0.149*	0.011	-0.043**	0.055	-0.031
	(1.372)	(1.339)	(0.023)	(0.021)	(0.112)	(9.301)	(0.071)	(0.081)	(0.013)	(0.021)	(0.071)	(0.024)
Under-confidence	-6.680***	2.244	-0.001	-0.057*	-0.136	-12.011*	-0.191***	-0.017	-0.050***	-0.004	-0.046	0.017
	(1.664)	(1.655)	(0.021)	(0.032)	(0.100)	(6.948)	(0.038)	(0.052)	(0.018)	(0.036)	(0.067)	(0.021)
Main of averages:												
Over-confidence	–	–	0.008	0.036	-0.143	-11.712	-0.082	0.036	-0.015	-0.001	-0.071	0.002
			(0.041)	(0.032)	(0.144)	(14.373)	(0.095)	(0.122)	(0.012)	(0.033)	(0.124)	(0.043)
Under-confidence	–	–	0.017	-0.040	-0.087	-5.400	-0.116***	-0.049	-0.038**	0.028	-0.047	-0.035*
			(0.022)	(0.033)	(0.107)	(7.458)	(0.042)	(0.055)	(0.018)	(0.039)	(0.075)	(0.021)
Average of main:												
Over-confidence	–	–	0.057	0.023	0.031	1.166	0.036	0.110	0.005	-0.023	0.031	-0.016
			(0.036)	(0.028)	(0.177)	(15.192)	(0.117)	(0.120)	(0.016)	(0.030)	(0.107)	(0.038)
Under-confidence	–	–	0.012	-0.054	-0.062	-6.511	-0.186***	-0.084	-0.051**	0.030	-0.133	-0.019
			(0.023)	(0.036)	(0.113)	(7.963)	(0.057)	(0.065)	(0.023)	(0.041)	(0.082)	(0.022)

Table A16: Robustness to definitions of confidence (continued)

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earn) (11)	Unempl (12)
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Main measure	2.827***	0.128	0.018*	0.033***	0.041	3.631	0.078***	0.022	0.018***	0.001	0.059**	-0.023**
	(0.569)	(0.580)	(0.010)	(0.011)	(0.046)	(3.417)	(0.023)	(0.025)	(0.006)	(0.012)	(0.029)	(0.009)
Percentiles	2.631***	0.356	0.016**	0.027***	0.047	3.814	0.062***	0.006	0.017***	-0.005	0.045*	-0.012
	(0.528)	(0.505)	(0.008)	(0.010)	(0.038)	(2.803)	(0.019)	(0.020)	(0.006)	(0.011)	(0.025)	(0.008)
Empirical dist	2.525***	0.318	0.018**	0.027***	0.027	2.517	0.065***	0.018	0.016***	0.006	0.050**	-0.019**
	(0.487)	(0.497)	(0.008)	(0.009)	(0.039)	(2.895)	(0.020)	(0.022)	(0.006)	(0.010)	(0.025)	(0.008)
Main of averages	–	–	0.016	0.019*	-0.048	-2.972	0.042*	0.047*	0.012*	-0.007	0.001	-0.009
			(0.010)	(0.011)	(0.048)	(3.633)	(0.025)	(0.026)	(0.007)	(0.012)	(0.031)	(0.010)
Average of main	–	–	0.017*	0.018	-0.038	-2.434	0.047*	0.040	0.010	-0.008	0.012	-0.009
			(0.010)	(0.011)	(0.048)	(3.609)	(0.025)	(0.026)	(0.007)	(0.012)	(0.031)	(0.010)

Notes: This table presents the robustness of our main results to changing our definitions of our main measures of confidence. All regressions estimate our main specification in the even-numbered columns of Tables 3, 4, and 5, replacing the main measure of over-confidence and under-confidence or the main measure of degrees of confidence with alternate definitions. Sample sizes for each regression are the same as in each main table. In Panel A, we iterate over our definitions of binary over- and under-confidence variables. Each pair of over- and under-confidence measures are estimated in the same regression. The definitions labeled 'more strict' or 'less strict' change the self-assessment and percentile score cutoffs in our main measure. The relative measure identifies children who score in the top or bottom 25 percent of test scores within each self-assessment bucket as under- or over-confident, respectively. The 'continuous tails' measure identifies over-confident children as those whose more continuous measure of confidence is between 3 and 6, and under-confident children as those whose continuous measure of confidence is between -6 and -3. Finally, the last two measures combine data over the two waves of the CDS where we observe confidence measurements, when available. The first averages test scores and self-reports over the two waves and then applies our main cutoffs, and the second averages the main measure over the two waves. In Panel B, we iterate over our definitions of the more continuous measure of biased beliefs. The one labeled 'percentiles' differences the percentile of children's self-assessment and their percentile score, and the one labeled 'empirical dist' assumes that children knew the empirical distribution of self-reports and should have correspondingly reported their self-assessments (instead of assuming a uniform distribution). Again, the last two measures combine data over the two waves were possible: the first averages test scores and self-reports over the two waves and then applies the transformation to the same scale, and the second averages the main measure over the two waves. Further iteration is presented in specification charts for each outcome, found in Appendix Figures A5-A16. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A17: Sample means by whether missing confidence variables in the CDS

	Sample	Non-Sample	*p-value*
	(1)	(2)	*(1)=(2)*
Panel A: Demographic Characteristics			
Child is female	0.497	0.458	0.089
	(2985)	(578)	
Child is white	0.458	0.481	0.303
	(2985)	(578)	
Child is black	0.417	0.370	0.036
	(2985)	(578)	
Child is hispanic	0.075	0.076	0.928
	(2985)	(578)	
Child's birth order	1.625	1.521	0.067
	(2645)	(457)	
Child's birth year	1990.020	1991.478	0.000
	(2985)	(573)	
Panel B: Parent and Family Characteristics			
Father at least graduated high school	0.835	0.811	0.256
	(1882)	(365)	
Father at least has bachelors	0.257	0.260	0.880
	(1883)	(365)	
Mother at least graduated high school	0.817	0.735	0.000
	(2657)	(499)	
Mother at least has bachelors	0.134	0.123	0.490
	(2648)	(496)	
Total taxable family income (thous 2016 USD)	69.777	61.977	0.032
	(2985)	(578)	
HH lives in public housing	0.061	0.076	0.178
	(2985)	(577)	
HH receives food stamps	0.198	0.213	0.404
	(2985)	(577)	
Two adults in HH	0.645	0.666	0.329
	(2985)	(578)	
Panel C: Other Child Characteristics			
Child ever in gifted prog	0.243	0.077	0.000
	(2893)	(568)	
Child ever in special ed prog	0.127	0.067	0.000
	(2888)	(568)	
Child has repeated grade	0.122	0.033	0.000
	(2921)	(568)	
Child qualifies for FRP lunch	0.598	0.510	0.005
	(2220)	(288)	
Parent's rating of child health	0.014	0.025	0.802
	(2969)	(572)	
# Siblings in the HH	1.415	1.178	0.000
	(2985)	(578)	
Big 5 personality scores (indices)			
Conscientiousness	-0.003	0.014	0.754
	(2964)	(153)	
Extroversion	0.005	0.060	0.282
	(2970)	(154)	
Neuroticism	-0.006	0.040	0.374
	(2957)	(152)	
Agreeableness	-0.003	0.025	0.605
	(2962)	(153)	
Openness to experiences	-0.001	-0.069	0.102
	(2955)	(151)	
Panel D: Child Ability Measures			
Math score percentile	58.477	49.256	0.000
	(2985)	(156)	
Reading score percentile	55.386	52.587	0.449
	(2973)	(63)	
Digit span score	14.246	7.403	0.000
	(2863)	(149)	

Notes: This table regresses an indicator for whether a child is in our final sample on child characteristics. The sample is all 3563 children in the CDS survey. 578 children are dropped from our analysis sample. These are children for whom we never observe *both* a self-assessed and observed ability measure. Of those, 99 percent are missing a self-assessed measure and 73 percent are missing a math test score.

Table A18: Correlations between math confidence and other attitudes

	Over-Conf	Under-Conf	Math Confidence (sd)	General Conf
Panel A: Other Math Attitudes				
Math skill relative to peers	0.275	-0.307	0.534	0.242
Expected performance in math this year	0.189	-0.219	0.403	0.209
How good at learning new thing in math	0.152	-0.190	0.316	0.242
How easy is math for you	0.078	-0.095	0.166	0.044
How useful is what you learn in math	0.055	-0.056	0.130	0.207
Being good in math is important	0.063	-0.057	0.146	0.199
Working on math is interesting	0.157	-0.120	0.291	0.167
How much do you like math	0.240	-0.189	0.404	0.142
Panel B: Social and School Performance				
Do you feel like part of your school	-0.000	-0.036	0.054	0.213
Do you feel close to people at your school	0.001	-0.051	0.049	0.247

Note: This table shows the partial correlations between over- and under-confidence in math and general confidence and children's other attitudes towards math and social experiences at school after partialling out the relationship with math test score deciles.

Table A19: Childhood math confidence and average employment outcomes from age 28-33

Dependent variable:	Works in STEM (non-health)		Non-STEM high-educ occ.		Ln(Earnings)		Unemployed this year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Independent variables are binary measures of over- and under-confidence								
Over-confidence	0.011	0.015	-0.034	-0.041	0.122	0.132	-0.031	-0.035
	(0.017)	(0.017)	(0.030)	(0.032)	(0.109)	(0.109)	(0.035)	(0.035)
Under-confidence	-0.039**	-0.040**	0.022	0.020	-0.027	-0.030	0.025	0.021
	(0.019)	(0.019)	(0.039)	(0.039)	(0.073)	(0.075)	(0.022)	(0.022)
N	1301	1301	1301	1301	1269	1269	1364	1364
OC = -1*UC? *p-value*:	0.262	0.318	0.803	0.687	0.463	0.436	0.881	0.735
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units								
Confidence	0.017**	0.017**	0.003	0.004	0.084**	0.089**	-0.024*	-0.024*
	(0.007)	(0.007)	(0.013)	(0.014)	(0.038)	(0.039)	(0.013)	(0.013)
N	1301	1301	1301	1301	1269	1269	1364	1364
Sample mean of dep. var.	0.043		0.167		10.227		0.141	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓

Notes: This table regresses employment outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A20: Comparing predictiveness of biased beliefs and Big-Five traits

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality ind. (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earnings) (11)	Unempl (12)
Panel A: Independent variables are binary measures of over- and under-confidence												
Math over-confidence	2.485*	-0.440	0.061**	0.023	-0.004	0.517	0.053	0.062	0.013	-0.033	0.072	-0.033
	(1.484)	(1.387)	(0.026)	(0.025)	(0.147)	(12.009)	(0.102)	(0.087)	(0.017)	(0.027)	(0.086)	(0.030)
Math under-confidence	-5.633***	0.473	0.023	-0.055**	-0.107	-10.609*	-0.155***	0.021	-0.049***	0.025	-0.078	0.008
	(1.495)	(1.465)	(0.017)	(0.028)	(0.081)	(5.941)	(0.038)	(0.048)	(0.016)	(0.032)	(0.057)	(0.017)
Conscientiousness	-0.944	-0.026	0.021	0.051***	0.130**	6.120	0.009	0.037	0.002	0.038**	0.030	-0.004
	(0.913)	(0.924)	(0.014)	(0.015)	(0.064)	(4.777)	(0.039)	(0.037)	(0.010)	(0.016)	(0.045)	(0.014)
Extroversion	-0.109	0.020	0.001	0.007	-0.069	-3.263	-0.057**	0.011	-0.016**	0.012	0.029	-0.019*
	(0.712)	(0.695)	(0.011)	(0.012)	(0.053)	(4.101)	(0.029)	(0.029)	(0.007)	(0.013)	(0.033)	(0.010)
Agreeableness	-0.611	-0.342	0.039**	0.020	-0.052	-4.402	-0.041	0.039	-0.006	0.002	0.063	-0.002
	(1.086)	(1.084)	(0.018)	(0.017)	(0.072)	(5.848)	(0.051)	(0.045)	(0.011)	(0.019)	(0.048)	(0.015)
Openness	-0.789	0.706	-0.036**	-0.036**	-0.035	-0.270	0.021	-0.078**	-0.019**	-0.027*	-0.023	0.036***
	(0.922)	(0.950)	(0.014)	(0.015)	(0.063)	(4.818)	(0.038)	(0.034)	(0.009)	(0.016)	(0.042)	(0.013)
Neuroticism	-0.640	0.549	0.015	0.051***	0.019	-1.238	-0.069**	0.016	-0.005	0.009	-0.084*	0.006
	(0.933)	(0.964)	(0.015)	(0.017)	(0.066)	(4.904)	(0.032)	(0.034)	(0.008)	(0.019)	(0.045)	(0.014)
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units												
Math confidence	2.775***	0.080	0.014	0.027**	0.030	3.262	0.076***	0.014	0.019***	-0.004	0.054*	-0.023**
	(0.568)	(0.581)	(0.010)	(0.011)	(0.046)	(3.442)	(0.025)	(0.025)	(0.006)	(0.012)	(0.028)	(0.009)
Conscientiousness	-1.007	-0.042	0.021	0.051***	0.128**	5.933	0.004	0.036	0.003	0.037**	0.030	-0.003
	(0.914)	(0.923)	(0.014)	(0.016)	(0.064)	(4.749)	(0.040)	(0.038)	(0.010)	(0.016)	(0.045)	(0.014)
Extroversion	-0.191	0.017	0.000	0.007	-0.070	-3.356	-0.059**	0.010	-0.016**	0.013	0.028	-0.018*
	(0.713)	(0.697)	(0.011)	(0.012)	(0.053)	(4.100)	(0.030)	(0.029)	(0.007)	(0.013)	(0.032)	(0.010)
Agreeableness	-0.684	-0.315	0.039**	0.020	-0.053	-4.516	-0.042	0.039	-0.007	0.003	0.061	-0.002
	(1.086)	(1.082)	(0.018)	(0.017)	(0.072)	(5.796)	(0.051)	(0.045)	(0.012)	(0.019)	(0.048)	(0.015)
Openness	-0.718	0.715	-0.036**	-0.034**	-0.035	-0.246	0.020	-0.076**	-0.018**	-0.028*	-0.023	0.035***
	(0.925)	(0.950)	(0.014)	(0.015)	(0.062)	(4.826)	(0.038)	(0.033)	(0.009)	(0.016)	(0.042)	(0.013)
Neuroticism	-0.700	0.552	0.015	0.051***	0.019	-1.215	-0.069**	0.017	-0.005	0.009	-0.085*	0.006
	(0.935)	(0.964)	(0.015)	(0.017)	(0.066)	(4.924)	(0.032)	(0.034)	(0.008)	(0.019)	(0.045)	(0.014)

Notes: This table presents the estimates from Panel A of Table A13 but includes the coefficients on the childhood personality measures. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A21: Math confidence and young adult social outcomes

Dependent variable:	In a romantic relationship		Mental health		Social anxiety		Drinks alcohol often		Dangerous behavior	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Independent variables are binary measures of over- and under-confidence										
Over-confidence	-0.060**	-0.063**	-0.038	-0.023	-0.040	-0.043	0.020	0.019	0.048	0.051
	(0.026)	(0.026)	(0.048)	(0.048)	(0.055)	(0.055)	(0.014)	(0.014)	(0.043)	(0.044)
Under-confidence	-0.004	-0.001	-0.034	-0.036	-0.001	-0.000	0.013	0.015	-0.022	-0.022
	(0.023)	(0.023)	(0.034)	(0.034)	(0.044)	(0.044)	(0.015)	(0.015)	(0.029)	(0.029)
N	10389	10389	10360	10360	10374	10374	10360	10360	10277	10277
OC = -1*UC? *p-value*:	0.064	0.066	0.218	0.314	0.553	0.535	0.113	0.102	0.621	0.584
Panel B: Independent variable is degrees of over- and under-confidence in standard deviation units										
Confidence	-0.005	-0.005	0.008	0.010	-0.016	-0.015	-0.001	-0.002	0.018	0.018
	(0.011)	(0.010)	(0.017)	(0.017)	(0.021)	(0.021)	(0.006)	(0.006)	(0.015)	(0.015)
N	10389	10389	10360	10360	10374	10374	10360	10360	10277	10277
Sample mean of dep. var.	0.551		0.001		0.000		0.105		-0.002	
Basic controls:	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Added background controls:		✓		✓		✓		✓		✓

Notes: This table regresses social (placebo) outcomes outcomes on childhood biased beliefs with various controls. Biased beliefs are measured in the earlies observed wave in the CDS with non-missing test scores and self-assessed ability. In Panel A, the outcome is regressed on an indicator for over-confidence, an indicator for under-confidence and our basic set of controls (in odd-numbered columns) and our extended set of controls (in even-numbered columns). The p-value listed tests whether the coefficient on the over-confidence indicator is equal to -1 times the coefficient on the under-confidence indicator. In Panel B, the outcome is regressed on our more continous measure of biased beliefs which has been standardized to have mean zero and standard deviation one in our sample and the same sets of controls. All controls are the same as described in Table 1, minus the controls for adolescent test score deciles. Basic controls also include year fixed effects when the outcome is observed in a panel. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

Table A22: Heterogeneity by over- and under-confidence using the degrees of confidence measure

	Math Score (1)	Reading Score (2)	HS grad (3)	College grad (4)	College quality (5)	College math SAT 75p (6)	STEM major (7)	Grad degree (8)	STEM occup (9)	High-educ occup (10)	ln(Earnings) (11)
Math confidence (sd)	3.100***	0.905	0.022	0.047**	-0.002	-0.714	0.015	0.009	0.012	0.034*	0.100**
	(1.042)'	(1.018)	(0.016)	(0.019)	(0.065)	(4.979)	(0.039)	(0.037)	(0.010)	(0.020)	(0.042)
Confidence (sd)*Over	-1.662	-3.046	0.011	-0.023	-0.091	-6.059	-0.193	-0.232*	-0.001	-0.055*	-0.198
	(1.875)'	(1.889)	(0.046)	(0.035)	(0.234)	(18.494)	(0.142)	(0.132)	(0.024)	(0.030)	(0.146)
Confidence (sd)*Under	-0.254	-1.621	-0.067	0.008	-0.279	-12.113	0.108	-0.097	0.010	-0.133	-0.033
	(3.621)'	(4.425)	(0.041)	(0.084)	(0.277)	(17.160)	(0.089)	(0.159)	(0.030)	(0.097)	(0.176)
Math over-confidence	-0.389	5.185*	-0.010	-0.016	0.225	16.374	0.346	0.472**	-0.002	-0.003	0.226
	(2.999)'	(3.074)	(0.069)	(0.052)	(0.334)	(27.290)	(0.218)	(0.227)	(0.034)	(0.042)	(0.220)
Math under-confidence	-3.173	0.584	-0.091	0.015	-0.620	-33.718	0.013	-0.172	-0.018	-0.193	0.014
	(6.610)'	(7.752)	(0.072)	(0.147)	(0.471)	(29.471)	(0.144)	(0.270)	(0.058)	(0.171)	(0.302)
N	1747	1745	2714	2725	1107	1117	736	810	4592	4592	4423
$Slope_{Over} = Slope_{Under}$	0.712	0.760	0.181	0.727	0.589	0.802	0.065	0.500	0.762	0.428	0.454
$Slope_{Over} = Slope_{Under} = Slope_{Neither}$	0.674	0.266	0.233	0.801	0.574	0.752	0.171	0.192	0.943	0.101	0.396

Notes: This table estimates our main specification for our more continuous degrees of confidence measure in standard deviation units (found in the even-numbered columns of Panel B in Tables 3, 4, and 5), but adds indicators for being over- and under-confident according to this measure and the interactions between the degrees of confidence measure and the indicators. Any student whose difference between their self-assessed bin from 1-7 and the bin they should have reported given their test scores is greater than two is considered over-confident and any student whose difference is less than negative two is considered under-confident. Standard errors are clustered at the family level and included in parentheses below each estimate. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 percent level, respectively.

B Appendix B: Constructing Indices

College quality measures:

Using restricted data from the TAS, we identify the first college that each child in our sample attended, if they attended college. We then match those schools to college quality data from the first year that they attended the college. Following Cohodes and Goodman (2014), we construct an index of college quality by taking the first component from a principal component analysis of colleges' 75th-percentile math SAT scores among incoming freshmen, graduation rates, and per-pupil instructional expenditures, separately by year from 2005-2019.

We impute SAT scores where possible to increase our sample size; some schools report 75th-percentile math ACT scores but not SAT scores. For those schools, we impute 75th-percentile math SAT scores as predicted values from a regression of 75th-percentile SAT scores on 75th-percentile ACT scores among schools with both measures. We also use the 6-year graduation rate rather than the 4-year graduation rate because the 6-year rate is available for more schools and the two measures are highly correlated.

Depending on the year, the first principal component captures 70-80 percent of the variation between these three variables and assigns nearly equal weights to all three variables in all years. We standardize the first principal component to have mean 0 and standard deviation 1 in the full sample of four-year colleges in the US by year, and call this our college quality index.

Secondary outcome variables:

To minimize the number of outcomes and controls include in our analysis, we create many indices of similar variables. Here we list each index with its underlying variables. All underlying variables are scales from 1-3, 1-4, 1-5, or 1-7. When the variable applies to parents, we standardize by year before taking the average. When the variable applies to children, we standardize by year and age group (8-11, 12-14, and 15-19).

Adult math confidence:

- How good would you be in a career that required you to use math?

- How good would you be in a career that required you to use physical science or technology?

Adult reading confidence:

- How good would you be in a career that required you to read and write a lot?

Adult general academic confidence:

- How good would you be in a career that required you to be creative?
- How good are you at solving problems you encounter?
- How good are you at logical, analytic thinking?
- How intelligent are you, compared to others?
- How good are you at listening to and understanding others?
- How good are you at teaching and explaining to others?

Adult career confidence:

- How successful do you think you could be in the type of job you most want?
- How likely do you think you are to end up in the job you most want at age 30?

Adult general confidence: (variables marked with a * are flipped so that a higher score means more confident)

- How confident are you, compared with others?
- How decisive are you, compared with others?
- How independent are you, compared with others?
- How good are you at being a leader?
- How good are you at supervising others?
- How often do you feel discouraged about the future*
- How often, in the last month, did you feel that you had something important to contribute to society?
- How often, in the last month, did you feel good at managing responsibilities of daily life?
- How often, in the last month, did you feel confident to think or express your own ideas and opinions?

- How often, in the last month, did you feel that your life had a direction or purpose?

Adult Big-Five personality measures: (variables marked with a * are flipped so that a higher score means more conscientious, agreeable, etc.)

- Conscientiousness:
  - You are someone who does a thorough job. Does this describe you not at all, a little, some, or a lot?
  - You are someone who tends to be lazy. Does this...?*
  - You are someone who does things efficiently. Does this...?
- Agreeableness:
  - You are someone who is sometimes rude to others. Does this...?*
  - You are someone who has a forgiving nature. Does this...?
  - You are someone who is considerate and kind to almost everyone. Does this...?
- Extroversion:
  - You are someone who is talkative. Does this...?
  - You are someone who is outgoing, sociable. Does this...?
  - You are someone who is reserved. Does this...?*
- Neuroticism:
  - You are someone who worries a lot. Does this...?
  - You are someone who gets nervous easily. Does this...?
  - You are someone who is relaxed, handles stress well. Does this...?*
- Openness to experience:
  - You are someone who is original, comes up with new ideas. Does this...?
  - You are someone who values artistic experiences. Does this...?
  - You are someone who has an active imagination. Does this...?

Dangerous behavior index:

- How often in the last 6 months did you do something you knew was dangerous just for the thrill of it?
- How often in the last 6 months did you damage public or private property?
- How often in the last 6 months did you get into a physical fight?
- How often in the last 6 months did you drive when you were drunk or high on drugs?
- How often in the last 6 months did you ride with a driver who had too much to drink?

Control variables:

Child general confidence:

- Does the statement never, sometimes, always apply to you ... I do things as well as most people
- ... When I do something, I do it well
- ... I'm as good as most other people
- ... A lot of things about me are good
- ... I have a lot to be proud of

Child Big-Five personality measures, reported by primary caregiver: (variables marked with a * are flipped so that a higher score means more conscientious, agreeable, etc.)

- Conscientiousness:
  - According to [child's] behavior, [he/she] cheats or tells lies*
  - ... [he/she] has difficulty concentrating, cannot pay attention for long*
  - Thinking about [child], tell me if [child] waits [his/her] turn in games and other activities
  - ... tell me if [child] does neat, careful work
  - ... tell me if [child] usually does what you tell [him/her] to do
- Agreeableness:
  - ... [he/she] argues too much*
  - ... [he/she] bullies or is cruel or mean to others*
  - ... [he/she] is disobedient*
  - ... [he/she] has trouble getting along with other children*

  - ... [he/she] is stubborn, sullen, or irritable*
  - ... [he/she] breaks things on purpose or deliberately destroys [his/her] own or another's things*
  - ... tell me if [child] is cheerful, happy
  - ... tell me if [child] gets along well with other children
  - ... tell me if [child] is admired and well-liked by other children

- Extroversion:
  - ... [he/she] is withdrawn, does not get involved with others*
  - ... [he/she] demands a lot of attention

- Neuroticism:
  - ... [he/she] has sudden changes in mood or feeling
  - ... [he/she] is rather high strung, tense and nervous
  - ... [he/she] is too fearful or anxious
  - ... [he/she] has a lot of difficulty getting [his/her] mind off certain thoughts
  - ... [he/she] feels others are out to get [him/her]
  - ... [he/she] worries too much
  - ... tell me if [child] can get over being upset quickly*

- Openness to experience:
  - ... [he/she] is impulsive, or acts without thinking
  - ... [he/she] clings to adults*
  - ... [he/she] hangs around with kids who get into trouble
  - ... tell me if [child] is curious and exploring, likes new experiences

Parent adherence to traditional gender norms: (variables marked with a * are flipped so that a higher score means more traditional gender norms)

- Most of the important decisions in the life of the family should be made by the man of the house
- Women are much happier if they stay at home and take care of their children
- There is some work that is men's and some that is women's and they should not be doing each other's

- It is much better for everyone if the man earns the living and the woman takes care of the home and family
- It is more important for a wife to help her husband's career than to have one herself
- Preschool children are likely to suffer if their mother is employed
- An employed mother can establish as warm and secure a relationship with her children as a mother who is not employed*
- Parents should encourage just as much independence in their daughters as their sons*
- A father should be as heavily involved in the care of his child as the mother*
- If a husband and wife both work full-time, they should share household tasks equally*

Parent aggravation in parenting:

- Thinking about [child], there are some things that [he/she] does that really bother me a lot
- ... I find myself giving up more of my life to meet [child's] needs than I ever expected
- ... I often feel angry with [child]
- Thinking about my child[ren], being a parent is harder than I thought it would be
- ... I feel trapped by my responsibilities as a parent
- ...I find that taking care of my child[ren] is much more work than pleasure
- ...I often feel tired, worn out, or exhausted from raising a family

Parent self-esteem: (variables marked with a * are flipped so that a higher score means higher self-esteem)

- I feel that I'm a person of worth, at least on an equal basis with others
- I feel that I have a number of good qualities
- All in all, I am inclined to feel that I am a failure*
- I am able to do things as well as most other people
- I feel I do not have much to be proud of*
- I take a positive attitude toward myself
- On the whole, I am satisfied with myself
- I wish I could have more respect for myself*

- I certainly feel useless at times*
- At times I think I am no good at all*

Parent self-efficacy: (variables marked with a * are flipped so that a higher score means higher self-efficacy)

- There is really no way I can solve some of the problems I have*
- Sometimes I fell that I'm being pushed around in life*
- I have little control over the things that happen to me*
- I can do just about anything I really set my mind to
- I often feel helpless in dealing with the problems of life*
- What happens to me in the future most depends on me
- There is little I can do to change many of the important things in my life*

C Biased Beliefs and Other Attitudes Towards School

Our confidence measures consistently correlate with children's other attitudes towards math in ways we would expect. Appendix Table A18 shows the pairwise correlations between children's attitudes towards math and school, our measures of over- and under-confidence in math, and the general confidence index. Over-confidence in math is positively correlated with children's self-assessed ability relative to their peers, their expected performance in math that year, how good they think they are at learning a new skill in math, how interesting they think math is, and how much they like math ($\rho \in [0.15, 0.28]$). All of the same correlations are negative and of similar magnitude for children who are under-confident in math. The correlations between over- and under-confidence and how easy, useful, or important math is are much smaller in magnitude but have the expected signs. There are very similar patterns using the more continuous measure of math confidence. General confidence is also positively correlated with these attitudes ($\rho \in [0.14, 0.24]$), except for how easy a child thinks math is.

On the other hand, whether children report feeling like part of their school community or close to their peers are both uncorrelated with math over- or under-confidence ($\rho < |0.05|$), but are positively correlated with our index of general confidence ($\rho \approx 0.23$). Together, these patterns suggest that our measures isolate over- and under-confidence in the particular domain of math, but our regressions also control for general confidence and other measures of child ability to further isolate the relationship between children's biased beliefs about their math ability and their medium- and long-run outcomes.

D Over- versus under-confidence

One ex-ante strength of our binary measures of biased beliefs is that they offer a clear way to test whether over- and under-confidence correlate with later-life outcomes with symmetric magnitudes; we display p-values for all of these comparisons at the bottom of Panel A in Tables 3, 4, and 5. In practice, we find that the coefficient magnitudes for over- and under-confidence are only significantly different for two of our twelve outcomes: high-school graduation and working in STEM. Over-confidence predicts high-school graduation significantly more strongly than does under-confidence, while only under-confidence predicts working in STEM.

We also test for heterogeneity in the direction of biased beliefs using our more continuous measure of degrees of confidence. In Appendix Table A22, we allow the coefficient on this measure to differ by whether a child is over-confident (assessing one's ability at least 3 bins, or 42 percentiles, too high), under-confident (assessing one's ability at least 3 bins too low), or neither. We cannot reject that the slope of the outcome with respect to the degrees of confidence variable is equal across these groups for any outcome, though we are likely under-powered to do so. This result supports the functional-form assumptions we make in Panel B of each of our main tables, where degrees of confidence enter linearly for all outcomes. More broadly, these results and those using our binary measures of over- and under-confidence suggest that over- and under-confidence largely predict similarly-sized, oppositely-signed gaps in long-term educational and employment outcomes.

E Measuring key confounders

Big-Five Personality Traits

In Section 7.1, we show that our results are robust to controlling for children's Big-Five personality traits. The CDS did not measure these traits using standard psychometric scales, so we approximate them using parents' reports of child behavior. See Appendix B for the variables that make up the index for each trait.

While our proxies for these traits may be noisy, they do correlate with other variables in expected ways. First, the TAS did collect standard psychometric scales to measure Big-5 traits among young adults, and our childhood measures correlate with these adult measures at levels similar to other estimates of the longitudinal persistence of the Big-Five traits (Hampson and Goldberg, 2006; Edmonds et al., 2013). The intercorrelations of our childhood Big-Five personality measures are also broadly similar to those found in studies that use more standard scales to measure these traits (van den Akker et al., 2014; Soto, 2016). Finally, if we regress contemporaneous math and reading cognitive test scores on our childhood Big-Five measures while controlling for IQ, race, and gender, the coefficients on the Big-Five characteristics follow similar patterns as those reported in Almlund et al. (2011) (results available upon request).

We also consider the extent to which the Big-Five traits predict long-term outcomes in our data. We present the coefficients on each personality trait in the specifications above in Appendix Table A20. Some correlations are consistent with prior estimates of the contemporaneous links between personality and economic outcomes (Almlund et al., 2011; Heckman et al., 2019), but we find fewer significant relationships than expected. These null results may reflect noise in our constructed measures of personality, or they could reflect that childhood personality traits only moderately persist into adulthood (Hampson and Goldberg, 2006; Edmonds et al., 2013).

Parent and Teacher Beliefs and Investment

In Section 7.1, we also test that our main results are robust to controlling for measures of parents' and teachers' investments and beliefs. We construct these controls using data from the CDS. We measure caregiver investment from self-reports of how often they do certain activities with their child (e.g. do homework, play games), and we observe both caregiver and teacher reports of the level of educational attainment they expect the child to achieve. Our data also include teacher ratings of the student's academic, social, and physical competence on a scale from 1 (extremely competent) to 4 (not at all competent); we standardize these ratings by year and age group as a measure of teacher perceptions, which likely relate to teacher investment. See Appendix Table A6 for summary statistics on these variables.

We have relatively low data coverage for teacher reports because the CDS only interviewed elementary school teachers, while the CDS sample includes older children, and because questionnaires were mailed to teachers and had relatively low response rates. In total, 54 percent of our final sample had a teacher respond in any wave of the CDS. We observe teacher predictions of educational attainment in the same year in which we observe biased beliefs for 34 percent of our sample, and we observe teacher reports of student competence for 20 percent of our sample (this variable was only recorded in the 1997 CDS). In contrast, we observe caregiver reports of investment and predicted educational attainment for more than 99 percent of our sample.

These measures of teacher and parent beliefs and investments correlate with children's beliefs in math in our sample, making them potential confounders of the main associations we estimate. Appendix Table A15 regresses childhood over- and under-confidence in math on our variables for teacher perceptions and expectations, parent investment and expectations, and child test scores. First, teacher expectations of educational attainment predict children's biased beliefs: children that teachers think are going to get a graduate degree are more confident, and in particular are less likely to be under-confident. Next, parent investment predicts children's under-confidence but not over-confidence: children whose parents read or do homework with them more than once per week are (marginally significantly) more likely to be under-confident in math, whereas we find suggestive evidence that children with parents who play sports or games with them are less likely to be under-confident. Similar to the results for teacher expectations of educational attainment, children whose parents think they are likely to get a graduate degree are less likely to be under-confident.

Overall, these results show that our measures of children's over- and under-confidence in math are correlated with parent and teacher beliefs and investment in largely expected ways, even when we control for children's ability and general confidence. This suggests that one mechanism through which childhood over- and under-confidence could relate to long-term outcomes could be through parent and teacher behavior. However, adding controls for these adult beliefs and behaviors does not change the relationship between children's over- and under-confidence and long-run outcomes – if anything, children's biased beliefs become *more* predictive of long-run outcomes when we condition on these variables.

School quality when confidence is measured

Finally, Section 7.1 tests that our results are robust to controlling for the quality of the school that children were attending when we observe their first measures of over- and under-confidence in math. We match respondents with school IDs using restricted data from the CDS.

Then, we collect data on free or reduced-price lunch and student-teacher ratios from the NCES, while we collect data on testing achievement from the Stanford Education Data Archive (SEDA; Fahle et al., 2021). The measures are scaled relative to national grade- and subject-specific test score distributions. SEDA's data for school test scores pools data from 2009-2018 and is unavailable in earlier years. The students in our sample attended these schools in 1997, 2003, or 2007; we are forced to assume that relative school quality was similar in the decade before we observe testing data. 60 percent of our sample attends a school where we observe test scores in 2009-2019; 80 (50) percent of students attend a school where we observe the student-teacher ratio (percent FRPL) in the year in which we observe confidence. We also include an indicator for missing an NCES School ID in the CDS data.

F Alternate definitions of childhood biased beliefs

This section describes the alternate definitions of over-confidence, under-confidence, and more continuous degrees of confidence to which we test robustness in Section 7 above. Throughout the following definitions, p refers to children's score percentiles in math and r refers to children's self-reported math ability from 1 to 7. The names referring to each definition match those used in the specification charts given in Appendix Figures A4-A16.

Section A. Over-confidence:

1. Main measure:
   - Over-confident if $\begin{cases} p < 25 & r \in \{6,7\} \\ p < 50 & r = 7 \end{cases}$
2. Main - more strict (1):
   - Over-confident if $\begin{cases} p < 15 & r \in \{6,7\} \\ p < 40 & r = 7 \end{cases}$
3. Main - less strict (1):
   - Over-confident if $\begin{cases} p < 35 & r \in \{6,7\} \\ p < 60 & r = 7 \end{cases}$
4. Main - less strict (2):
   - Over-confident if $\begin{cases} p < 15 & r \in \{5,6,7\} \\ p < 40 & r \in \{6,7\} \end{cases}$
   - Estimated with under-confidence measure *Original - less strict*
5. Relative:
   - Over-confident if $p <$ the 25th percentile of people who report the same self-reported ability (r) in the same age bucket and if $r < 5$

6. Continuous tails (3 to 6):

    - Over-confident if *Main* degrees of confidence measure (Section C #1) ≥ 3

7. Main of averages:

    - Take average of first and second self-reported ability (r) and first and second percentile scores (p)
    - Apply cutoffs of *Main* measure (Section A #1) to these averages

8. Average of main:

    - Take the average of the first- and second-observed *Main* (Section A #1) over-confidence measures

Section B. Under-confidence

1. Main measure::

    - Under-confident if $\begin{cases} p > 75 & r \in \{1,2,3,4\} \\ p > 50 & r \in \{1,2,3\} \end{cases}$

2. Main - more strict (1):

    - Under-confident if $\begin{cases} p > 85 & r \in \{1,2,3,4\} \\ p > 60 & r \in \{1,2,3\} \end{cases}$

3. Main - more strict (2):

    - Under-confident if $\begin{cases} p > 85 & r \in \{1,2,3,4\} \\ p > 60 & r \in \{1,2,3\} \end{cases}$
    - Estimated with over-confidence measure *Original - more strict*

4. Main - less strict (1):

- Under-confident if $\begin{cases} p > 65 & r \in \{1,2,3,4\} \\ p > 40 & r \in \{1,2,3\} \end{cases}$

5. Relative:

   - Under-confident if $p >$ the 75th percentile of people who report the same self-reported ability (r) in the same age bucket and if $r > 3$

6. Continuous tails (-6 to -3):

   - Over-confident if *Main* degrees of confidence measure (Section C #1) $\leq$ -3

7. Main of averages:

   - Take average of first and second self-reported ability (r) and first and second percentile scores (p)
   - Apply cutoffs of *Main* measure (Section B #1) to these averages

8. Average of main:

   - Take the average of the first- and second-observed *Main* (Section B #1) under-confidence measures

Section C. Degrees of confidence

- Main measure:
  - Assume that kids with accurate beliefs would have reported $r^* = 1$ if $p \in \{1,14\}$, $r^* = 2$ if $p \in \{15,28\}$, ... $r^* = 7$ if $p \in \{86,100\}$.
  - Confidence measure is self-reported ability ($r \in \{1,...,7\}$) minus what they would have reported if they had accurate beliefs ($r^* \in \{1,...,7\}$). This variable has range -6 to 6.
- Percentiles:

- Convert empirical distribution of self-reports (r) into percentiles from 0 to 100 (p_r)
- Degree of confidence = $p_r - p$, or percentile of self-reported ability minus actual score percentile in our sample

- Empirical distribution:
  - Assume that the empirical distribution of self-reported ability is correct, but kids may be wrong about their place in it. In other words, if the full sample had accurate beliefs, the bottom 4% of scorers in our sample would report $r^* = 1$, the next 2% would report $r^* = 2$, and the top 22% of scorers would report $r^* = 7$. These values come from the empirical distribution of r, graphed in Figure 1.
  - Degree of confidence = $r - r^*$, or self-reported ability minus what children would have reported if they had accurate beliefs by this measure. This variable has range -6 to 6.

- Main of averages:
  - Take average of first and second self-reported ability (r) and first and second percentile scores (p)
  - Apply the same rule as the *Main* measure (Section C #1) to these averages

- Averages of main:
  - Take the average of the first- and second-observed *Main* measures (Section C #1) of degrees of confidence.

- **To make the specification charts, we standardize all of these measures of degrees of confidence to have mean 0 and standard deviation 1 in our analysis sample so that they can be compared on the same scale.**