Online Appendix

A.1 Supplementary Analyses for Children Born 1983 to 1986

In supplemental analyses, I estimated models using separate self-regulation problems and social problems subscales. Patterns of results were similar to those presented in the main text. Consistent with most psycho-biologists' notion of self-regulation (see Blair and Diamond 2008), the self-regulation problems subscale includes a measure of concentration that is excluded from the overall externalizing problems scale: "How frequently does your child have difficulty concentrating?" (This item is excluded from the overall externalizing problems scale to preserve comparability to other externalizing problems scales, which do not typically include concentration [Duncan et al. 2007; Peterson and Zill 1986]).

Although the patterns of associations between both subscales and years of schooling were similar to those of the overall externalizing problems scale, early self-regulation problems are ultimately stronger drivers of educational attainment than are social problems. This finding is consistent with work by Duncan and colleagues (2007), which finds that attention skills (closely related to self-regulation skills) are the behaviors most strongly correlated with adult educational attainment. Specifically, in my analysis of gender differences in the associations among NLSY-C children born 1983 to 1986, I find that the magnitude of the relationship between early self-regulation problems and years of schooling among boys was nearly 40 percent larger than that between social problems and years of schooling among boys. Additionally, the gender difference in the magnitudes of the association was roughly 5 times larger for self-regulation problems than for social problems. This was largely because higher self-regulation problems were associated with a larger decrease in years of schooling among boys than higher social problems (described above). But, it was also because higher self-regulation problems were not associated with as large of a decline in years of schooling among girls compared to higher social problems. Each unit increase in self-regulation problems was associated with a .075 years decrease in schooling among girls, whereas each unit increase in social problems was associated with a .082 years decrease in schooling, net of early childhood factors and demographic controls.

I also used supplemental analyses to address features of the model design, to see if one could include early behaviors or key mediators at three or more developmental stages. Measures like externalizing behaviors, PPVT, school/peer and home context, and school performance were collected at additional time points beyond those at which they are modeled. In this study, my focus is documenting gender patterns rather than comprehensively documenting precise magnitudes of correlation of behaviors across all developmental stages. However, I did encounter issues of power and multicollinearity when attempting to estimate a path model with more than two measures of early behaviors and mediators like home or school context. Given the current model's complexity, I was only able to include measures at one or two of the key developmental stages at which their importance has been highlighted in recent research.

I conducted a third set of supplemental analyses as part of the treatment of missing data. Imputed predictor variables with high missingness included parental conflict (missing for 35 percent of cases), childhood behaviors (18 percent missing, almost all were also missing on educational attainment), household income at age 4 (16 percent), and home environment at ages 10 to 11 (12 percent). Given the extent of item missingness due largely to sample attrition, sensitivity analyses investigated whether the treatment of missing data introduced bias into parameter estimates: (1) replication with complete cases only; (2) replication with a second set of 20 imputed datasets in which the variances of imputed items were increased by 10 percent to test whether introducing noise in the multiple imputation procedure would significantly weaken associations of interest (Royston 2004), and (3) assessment of systematic biases in item-missingness by regressing a binary indicator for missingness on the dependent variable on observed covariates. This tests whether the dependent variable is missing-at-random as a function of observed predictors. Substantive results did not change. Sample attrition by age 26 to 29 was slightly higher among children with higher PPVT scores at age 3 to 4 and who were low birth weight, and was slightly lower among children who were cared for outside the home at age 3 to 4.

Finally, because it is common for children in the United States to enter kindergarten at age 5—the earliest point at which behavior problems were measured for roughly half of the sample in this study (i.e., the 1983 and 1985 birth cohorts)—I restricted analyses to children born in 1984 and 1986 for whom behaviors at age 4 are available. Substantive results did not change. In addition, I compared patterns of gender differences for the NLSY-C children with age 5 behavior assessments to NLSY-C children whose behavior problems were assessed at age 4, as well as to children in the nationally representative sample of children in the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) observed at age 4. Patterns were overwhelmingly similar across ages and samples, indicating that exposure to school contexts for the children assessed at age 5 did not influence mean gender differences compared to children measured at age 4 prior to kindergarten entry. This points to persistent patterns of gender difference over the 15 years between the mid-1980s and the late 1990s.

A.2 Supplementary Analyses for Children Born 1983 to 1993 to Include Children Born to Older Mothers

In supplementary analyses addressing generalizability of results to older mothers at birth, who tend to be from higher-SES backgrounds, I replicated the decomposition but pooling children born 1983 to 1993 to mothers age 18 to 35 years at birth (rather than 18 to 29 years). Table A2.1 shows results for the subset of 1983 to 1993 cohorts born to a wider SES range of mothers (those who were age 18 to 35 at birth). Because the youngest children (those born in 1993) are only 19 by 2012, here I decompose schooling completed by ages 19 to 22 across all 1983 to 1993 birth cohorts. This can be thought of as capturing

nearly on-time high school completion and potential college enrollment.

When including children born to older (and higher-SES) mothers, the gender gap in schooling is smaller (.54 years versus .75 years) than among the subset of 1983 to 1986 cohorts who were born to younger mothers of lower-SES backgrounds (see Appendix Table A2.1 versus Table 3 of the main text). Both boys' and girls' mean externalizing problems scores are lower (row 2, columns 1 to 3 of Table A2.1). In addition, boys experience a buffer with respect to the effect of early externalizing problems on their early adult attainment (columns 4 and 5).

Interestingly, when looking at relatively on-time high school completion and college enrollment among girls ages 19 to 22, there is a larger, negative effect of early externalizing problems, despite the inclusion of children born to older mothers (presumably also from higher-SES backgrounds). Based on a supplementary decomposition of schooling completed by ages 19 to 22 among the 1983 to 1986 cohorts (not shown), I speculate that this reflects a broader pattern in which the effects of early externalizing are similar for girls and boys when considering relatively on-time high school completion and college enrollment by age 19 to 22 (rather than 26 to 29). However, to more directly distinguish effects of maternal age at birth from those of family SES, future work should directly model variation by mother's education or household income rather than assuming mother's age at birth is a direct proxy for social class.

When including children born to mothers up through age 35 at birth, many from higher-SES backgrounds, early externalizing problems account for a much smaller share of the gender gap in on-time high school completion and college enrollment by age 19 to 22. Early externalizing problems account for 15.9 percent of the .54 years gap, or 3.9 percent of observed gap-widening early childhood and demographic factors. Boys' higher levels of early externalizing problems contribute to the gap in high school completion and college enrollment by age 19 to 22. Only 30 percent of the gap (.026 of the .086 years contributed by early externalizing problems) is due to the larger negative effect of early externalizing problems on boys' schooling. Roughly 70 percent (.60 of the .086 years) is due to boys' higher levels of early externalizing problems.

In supplementary analyses not shown, I also fit logit models examining high school completion by age 22 as a binary outcome. Although the association between early externalizing problems and high school completion is statistically significant when examining boys born to younger mothers and older mothers together, the magnitude is only roughly one-third as large. Including children born to older mothers, many from higher-income families, the gender difference in the association is only one-tenth as large (a difference between boys and girls of roughly 18 percentage points when including older mothers versus roughly 2 percentage points when examining only children born to younger mothers).

Online Appendix Table A1.1: Detailed Results for Two-Way Decomposition of Gender Differences in Levels and Effects of Externalizing Problems and Family Factors on Years of Schooling at Ages 26 to 29 for NLSY-C Children Born 1983 to 1986 (Reference: Boys)

	Means		Means Diff	OLS Regression Coefficients				Contrib-ution of Diff in Levels	Contrib-ution of Diff. in Coeff-icients	Total Contribution of Levels & Coefficients	Prop. of Positive Effects on Gap	Prop. of Negative Effects on Gap
	(1) F	(2) M	(3) F-M	(4) F	Sig	(5) M	Sig	(6)	(7)	(8)	(9)	(10)
Mother Years of Schooling at Birth	12.139	12.088	0.051	0.353	***	0.298	***	0.015	0.665	0.680	0.248	0.000
Externalizing Problems, Ages 4-5	1.971	2.603	-0.632	-0.038		-0.169	***	0.107	0.341	0.448	0.163	0.000
Birth Order	1.784	1.791	-0.007	-0.235		-0.442	**	0.003	0.371	0.374	0.136	0.000
African-American	0.179	0.154	0.026	0.690	**	0.008		0.000	0.105	0.105	0.038	0.000
Received Care Outside Home, Ages 4-5	0.542	0.521	0.021	0.141		-0.006		0.000	0.077	0.076	0.028	0.000
Low Birth Weight	0.068	0.050	0.018	0.500		-0.142		-0.003	0.032	0.030	0.011	0.000
Child Std. Score on PPVT, Ages 3-4	0.247	0.185	0.062	0.414	***	0.437	***	0.027	-0.004	0.023	0.008	0.000
Home Environment, Ages 3-5	-0.055	-0.031	-0.025	-0.693	*	-0.058		0.001	0.019	0.021	0.008	0.000
Hispanic	0.075	0.081	-0.006	0.249		0.088		-0.001	0.013	0.012	0.005	0.000
Mother Age at Birth	23.820	23.865	-0.045	0.085		0.156	*	-0.007	-1.694	-1.702	0.000	0.855
Number of Children in Family, Ages 4-5	2.218	2.255	-0.037	0.162		0.232		-0.009	-0.158	-0.166	0.000	0.084
Father Absent at Child's Birth	0.193	0.179	0.014	-0.792	**	-0.219		-0.003	-0.103	-0.106	0.000	0.053
Household Income in $2011 (/$10,000), Child Ages 4-5	7.288	6.028	1.260	0.010	*	0.016		0.020	-0.036	-0.016	0.000	0.008
Constant	1.000	1.000	0.000	7.170	***	6.196	***	0.000	0.974	0.974	0.355	0.000
Observations (N)	881	780		881		780						
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap in Years of Schooling:								0.152	0.601	0.754	1.000	1.000
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap as a Proportion of Gap Driven by Levels vs. Effects:								0.202	0.798	1.000		

*** p<0.001, **p<0.01, * p<0.05, + p<0.10 (two-tailed *t*-tests for a statistically significant difference from 0). This model uses boys' coefficients as the reference when calculating each variable's contribution to the gap in schooling due to gender differences in mean levels and boys' means as the reference when calculating each variable's contribution due to gender differences in coefficients (i.e., effects).

Source: The 1983 to 1986 birth cohorts of the Children of the National Longitudinal Survey of Youth:1979 (NLSY-C; https://www.nlsinfo.org/content/cohorts/nlsy79-children) and matched National Longitudinal Survey of Youth:1979 (mother sample). The low-income white and military oversamples are excluded.

Note: The National Longitudinal Survey of Youth-Child Supplement (NLSY-C) consists of a nationally representative sample of children born to women age 14 to 21 in 1979; after excluding the poor white and military oversamples, the working sample in this study is restricted to the 1,857 children born 1983 to 1986, whose mothers were 18 to 29 years at birth. Children born 1983 to 1986 were born early enough to be age 26 to 29 as of the 2012 follow-up survey, but late enough to have early behavior problems information measured at age 4 to 5 beginning in 1986, at which point these items were introduced for children age 4 to 16. I used multiple imputation of 20 datasets to handle item-missingness. Model estimates use inverse-probability weighting to deal with stratified sample design (minority oversampling) and sample attrition by the 2012 follow-up wave (weights are described at: https://www.nlsinfo.org/weights/nlsy79). Once inverse-probability survey weights are applied, the working sample with complete attainment and behavior information drops from 1,857 to 1,661 children (881 girls, 780 boys).

Online Appendix Table A1.2: Detailed Results for Two-Way Decomposition of Gender Differences in Levels and Effects of Externalizing Problems and Family Factors on Years of Schooling at Ages 26 to 29 for NLSY-C Children Born 1983 to 1986 (Reference: Girls)

	Means		Means Diff	OLS Regression Coefficients				Contrib-ution of Diff. in Levels	Contrib-ution of Diff. in Coeff-icients	Total Contribution of Levels & Coefficients	Prop. of Positive Effects on Gap	Prop. of Negative Effects on Gap
	(1) F	(2) M	(3) M-F	(4) F	Sig	(5) M	Sig	(6)	(7)	(8)	(9)	(10)
Mother Years of Schooling at Birth	12.139	12.088	-0.051	0.353	***	0.298	***	0.018	0.668	0.686	0.257	0.000
Birth Order	1.784	1.791	0.007	-0.235		-0.442	**	0.002	0.369	0.371	0.139	0.000
Externalizing Problems, Ages 4-5	1.971	2.603	0.632	-0.038		-0.169	***	0.024	0.258	0.282	0.106	0.000
African-American	0.179	0.154	-0.026	0.690	**	0.008		0.018	0.122	0.140	0.052	0.000
Received Care Outside Home, Ages 4-5	0.542	0.521	-0.021	0.141		-0.006		0.003	0.080	0.083	0.031	0.000
Low Birth Weight	0.068	0.050	-0.018	0.500		-0.142		0.009	0.044	0.053	0.020	0.000
Home Environment, Ages 3-5	-0.055	-0.031	0.025	-0.693	*	-0.058		0.017	0.035	0.052	0.020	0.000
Child Std. Score on PPVT, Ages 3-4	0.247	0.185	-0.062	0.414	***	0.437	***	0.026	-0.006	0.020	0.007	0.000
Hispanic	0.075	0.081	0.006	0.249		0.088		-0.002	0.012	0.010	0.004	0.000
Household Income in $2011 (/$10,000), Child Ages 4-5	7.288	6.028	-1.260	0.010	*	0.016		0.013	-0.044	-0.031	0.000	0.015
Mother Age at Birth	23.820	23.865	0.045	0.085		0.156	*	-0.004	-1.691	-1.695	0.000	0.844
Number of Children in Family, Ages 4-5	2.218	2.255	0.037	0.162		0.232		-0.006	-0.155	-0.161	0.000	0.080
Father Absent at Child's Birth	0.193	0.179	-0.014	-0.792	**	-0.219		-0.011	-0.110	-0.121	0.000	0.060
Constant	1.000	1.000	0.000	7.170	***	6.196	***	0.000	0.974	0.974	0.365	0.000
Observations (N)	881	780		881		780						
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap in Years of Schooling:								0.106	0.556	0.662	1.000	1.000
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap as a Proportion of Gap Driven by Levels vs. Effects:								0.161	0.839	1.000		

*** p<0.001, **p<0.01, * p<0.05, + p<0.10 (two-tailed *t*-tests for a statistically significant difference from 0). This model uses girls' coefficients as the reference when calculating each variable's contribution to the gap in schooling due to gender differences in mean levels and girls' means as the reference when calculating each variable's contribution due to gender differences in coefficients (i.e., effects).

Source: The 1983 to 1986 birth cohorts of the Children of the National Longitudinal Survey of Youth:1979 (NLSY-C; https://www.nlsinfo.org/content/cohorts/nlsy79-children) and matched National Longitudinal Survey of Youth:1979 (mother sample). The low-income white and military oversamples are excluded.

Note: The National Longitudinal Survey of Youth-Child Supplement (NLSY-C) consists of a nationally representative sample of children born to women age 14 to 21 in 1979; after excluding the poor white and military oversamples, the working sample in this study is restricted to the 1,857 children born 1983 to 1986, whose mothers were 18 to 29 years at birth. Children born 1983 to 1986 were born early enough to be age 26 to 29 as of the 2012 follow-up survey, but late enough to have early behavior problems information measured at age 4 to 5 beginning in 1986, at which point these items were introduced for children age 4 to 16. I used multiple imputation of 20 datasets to handle item-missingness. Model estimates use inverse-probability weighting to deal with stratified sample design (minority oversampling) and sample attrition by the 2012 follow-up wave (weights are described at: https://www.nlsinfo.org/weights/nlsy79). Once inverse-probability survey weights are applied, the working sample with complete attainment and behavior information drops from 1,857 to 1,661 children (881 girls, 780 boys).

Online Appendix Table A1.3: Detailed Results from Path Analysis of Indirect Pathways between Externalizing Problems at Ages 4 to 5 and Years of Schooling among NLSY-C Children Born 1983 to 1986, by Gender

				Males (N=780)				Females (N=881)			
				β	SE	T---Value	P---Value[1]	β	SE	T---Value	P---Value[1]
Total[1]				---0.110	0.045	---2.427	0.015	---0.066	0.048	---1.380	0.168
Total Indirect (Extern ---> X' ---> HGC)				---0.129	0.028	---4.567	0.000	---0.101	0.029	---3.542	0.000
X' (Vector of Specific Indirect Pathways: A, B, C, D)											
First-Order Indirect Path (A) between Externalizing Problems at Ages 4 to 5 and Years of Schooling (i.e., Highest Grade Completed)											
A	*B*	*C*	*D*								
READP68				---0.001	0.007	---0.144	0.885	0.001	0.004	0.275	0.784
MATHP68				---0.004	0.007	---0.594	0.552	---0.001	0.003	---0.421	0.674
EXPECT14				0.003	0.004	0.907	0.364	0.004	0.006	0.733	0.464
PEER1012				---0.004	0.004	---0.969	0.333	---0.003	0.003	---0.899	0.369
EXTERN12				---0.051	0.018	---2.831	0.005	---0.065	0.018	---3.621	0.000
READP12				---0.007	0.005	---1.291	0.197	---0.003	0.004	---0.817	0.414
MATHP12				---0.011	0.007	---1.460	0.144	---0.002	0.007	---0.292	0.770
EFFORT12				0.003	0.003	0.796	0.426	---0.004	0.004	---1.008	0.313
REPGR1417				---0.012	0.013	---0.884	0.377	0.001	0.012	0.068	0.946
Second-Order Indirect Paths (A & B) between Externalizing Problems at Ages 4 to 5 and Years of Schooling (i.e., Highest Grade Completed)											
A	*B*	*C*	*D*								
READP68	EXPECT14			0.000	0.001	---0.660	0.509	0.000	0.000	---0.111	0.911
MATHP68	EXPECT14			---0.002	0.001	---1.424	0.154	---0.001	0.001	---1.085	0.278
PEER1012	EXPECT14			0.000	0.000	---0.513	0.608	0.000	0.000	0.739	0.460
EXTERN12	EXPECT14			---0.005	0.003	---1.681	0.093	---0.006	0.003	---2.110	0.035
READP12	EXPECT14			0.000	0.000	---0.672	0.501	0.000	0.000	0.586	0.558
MATHP12	EXPECT14			---0.001	0.001	---1.058	0.290	0.000	0.001	---0.292	0.770
EFFORT12	EXPECT14			0.000	0.000	---0.834	0.404	0.000	0.000	0.642	0.521
READP68	PEER1012			0.000	0.000	0.814	0.416	---0.001	0.000	---1.129	0.259
MATHP68	PEER1012			0.000	0.000	---0.643	0.520	0.000	0.000	---0.801	0.423
READP68	COGN1012			0.000	0.001	---0.006	0.995	---0.001	0.001	---0.880	0.379
MATHP68	COGN1012			---0.002	0.001	---1.385	0.166	---0.001	0.001	---0.938	0.348
READP68	EMOT1012			0.000	0.000	0.137	0.891	0.000	0.000	---0.685	0.493
MATHP68	EMOT1012			0.000	0.000	0.132	0.895	0.000	0.000	---0.435	0.664
READP68	EXTERN12			---0.001	0.001	---1.241	0.215	0.000	0.001	---0.125	0.901
MATHP68	EXTERN12			0.000	0.001	---0.077	0.939	0.000	0.001	---0.813	0.416
PEER1012	EXTERN12			0.000	0.000	---0.870	0.384	0.000	0.000	---0.377	0.706
READP68	READP12			---0.007	0.004	---1.806	0.071	---0.002	0.002	---1.192	0.233
MATHP68	READP12			---0.002	0.001	---1.639	0.101	0.000	0.000	---0.881	0.378
PEER1012	READP12			---0.001	0.001	---1.224	0.221	0.000	0.000	---0.868	0.385
READP68	MATHP12			---0.004	0.002	---2.196	0.028	---0.002	0.002	---1.588	0.112
MATHP68	MATHP12			---0.006	0.003	---2.193	0.028	---0.003	0.002	---1.277	0.202
PEER1012	MATHP12			---0.002	0.001	---1.401	0.161	---0.001	0.001	---0.986	0.324
READP68	EFFORT12			0.000	0.000	---0.617	0.537	0.000	0.000	---1.046	0.296
MATHP68	EFFORT12			0.000	0.000	---0.313	0.754	0.000	0.000	---0.442	0.658
PEER1012	EFFORT12			0.000	0.000	---0.165	0.869	0.000	0.000	---0.091	0.927
READP68	REPGR1417			---0.002	0.002	---1.110	0.267	---0.002	0.001	---1.101	0.271
MATHP68	REPGR1417			0.001	0.002	0.564	0.573	0.000	0.001	0.153	0.878
PEER1012	REPGR1417			---0.002	0.002	---1.116	0.264	0.000	0.000	0.073	0.942
EXTERN12	REPGR1417			---0.007	0.005	---1.269	0.205	---0.008	0.004	---1.839	0.066
READP12	REPGR1417			0.000	0.001	0.437	0.662	0.001	0.001	0.848	0.397
MATHP12	REPGR1417			---0.001	0.001	---1.043	0.297	0.000	0.002	---0.291	0.771
EFFORT12	REPGR1417			0.001	0.001	0.999	0.318	0.000	0.001	0.291	0.771

Online Appendix Table A1.3: [Continued] Detailed Results from Path Analysis of Indirect Pathways between Externalizing Problems at Ages 4 to 5 and Years of Schooling among NLSY-C Children Born 1983 to 1986, by Gender

				Males (N=780)				Females (N=881)			
				β	SE	T---Value	P---Value[1]	β	SE	T---Value	P---Value[1]
Third-Order Indirect Paths (A---C) between Externalizing Problems at Ages 4 to 5 and Years of Schooling (i.e., Highest Grade Completed)											
A	*B*	*C*	*D*								
READP68	PEER1012	EXPECT14		0.000	0.000	0.534	0.593	0.000	0.000	0.862	0.389
MATHP68	PEER1012	EXPECT14		0.000	0.000	---0.465	0.642	0.000	0.000	0.666	0.505
READP68	EXTERN12	EXPECT14		0.000	0.000	---1.106	0.269	0.000	0.000	---0.125	0.901
MATHP68	EXTERN12	EXPECT14		0.000	0.000	---0.076	0.939	0.000	0.000	---0.808	0.419
PEER1012	EXTERN12	EXPECT14		0.000	0.000	---0.808	0.419	0.000	0.000	---0.387	0.699
READP68	READP12	EXPECT14		0.000	0.000	---0.702	0.483	0.000	0.000	0.649	0.517
MATHP68	READP12	EXPECT14		0.000	0.000	---0.679	0.497	0.000	0.000	0.572	0.568
PEER1012	READP12	EXPECT14		0.000	0.000	---0.646	0.519	0.000	0.000	0.568	0.570
READP68	MATHP12	EXPECT14		0.000	0.000	---1.267	0.205	0.000	0.000	---1.363	0.173
MATHP68	MATHP12	EXPECT14		0.000	0.000	---1.272	0.203	0.000	0.000	---1.189	0.234
PEER1012	MATHP12	EXPECT14		0.000	0.000	---1.063	0.288	0.000	0.000	---0.970	0.332
READP68	EFFORT12	EXPECT14		0.000	0.000	0.594	0.553	0.000	0.000	0.736	0.462
MATHP68	EFFORT12	EXPECT14		0.000	0.000	0.323	0.747	0.000	0.000	0.418	0.676
PEER1012	EFFORT12	EXPECT14		0.000	0.000	0.161	0.872	0.000	0.000	0.091	0.927
READP68	PEER1012	EXTERN12		0.000	0.000	0.769	0.442	0.000	0.000	---0.413	0.679
MATHP68	PEER1012	EXTERN12		0.000	0.000	---0.598	0.550	0.000	0.000	---0.365	0.715
READP68	COGN1012	EXTERN12		0.000	0.000	0.006	0.995	0.000	0.000	---0.430	0.667
MATHP68	COGN1012	EXTERN12		0.000	0.000	0.521	0.602	0.000	0.000	---0.440	0.660
READP68	EMOT1012	EXTERN12		0.000	0.000	0.404	0.686	0.000	0.000	---0.266	0.790
MATHP68	EMOT1012	EXTERN12		0.000	0.000	0.246	0.806	0.000	0.000	---0.230	0.818
READP68	PEER1012	READP12		0.000	0.000	0.989	0.322	0.000	0.000	---1.004	0.315
MATHP68	PEER1012	READP12		0.000	0.000	---0.700	0.484	0.000	0.000	---0.775	0.438
READP68	COGN1012	READP12		0.000	0.000	---0.006	0.995	0.000	0.000	---0.595	0.552
MATHP68	COGN1012	READP12		0.000	0.000	---1.126	0.260	0.000	0.000	---0.666	0.505
READP68	EMOT1012	READP12		0.000	0.000	0.361	0.718	0.000	0.000	0.607	0.544
MATHP68	EMOT1012	READP12		0.000	0.000	0.222	0.824	0.000	0.000	0.442	0.658
READP68	PEER1012	MATHP12		0.000	0.000	1.106	0.269	0.000	0.000	---1.294	0.196
MATHP68	PEER1012	MATHP12		0.000	0.000	---0.778	0.436	0.000	0.000	---0.857	0.392
READP68	COGN1012	MATHP12		0.000	0.000	---0.006	0.995	0.000	0.000	---0.707	0.480
MATHP68	COGN1012	MATHP12		0.000	0.000	---0.157	0.876	0.000	0.000	---0.823	0.410
READP68	EMOT1012	MATHP12		0.000	0.000	---0.342	0.732	0.000	0.000	0.623	0.533
MATHP68	EMOT1012	MATHP12		0.000	0.000	---0.216	0.829	0.000	0.000	0.445	0.657
READP68	PEER1012	EFFORT12		0.000	0.000	0.161	0.872	0.000	0.000	---0.091	0.928
MATHP68	PEER1012	EFFORT12		0.000	0.000	---0.161	0.872	0.000	0.000	---0.091	0.928
READP68	PEER1012	REPGR1417		0.000	0.000	1.006	0.314	0.000	0.000	0.074	0.941
MATHP68	PEER1012	REPGR1417		0.000	0.000	---0.676	0.499	0.000	0.000	0.073	0.942
READP68	COGN1012	REPGR1417		0.000	0.000	---0.006	0.995	0.000	0.000	---0.145	0.884
MATHP68	COGN1012	REPGR1417		---0.001	0.000	---1.377	0.168	0.000	0.000	---0.148	0.882
READP68	EMOT1012	REPGR1417		0.000	0.000	0.392	0.695	0.000	0.000	0.681	0.496
MATHP68	EMOT1012	REPGR1417		0.000	0.000	0.246	0.806	0.000	0.000	0.449	0.653
READP68	EXTERN12	REPGR1417		0.000	0.000	---0.901	0.368	0.000	0.000	---0.123	0.902
MATHP68	EXTERN12	REPGR1417		0.000	0.000	---0.077	0.939	0.000	0.000	---0.814	0.415
PEER1012	EXTERN12	REPGR1417		0.000	0.000	---0.735	0.462	0.000	0.000	---0.371	0.711
READP68	READP12	REPGR1417		0.000	0.001	0.455	0.649	0.001	0.000	1.121	0.262
MATHP68	READP12	REPGR1417		0.000	0.000	0.452	0.651	0.000	0.000	0.865	0.387
PEER1012	READP12	REPGR1417		0.000	0.000	0.436	0.663	0.000	0.000	0.815	0.415
READP68	MATHP12	REPGR1417		---0.001	0.000	---1.240	0.215	---0.001	0.000	---1.462	0.144
MATHP68	MATHP12	REPGR1417		---0.001	0.001	---1.216	0.224	---0.001	0.001	---1.233	0.218
PEER1012	MATHP12	REPGR1417		0.000	0.000	---1.060	0.289	0.000	0.000	---0.981	0.327
READP68	EFFORT12	REPGR1417		0.000	0.000	---0.579	0.563	0.000	0.000	0.312	0.755
MATHP68	EFFORT12	REPGR1417		0.000	0.000	---0.352	0.725	0.000	0.000	0.298	0.766
PEER1012	EFFORT12	REPGR1417		0.000	0.000	---0.158	0.875	0.000	0.000	0.093	0.926

Online Appendix Table A1.3: [Continued] Detailed Results from Path Analysis of Indirect Pathways between Externalizing Problems at Ages 4 to 5 and Years of Schooling among NLSY-C Children Born 1983 to 1986, by Gender

				Males (N=780)				Females (N=881)			
				β	SE	T-Value	P-Value[1]	β	SE	T-Value	P-Value[1]
Fourth-Order Indirect Paths (A-D) between Externalizing Problems at Ages 4 to 5 and Years of Schooling (i.e., Highest Grade Completed)											
A	*B*	*C*	*D*								
READP68	PEER1012	EXTERN12	EXPECT14	0.000	0.000	0.749	0.454	0.000	0.000	-0.422	0.673
MATHP68	PEER1012	EXTERN12	EXPECT14	0.000	0.000	-0.587	0.558	0.000	0.000	-0.370	0.711
READP68	COGN1012	EXTERN12	EXPECT14	0.000	0.000	0.006	0.995	0.000	0.000	-0.414	0.679
MATHP68	COGN1012	EXTERN12	EXPECT14	0.000	0.000	0.492	0.623	0.000	0.000	-0.423	0.672
READP68	EMOT1012	EXTERN12	EXPECT14	0.000	0.000	0.389	0.697	0.000	0.000	-0.266	0.790
MATHP68	EMOT1012	EXTERN12	EXPECT14	0.000	0.000	0.241	0.809	0.000	0.000	-0.230	0.818
READP68	PEER1012	READP12	EXPECT14	0.000	0.000	0.624	0.533	0.000	0.000	0.602	0.547
MATHP68	PEER1012	READP12	EXPECT14	0.000	0.000	-0.533	0.594	0.000	0.000	0.570	0.569
READP68	COGN1012	READP12	EXPECT14	0.000	0.000	-0.006	0.995	0.000	0.000	0.507	0.612
MATHP68	COGN1012	READP12	EXPECT14	0.000	0.000	-0.608	0.543	0.000	0.000	0.564	0.573
READP68	EMOT1012	READP12	EXPECT14	0.000	0.000	0.306	0.760	0.000	0.000	-0.465	0.642
MATHP68	EMOT1012	READP12	EXPECT14	0.000	0.000	0.204	0.838	0.000	0.000	-0.378	0.706
READP68	PEER1012	MATHP12	EXPECT14	0.000	0.000	0.964	0.335	0.000	0.000	-1.196	0.232
MATHP68	PEER1012	MATHP12	EXPECT14	0.000	0.000	-0.703	0.482	0.000	0.000	-0.849	0.396
READP68	COGN1012	MATHP12	EXPECT14	0.000	0.000	-0.006	0.995	0.000	0.000	-0.691	0.490
MATHP68	COGN1012	MATHP12	EXPECT14	0.000	0.000	-0.155	0.876	0.000	0.000	-0.802	0.422
READP68	EMOT1012	MATHP12	EXPECT14	0.000	0.000	-0.336	0.737	0.000	0.000	0.611	0.541
MATHP68	EMOT1012	MATHP12	EXPECT14	0.000	0.000	-0.215	0.830	0.000	0.000	0.445	0.657
READP68	PEER1012	EFFORT12	EXPECT14	0.000	0.000	-0.158	0.875	0.000	0.000	0.091	0.927
MATHP68	PEER1012	EFFORT12	EXPECT14	0.000	0.000	0.157	0.875	0.000	0.000	0.091	0.927
READP68	PEER1012	EXTERN12	REPGR1417	0.000	0.000	0.667	0.505	0.000	0.000	-0.406	0.685
MATHP68	PEER1012	EXTERN12	REPGR1417	0.000	0.000	-0.543	0.587	0.000	0.000	-0.360	0.719
READP68	COGN1012	EXTERN12	REPGR1417	0.000	0.000	0.006	0.995	0.000	0.000	-0.417	0.677
MATHP68	COGN1012	EXTERN12	REPGR1417	0.000	0.000	0.486	0.627	0.000	0.000	-0.425	0.671
READP68	EMOT1012	EXTERN12	REPGR1417	0.000	0.000	0.385	0.700	0.000	0.000	-0.266	0.790
MATHP68	EMOT1012	EXTERN12	REPGR1417	0.000	0.000	0.239	0.811	0.000	0.000	-0.229	0.819
READP68	PEER1012	READP12	REPGR1417	0.000	0.000	-0.432	0.666	0.000	0.000	0.939	0.348
MATHP68	PEER1012	READP12	REPGR1417	0.000	0.000	0.395	0.693	0.000	0.000	0.758	0.448
READP68	COGN1012	READP12	REPGR1417	0.000	0.000	0.006	0.995	0.000	0.000	0.572	0.568
MATHP68	COGN1012	READP12	REPGR1417	0.000	0.000	0.431	0.667	0.000	0.000	0.661	0.509
READP68	EMOT1012	READP12	REPGR1417	0.000	0.000	-0.300	0.764	0.000	0.000	-0.609	0.543
MATHP68	EMOT1012	READP12	REPGR1417	0.000	0.000	-0.198	0.843	0.000	0.000	-0.435	0.663
READP68	PEER1012	MATHP12	REPGR1417	0.000	0.000	0.898	0.369	0.000	0.000	-1.221	0.222
MATHP68	PEER1012	MATHP12	REPGR1417	0.000	0.000	-0.658	0.510	0.000	0.000	-0.825	0.409
READP68	COGN1012	MATHP12	REPGR1417	0.000	0.000	-0.006	0.995	0.000	0.000	-0.698	0.485
MATHP68	COGN1012	MATHP12	REPGR1417	0.000	0.000	-0.157	0.875	0.000	0.000	-0.810	0.418
READP68	EMOT1012	MATHP12	REPGR1417	0.000	0.000	-0.334	0.738	0.000	0.000	0.597	0.550
MATHP68	EMOT1012	MATHP12	REPGR1417	0.000	0.000	-0.215	0.830	0.000	0.000	0.451	0.652
READP68	PEER1012	EFFORT12	REPGR1417	0.000	0.000	0.155	0.877	0.000	0.000	0.093	0.926
MATHP68	PEER1012	EFFORT12	REPGR1417	0.000	0.000	-0.154	0.877	0.000	0.000	0.093	0.926
Direct Path between Externalizing Problems at Ages 4-5 and Years of Schooling				0.019	0.042	0.448	0.654	0.035	0.041	0.866	0.386
Chi-square (df)							147.99 (30)***				
RMSEA							0.05				
CFI							0.974				

*** p<0.001, **p<0.01, * p<0.05, + p<0.10. Estimates display the standardized total, total indirect, specific indirect, and direct effects (i.e., associations) between externalizing problems at ages 4 to 5 and years of schooling at ages 26 to 29.

[1] Note that the total effect (T) is less than the total indirect effect (I) because the net direct effect (D) between externalizing problems at age 4 to 5 and years of schooling becomes positive (although non-statistically significantly from 0) for both genders after controlling for all proximate mediating factors. For example, since T=D+I, $D_{females}$= 0.035, $T_{females}$= -0.101 + 0.035 = -0.066.

Source: The 1983 to 1986 birth cohorts of the Children of the National Longitudinal Survey of Youth:1979 (NLSY-C; https://www.nlsinfo.org/content/cohorts/nlsy79-children) and matched National Longitudinal Survey of Youth:1979 (mother sample). The low-income white and military oversamples are excluded.

Note: The National Longitudinal Survey of Youth-Child Supplement (NLSY-C) consists of a nationally representative sample of children born to women age 14 to 21 in 1979; after excluding the poor white and military oversamples, the working sample in this study is restricted to the 1,857 children born 1983 to 1986, whose mothers were 18 to 29 years at birth. Children born 1983 to 1986 were born early enough to be age 26 to 29 as of the 2012 follow-up survey, but late enough to have early behavior problems information measured at age 4 to 5 beginning in 1986, at which point these items were introduced for children age 4 to 16. I used multiple imputation of 20 datasets to handle item-missingness. Model estimates use inverse-probability weighting to deal with stratified sample design (minority oversampling) and sample attrition by the 2012 follow-up wave (weights are described at: https://www.nlsinfo.org/weights/nlsy79). Once inverse-probability survey weights are applied, the working sample with complete attainment and behavior information drops from 1,857 to 1,661 children (881 girls, 780 boys).

Online Appendix Table A2.1: Detailed Results for Two-Way Decomposition of Gender Differences in Levels and Effects of Externalizing Problems and Family Factors on Years of Schooling at Age 19 to 22 for NLSY-C Children Born 1983 to 1993 (Reference: Boys)

	Means		Means Diff	OLS Regression Coefficients			Contrib-ution of Diff. in Levels	Contrib-ution of Diff. in Coeff-icients	Total Contribution of Levels & Coefficients	Prop. of Positive Effects on Gap	Prop. of Negative Effects on Gap
	(1) F	(2) M	(3) F-M	(4) F	Sig	(5) M Sig	(6)	(7)	(8)	(9)	(10)
Birth Order	1.983	1.992	-0.009	-0.003		-0.252 **	0.002	0.496	0.498	0.230	0.000
Mother Years of Schooling at Birth	12.852	12.722	0.130	0.272 ***		0.261 ***	0.034	0.140	0.174	0.080	0.000
Externalizing Problems, Ages 4-5	**1.800**	**2.336**	**-0.536**	**-0.101 *** **		**-0.112 *** **	**0.060**	**0.026**	**0.086**	**0.039**	0.000
African-American	0.148	0.136	0.012	0.260		-0.120	-0.001	0.052	0.050	0.023	0.000
Home Environment, Ages 3-5	-0.171	-0.142	-0.029	-0.467 **		-0.164	0.005	0.043	0.048	0.022	0.000
Child Std. Score on PPVT, Ages 3-4	0.393	0.273	0.120	0.323 ***		0.343 ***	0.041	-0.005	0.036	0.016	0.000
Low Birth Weight	0.075	0.056	0.019	0.065		-0.293	-0.006	0.020	0.014	0.007	0.000
Hispanic	0.069	0.075	-0.006	0.099		-0.002	0.000	0.008	0.008	0.003	0.000
Received Care Outside Home, Ages 4-5	0.536	0.513	0.023	-0.155		-0.088	-0.002	-0.034	-0.036	0.000	0.022
Mother Age at Birth	26.787	26.807	-0.020	-0.072 **		-0.021	0.000	-1.367	-1.367	0.000	0.839
Number of Children in Family, Ages 4-5	2.620	2.628	-0.008	-0.091		-0.017	0.000	-0.194	-0.194	0.000	0.119
Father Absent at Child's Birth	0.172	0.153	0.019	-0.470 **		-0.298 *	-0.006	-0.026	-0.032	0.000	0.020
Household Income in $2011, Child Ages 4-5	10.067	9.203	0.864	0.002		-0.008	-0.007	0.092	0.085	0.039	0.000
Constant	1.000	1.000	0.000	12.285 ***		1.113 ***	0.000	1.172	1.172	0.540	0.000
Observations (N)	2104	2074		2104		2074					
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap in Years of Schooling:							0.121	0.420	0.541	1.000	1.000
Overall Contribution of Early Externalizing Problems, Early Childhood Factors, and Controls to the Gender Gap as a Proportion of Gap Driven by Levels vs. Effects:							0.224	0.776	1.000		

*** p<0.001, **p<0.01, * p<0.05, + p<0.10 (two-tailed *t*-tests for a statistically significant difference from 0). This model uses boys' coefficients as the reference when calculating each variable's contribution to the gap in schooling due to gender differences in mean levels and boys' means as the reference when calculating each variable's contribution due to gender differences in coefficients (i.e., effects).

Source: The 1983 to 1993 birth cohorts of the Children of the National Longitudinal Survey of Youth:1979 (NLSY-C; https://www.nlsinfo.org/content/cohorts/nlsy79-children) and matched National Longitudinal Survey of Youth:1979 (mother sample). The low-income white and military oversamples are excluded.

Note: The National Longitudinal Survey of Youth-Child Supplement (NLSY-C) consists of a nationally representative sample of children born to women age 14 to 21 in 1979; after excluding the poor white and military oversamples, the working sample in this study is restricted to the 1,857 children born 1983 to 1986, whose mothers were 18 to 29 years at birth. Children born 1983 to 1986 were born early enough to be age 26 to 29 as of the 2012 follow-up survey, but late enoughto have early behavior problems information measured at age 4 to 5 beginning in 1986, at which point these items were introduced for children age 4 to 16. I used multiple imputation of 20 datasets to handle item-missingness. Model estimates use inverse-probability weighting to deal with stratified sample design (minority oversampling) and sample attrition by the 2012 follow-up wave (weights are described at: https://www.nlsinfo.org/weights/nlsy79). Once inverse-probability survey weights are applied, the working sample with complete attainment and behavior information drops from 1,857 to 1,661 children (881 girls, 780 boys).