

PRESCHOOL TELEVISION VIEWING AND ADOLESCENT TEST SCORES: HISTORICAL EVIDENCE FROM THE COLEMAN STUDY

MATTHEW GENTZKOW AND JESSE M. SHAPIRO

We use heterogeneity in the timing of television's introduction to different local markets to identify the effect of preschool television exposure on standardized test scores during adolescence. Our preferred point estimate indicates that an additional year of preschool television exposure *raises* average adolescent test scores by about 0.02 standard deviations. We are able to reject negative effects larger than about 0.03 standard deviations per year of television exposure. For reading and general knowledge scores, the positive effects we find are marginally statistically significant, and these effects are largest for children from households where English is not the primary language, for children whose mothers have less than a high school education, and for nonwhite children.

I. INTRODUCTION

Television has attracted young viewers since broadcasting began in the 1940s. Concerns about its effects on the cognitive development of young children emerged almost immediately and have been fueled by academic research showing a negative association between early-childhood television viewing and later academic achievement.¹ These findings have contributed to a belief among the vast majority of pediatricians that television has “negative effects on brain development” of children below age five (Gentile et al. 2004). They have also provided partial motivation for recent recommendations that preschool children's television viewing time be severely restricted (American Academy of Pediatrics 2001). According to a widely cited report on media use by young

* We are grateful to Dominic Brewer, John Collins, Ronald Ehrenberg, Eric Hanushek, and Mary Morris (at ICPSR) for assistance with Coleman study data, and to Christopher Berry for supplying data on school quality. Lisa Furchtgott, Jennifer Paniza, and Mike Sinkinson provided outstanding research assistance. We thank Marianne Bertrand, Stefano DellaVigna, Ed Glaeser, Austan Goolsbee, Jim Heckman, Caroline Hoxby, Larry Katz, Steve Levitt, Ethan Lieber, Jens Ludwig, Kevin M. Murphy, Emily Oster, Matthew Rabin, Andrei Shleifer, Chad Syverson, Bob Topel, workshop participants at the University of Chicago, Harvard University, UC Berkeley, the NBER, the University of Notre Dame, and the APPAM, and four anonymous referees for helpful comments. e-mail: gentzkow@ChicagoGSB.edu, jmshapir@uchicago.edu.

1. Recent studies showing negative correlations between early childhood viewing and later performance include Christakis et al. (2004), Hancox, Milne, and Poulton (2005) and Zimmerman and Christakis (2005). An older literature finds more mixed results, but reviewers conclude that the overall thrust of the evidence points toward negative effects of television (Strasburger 1986; Beentjes and Van der Voort 1988; Van Evra 1998).

children, “Many experts have argued that it is especially critical to understand media use by the youngest children . . . because social and intellectual development are more malleable in these early years” (Rideout, Vandewater, and Wartella 2003). This view is supported by randomized studies demonstrating large long-run effects of preschool interventions on children’s cognitive skills (Campbell and Ramey 1995; Currie 2001; Schweinhart et al. 2005).

Evidence of negative cognitive effects has made the growth of television a popular explanation for trends such as the decline in average verbal SAT scores during the 1970s (Wirtz et al. 1977; Winn 2002) and the secular decline in verbal ability across cohorts (Glenn 1994). Given the important role that cognitive skills play in individual (Griliches and Mason 1972) and aggregate (Bishop 1989) labor market performance, understanding the cognitive effects of television viewing may have significant implications for public policy and household behavior.

In this paper, we identify the effect of preschool exposure to television on adolescent cognitive skills by exploiting variation in the timing of television’s introduction to U.S. cities.² Most cities first received television between the early 1940s and the mid-1950s. The exact timing was affected by a number of exogenous events, most notably a four-year freeze on licensing prompted by problems with the allocation of broadcast spectrum across cities. Once it was introduced, television was adopted rapidly by families with children. Survey evidence suggests that young children who had television in their homes during this period watched as much as three and a half hours per day, and contemporary time-use studies show reductions in a wide range of alternative activities, including sleep, homework, and outdoor play. Evidence on television ownership suggests that the diffusion of television was broad-based, reaching families in many different socioeconomic strata. Together, these facts create a promising laboratory in which to study the effects of television on children.

To conduct our analysis, we use data from a 1965 survey of American schools and school children, commonly referred to as the Coleman Study. The data include standardized test scores of over 300,000 students who were in grades 6, 9, and 12 in 1965. These students were born between 1948 and 1954, just as television was

2. We build on the identification strategy developed by Gentzkow (2006). For earlier papers exploiting the timing of television’s introduction, see Parker (1963) and Hennigan et al. (1982).

expanding throughout the United States. Because television entered different U.S. markets at different times, students were exposed to varying amounts of television as preschoolers. Students in our sample range from those who had television in their local areas throughout their lives (for example, sixth graders whose areas got television between 1945 and 1954) to those whose areas only began receiving broadcasts after they reached age 6 (twelfth graders whose areas got television in 1954). Because the Coleman sample includes students of different ages within the same television market, we can identify the effects of television by comparing test scores across cohorts within a given area. This differences-in-differences approach allows us to estimate the effect of preschool television exposure on adolescent test scores, while holding constant fixed characteristics of a locale that affect test scores and might also be correlated with the timing of television introduction.

We find strong evidence against the view that childhood television viewing harms the cognitive or educational development of preschoolers. Our preferred point estimate indicates that an additional year of preschool television exposure *raises* average adolescent test scores by about 0.02 standard deviations. We are able to reject negative effects larger than about 0.03 standard deviations per year of television exposure.³ For reading and general knowledge scores—domains where intuition and existing evidence suggest that learning from television could be important—the positive effects we find are marginally statistically significant. In addition, we present evidence on the extent to which childhood viewing affects later noncognitive outcomes such as time spent on homework and desired school completion, again finding no consistent evidence of negative effects.

A number of specification checks support the identifying assumption that the timing of television's entry is uncorrelated with direct determinants of test scores. Most importantly, we find that the within-area cross-cohort variation in television exposure that identifies our models does not correlate with demographic variables that affect test scores. We also find that the timing of television introduction is uncorrelated with trends in area school quality, teacher characteristics, and demographics. Thus, although by definition we cannot test that our key exposure

3. For comparison, the early childhood interventions we discuss in Section V.B had long-term effects on achievement of approximately 0.07 to 0.25 standard deviations per year of intervention (Campbell and Ramey 1995; Schweinhart et al. 2005).

measures are orthogonal to unobservable variation in student ability, we show that these measures are unrelated to many observable correlates of ability.

Our final set of results addresses heterogeneity in the effects of television on test scores. The effects on verbal, reading, and general knowledge scores are most positive for children from households where English is not the primary language, for children whose mothers have less than a high school education, and for nonwhite children. When we combine student observables into a single index of parental investment—the time parents spent reading to their children in early childhood—we find that the effect of television is significantly more positive the lower is parental investment. Consistent with a rational-choice model, families in which television has relatively positive effects on learning also allocate more time to viewing.⁴

These findings point toward an important economic intuition that is often overlooked in the popular debate about television: the cognitive effects of television exposure depend critically on the educational value of the alternative activities that it crowds out. Like other early-childhood interventions (Currie 2001), television seems to be most beneficial for children who are relatively disadvantaged. For children with highly educated parents and rich home environments, the cognitive effects of television appear to be smaller and may even be negative. These results cast doubt on policies such as the American Academy of Pediatrics recommendations cited above that advocate a uniform standard of viewing for all young children. They also suggest that endogenous choice of viewing hours is likely to tilt the aggregate impact of television in a positive direction.

We wish to stress three important caveats. First, our identification strategy only allows us to speak to the effects of early childhood exposure. The effects of viewing by school-age children are also clearly important for policy, and our results do not directly inform that debate. Second, we can only identify long-run effects. Although concern about the cognitive effects of early-childhood viewing has been largely motivated by the possibility of harm to long-run development, there are other potential effects of television—on violence or obesity, for example—for which contemporaneous effects may be more relevant. Finally, we measure

4. In this respect, our paper relates to the literature on empirical selection into behaviors (Roy 1951; Heckman and Sedlacek 1985; Heckman 1996).

only the impact of 1950s-era television. Changes in content such as the increased availability of both educational and violent programming, as well as changes in the nontelelevision alternatives available to young children, could mean that the effects of television viewing today are different from those we estimate.

Our study contributes to a large literature on the cognitive effects of television, most of which identifies the effect of television using cross-sectional variation in children's viewing intensity.⁵ It also contributes to a growing economic literature on the effects of media on children (Dahl and DellaVigna 2006), and on the effects of mass media more generally (see, for example, Djankov et al. [2003]; Gentzkow and Shapiro [2004, 2006]; Stromberg [2004]; Gentzkow [2006]; Olken [2006]; and DellaVigna and Kaplan [2007]).

The remainder of the paper is organized as follows. Section II discusses the history of the introduction and diffusion of television. Section III presents our data. Section IV discusses our identification strategy and reduced-form findings. Section V presents estimates of the effect of preschool television exposure on cognitive development and student achievement, and Section VI presents an analysis of heterogeneity across students. Section VII concludes.

II. THE INTRODUCTION AND DIFFUSION OF TELEVISION

The Federal Communications Commission (FCC) first licensed television for full-scale commercial broadcasting on July 1, 1941.⁶ Two unexpected events intervened to delay television's expansion. The first was World War II: less than a year after the FCC authorization, the government issued a ban on new television station construction to preserve materials for the war effort. Although some existing stations continued to broadcast, the total

5. See Strasburger (1986), Beentjes and Van der Voort (1988), and Van Evra (1998) for reviews, and Zavodny (2006) for panel evidence. Two previous studies have used natural-experiment designs. Schramm, Lyle, and Parker (1961) compare two small towns in western Canada, one of which had access to television and the other of which did not. Harrison and Williams (1986) analyze data from three small Canadian towns, both before and after one of the towns received television. Neither study finds evidence of strong cognitive effects, although both find weak evidence that access to television improves young children's vocabulary. Our paper employs a similar source of variation to these studies but on a much larger scale. See Cook et al. (1975) and Diaz-Guerrero et al. (1976) for randomized studies of the effects of specific programming content.

6. This section draws primarily on Barnouw (1990) and Sterling and Kittross (2001). For details on the regulatory process, see also Slotten (2000).

number of sets in use during the war was less than 20,000. After the war, television expanded rapidly. Over 100 new licenses were issued between 1946 and 1948, so that by 1950 half of the country's population was reached by television signals. This growth was again halted, however, by an FCC-imposed freeze on new television licenses in September 1948. The FCC had determined that spectrum allocations did not leave sufficient space between adjacent markets, causing excessive interference. The process of redesigning the spectrum allocation took four years, and it was not until April 1952 that the freeze was lifted and new licenses began to be issued.

The diffusion of television ownership was rapid and demographically broad. Contemporaneous polling data show that television penetration rose from 8% to 82% from 1949 to 1955 among those with high school degrees, and from 4% to 66% among those without. Other demographic groups tend to show a similar pattern: television diffusion was rapid among both whites and non-whites, and among both elderly and nonelderly Americans.⁷ In households with television, viewership had already surpassed four and a half hours per day by 1950 (Television Bureau of Advertising 2003).

Children were among the most enthusiastic early viewers of television. Programs targeted specifically at children were introduced early, with *Howdy Doody* making its debut in 1947 and a number of popular series such as *Kukla, Fran, and Ollie*; *Jamboree Room*; and *Children's Matinee* on the air by 1948 (Television Magazine 1948). Children's programs accounted for more time on network television than any other category in 1950 (Roslow 1952), and by 1951 advertisers were spending \$400,000 per week to reach the children's market (Television Magazine 1951). Furthermore, children were frequent viewers of programming primarily targeted at adults—to take one example, *I Love Lucy* was ranked the most favored program among elementary-school students in 1952, 1953, and 1954 surveys (Television Magazine 1955).⁸

There were no large-scale studies of children's viewing hours in the 1950s, but a series of small surveys make clear that intense viewing was common from television's earliest years. Median daily viewership in samples of elementary-school children ranged from

7. Based on Gallup polls of American households (Roper Center for Public Opinion Research 1949–1955).

8. A 1960 study found that 40% of children's viewing was devoted to adult programs (Schramm, Lyle, and Parker 1961).

2.0 hours to 3.7 hours per day, with the earliest studies showing 3.1 hours per day in 1948 (ages 6–12), 3.7 hours per day in 1950–51 (grades 6–7), 2.7 hours per day in 1951 (elementary ages), 3.3 hours in 1953 (elementary ages), 3.7 hours in 1954 (grades 4–8), and 3.4 hours in 1955 (elementary ages).⁹ The only evidence we are aware of on *preschool* viewing—a small survey of families in San Francisco in 1958—found that weekday viewing averaged 0.7 hours per day for three-year-olds, 1.6 hours per day for four-year-olds, and 2.3 hours per day for five-year-olds, with weekend viewing on average half an hour to an hour higher (Schramm, Lyle, and Parker 1961).

Two studies from the period document the dramatic changes that television brought to children's allocation of time. First, Maccoby (1951) surveyed 622 children in Boston in 1950 and 1951 and matched children with and without television by age, sex, and socioeconomic status. The study found that radio listening, movie watching, and reading were substantially lower in the television group, but also that total media time was greater by approximately an hour and a half per day.¹⁰ The television group went to bed almost half an hour later and spent less time on homework and active play. The second study, conducted in 1959, surveyed children in two similar towns in Western Canada of which only one had television available (Schramm, Lyle, and Parker 1961). First-grade children in the town with television watched for an average of an hour and 40 minutes per day. They spent 35 fewer minutes listening to radio, 33 fewer minutes at play, 13 fewer minutes sleeping, and 20 fewer minutes reading and watching movies. Sixth-grade children showed similar shifts in time allocation and also spent 15 fewer minutes on homework.

III. MEASURING TEST SCORES AND TELEVISION EXPOSURE

III.A. *Test Scores in Grades 6–12*

Our data on test scores will come from the the Coleman Study, formally titled *Equality of Educational Opportunity* (Coleman

9. See Schramm, Lyle, and Parker (1961) for a review of this evidence.

10. The conclusion that the time devoted to television did not simply replace radio is supported by a number of studies suggesting that even in the 1930s radio listening averaged little more than an hour per day among elementary-age children (Fox Meadow School PTA 1933; Eisenberg 1936).

1966).¹¹ The study includes data on 567,148 students who were in grades 1, 3, 6, 9, or 12 in 1965. Sampling was conducted through the construction of primary sampling units (PSUs) consisting of either counties or metropolitan areas. Because racial differences were a primary focus of the study, PSUs, school districts, and schools were selected so that nonwhite students were oversampled relative to the U.S. population.

Within sample schools, all students were included in the study. Each student completed a survey and an exam, both of which were administered in the fall of 1965. We will focus our analysis on sixth, ninth, and twelfth graders because these students' birth cohorts (1948–1954) span most of the period during which television was introduced, and because exam style and format were fairly similar across these different grades. Exams for sixth, ninth, and twelfth graders contained sections on mathematics, spatial reasoning, verbal ability (vocabulary), and reading; ninth and twelfth graders completed an additional section on general knowledge. In addition to information on test scores, we extracted data on demographic characteristics from the student surveys. We tried to include all characteristics that were available and reasonably comparable across all three grades.

To select sample schools, the surveyors first chose schools with twelfth grades. Then, for each school containing a twelfth grade, they identified the middle and elementary schools that “fed” their students into the secondary school. If a lower-grade school fed more than 90% of its students into the selected twelfth-grade school, then it was sampled with certainty; other lower-grade schools were sampled in proportion to the share of their students who were fed into the twelfth-grade school. The Coleman data contain a school identifier variable unique to each sampled school containing a twelfth grade. For students in lower-grade schools, this identifier refers to the sampled twelfth-grade school into which the students were fed. We will employ this identifier to estimate specifications with “school” fixed effects, though we note that in the case of sixth graders attending schools without a twelfth grade, it may be better thought of as a school district fixed effect.

For schools located in metropolitan areas, our data match the school identifier to the Standard Metropolitan Statistical Area (SMSA) in which the school was located in 1965. For all other

11. For examples of other studies by economists using data from this study, see Hanushek and Kain (1972) and Ehrenberg and Brewer (1995).

schools, the data identify the county in which the school was located.¹² To estimate the extent to which students in the Coleman sample were exposed to television during early childhood, we will assume that the television market where a student currently attends school is the same as the one where he or she grew up. In Section V.C below, we use direct data on students' mobility since early childhood to show that this assignment is likely to be accurate for the vast majority of students, and that our conclusions are, if anything, strengthened by excluding those who are most likely to have grown up in a different market.

III.B. Television Availability in Local Markets

Our estimation strategy relies on information about the availability of television in U.S. cities beginning in 1946. We use data from Gentzkow (2006) on the year in which the first television station appeared in a given market.¹³ These data were compiled from annual editions of the *Television Factbook*. We define television markets using the Designated Market Area (DMA) concept designed by Nielsen Media Research (NMR). NMR assigns every county in the U.S. to a television market, such that all counties in a given market have a majority of their measured viewing hours on stations broadcasting from that market.¹⁴ We define the year television was introduced to a given county or SMSA to be the first year in which a station in its DMA broadcast for at least four months.

For the purposes of estimation, we will divide DMAs into three groups according to the year in which they began receiving television broadcasts: *early adopters* (broadcasts begin in 1948 or earlier), *middle adopters* (1949 to 1951), and *late adopters* (1952 or later). These categories, which correspond to the periods before, during, and after the FCC freeze, capture most of the relevant variation in the data.

12. Approximately 62% of the students in our sample live in metropolitan areas.

13. In most cases, we use the date that a station began commercial broadcasts, as regulated by the FCC. The exceptions are two stations—KTLA in Los Angeles and WTTG in Washington, DC—that began large-scale experimental broadcasts and subsequently converted to become commercial stations. In these cases, we use the stations' experimental start dates.

14. These definitions are based on viewership as of 2003, rather than in the historical period we are analyzing. However, because the broadcasting strength of stations is regulated by the FCC to avoid interference with neighboring markets, the area reached by particular stations has remained relatively constant over time. This has been verified by spot-checking the DMA definitions against coverage maps from the 1960s.

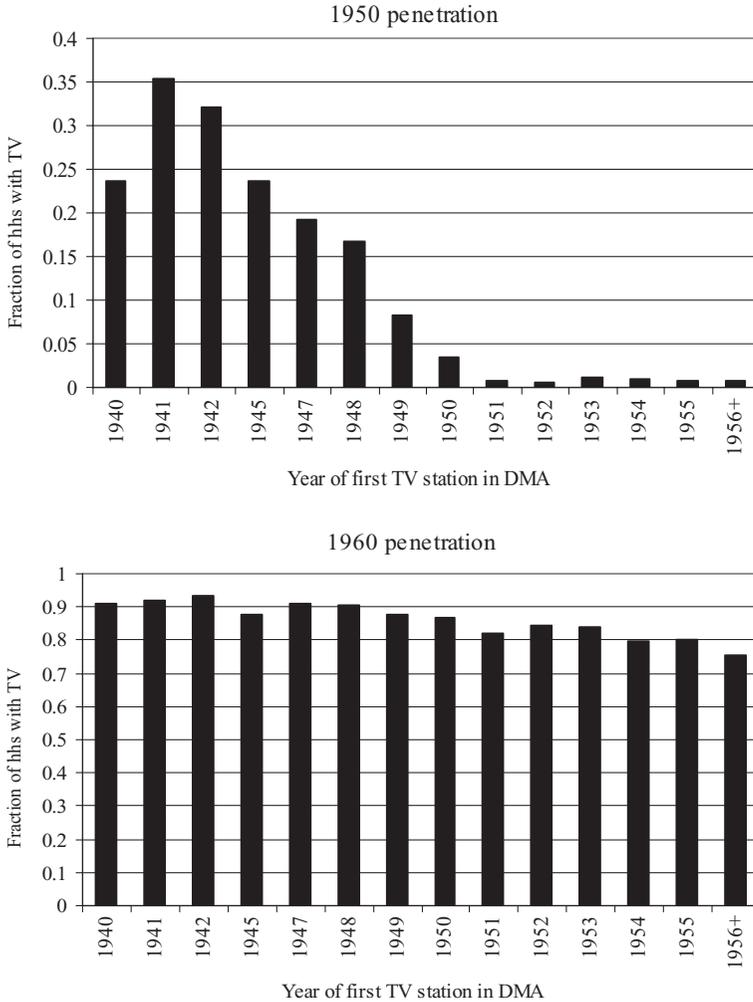


FIGURE I

Television Penetration in 1950 and 1960 by Year of Television Introduction

Source. *Television Factbook*, various years; 1950 and 1960 U.S. Censuses.

Notes. The height of each bar is the fraction of households with television sets as recorded in the 1950 or 1960 census, averaged over all DMAs that received television in the given year. Years in which no county received its first station are omitted from the figure.

To illustrate the impact of broadcast availability on television ownership, we compare our availability measure with data on television ownership from the 1950 and 1960 U.S. Censuses. Figure I shows the share of households owning televisions as a function of the year in which television broadcasts began in the

DMA. The first graph, which shows penetration in 1950, reveals a clear distinction between counties that had a station in their DMA and those that did not. The average penetration in DMAs whose first station began broadcasting before 1950 ranges from 8% in the 1949 group to over 35% in the 1941 group, whereas the average for groups getting television after 1950 never exceeds 1%. The second graph shows that, by 1960, differences in penetration across these DMAs had largely disappeared. Differences in the timing of introduction of television into different areas thus had a large initial impact, but by 1960 most late-adopting DMAs had caught up to those that began receiving broadcasts early. These patterns will be crucial to allowing us to identify the effect of television using differences across birth cohorts within a DMA.

An examination of historical records suggests two potential sources of endogeneity in the timing of television's introduction to a market. First, the FCC sought to maximize the number of people who could receive a commercial television signal. Conditional on the quality of existing coverage in a market, the FCC therefore handled applications to begin broadcasting in order of the market's total population (Television Digest 1953). Second, because a station's profitability was determined largely by advertising revenue, which in turn depended on the spending power of the market's population, commercial interest in operating stations in a given market was highly related to the market's total retail sales or income.

The data confirm the expected role of population and income. Early- and middle-adopting DMAs had, on average, five times larger populations and 24% larger per capita incomes than late-adopting DMAs. After controlling for log population and income, however, differences between early and late adopters appear much more idiosyncratic. Indeed, in regressions controlling for log population and income, *F*-tests show no statistically significant relationship between television adoption category and percent high school educated, median age, or percent nonwhite at the DMA level. (See the online appendix to this paper for details.) All of the models we estimate below will control for DMA-level log population and income, so the parameters will be driven solely by variation in the availability of television orthogonal to these two factors.¹⁵ In Section V.C below, we show formally that the remaining

15. The patterns in Figure I are substantively unchanged if we measure penetration using the residual from a regression of penetration on $\log(\text{income})$ and $\log(\text{population})$ at the DMA level.

variation in television adoption timing is not systematically related to student-level observables.

III.C. Childhood Exposure to Television

The data described above allow us to calculate the number of years of a given student's early childhood in which television signals were *available*. In order to make the magnitudes we measure in the analysis below more easily interpretable, we will also use data on the rate at which television ownership actually diffused among households in each county. We will use the term television *exposure* to refer to the expected number of years a child's household owned a television during the child's preschool years.

To construct our measure of exposure, we collect annual data on television penetration for U.S. counties. We combine the 1950 and 1960 U.S. Census data mentioned above with data from industry sources covering 1953 to 1959.¹⁶ For years with missing data, we used a linear interpolation (or extrapolation) from the surrounding years, with a transformation that restricts penetration shares to fall between zero and one.¹⁷

We use these penetration data to compute the expected years of television exposure during ages two through six in each county for each Coleman Study cohort.¹⁸ For example, consider students in some county who were in grade 12 in 1965, the year of the Coleman Study. Most students in this group were born in 1948. Suppose that television penetration in the county was 10% in 1950

16. We use data from Television magazine's *Market Book* for the years 1954–1959 and separate county-level data from the *Television Factbook* for 1953. These sources combine information from the Advertising Research Foundation, A.C. Nielsen, NBC, and CBS, as well as television shipments data, to construct annual estimates of penetration by county. The correlation between *Television's* county-level penetration estimates for 1959 and the U.S. Census counts for 1960 is a highly statistically significant 0.64 ($p < .0001$). Given that *Television* did not yet have access to the Census reports when producing these figures, this correlation suggests reasonably high reliability.

17. In particular, we computed the transformation $\log(\text{penetration}/(1-\text{penetration}))$ and imputed missing values using a linear interpolation (or extrapolation) of this transformed measure. We then used the inverse function to retransform the imputed values to a 0–1 scale. This approach amounts to assuming that television diffusion follows an S-shaped logistic process in years with missing data (Griliches, 1957).

18. In results not reported, we have also experimented with separate measures of television exposure during ages 0 through 3 and ages 4 through 6. In the specifications where we find marginally significant evidence of positive effects of television—reading and general knowledge—the effects tend to be larger for exposure at ages 4 through 6 than for exposure at ages 0 through 3 (although these differences are not statistically significant and this pattern does not hold for all tests). This finding is consistent with historical evidence that older children watched more hours of television.

(age two), 11% in 1951 (age three), 12% in 1952 (age four), 13% in 1953 (age five), and 14% in 1954 (age six). Then we calculate the total years of preschool television exposure for twelfth graders in this county as $(0.10 + 0.11 + 0.12 + 0.13 + 0.14) = 0.6$.¹⁹

We have chosen to ignore ages below two because there is relatively less information about viewing patterns in those ages. We restrict attention to ages six and below because by age six almost every student in our sample lived in a market in which television broadcasts were available.

Please see Table A.1 for summary statistics.

IV. IDENTIFICATION AND REDUCED-FORM EVIDENCE

IV.A. Identification

The key advance of this study relative to previous work is to identify the effect of television on test scores using variation across local markets in the timing of television's introduction. In Appendix I, we use the Coleman data to examine the potential biases in an approach that uses cross-sectional correlations between television viewing and test scores, as is done in the bulk of the existing literature. We show that virtually every observable characteristic in our data that is related to test scores is also strongly correlated with television viewing hours. Depending on which set of characteristics we include as controls, we can reproduce highly significant partial correlations of television and test scores that are either positive or negative. This suggests that inferring causal relationships from such correlations is a dubious enterprise.

To illustrate our approach to identification, suppose that childhood television viewing has a negative effect on test scores. Consider two cities, an early adopter where television was introduced in 1948, and a late adopter where it was introduced in 1954. In the first city, sixth, ninth, and twelfth graders were all able to watch television throughout childhood (recall that twelfth graders in the Coleman Study were born in 1948). In the second city, sixth graders had television available throughout their lives, but ninth graders only had access to it starting at age three and twelfth graders only at age six. We would therefore expect twelfth

19. For those students for whom we know SMSA but not county, we compute the analogous measure at the SMSA level. Because our measure of television exposure is more precise for the 38% of students for whom the county is known, we have verified that our qualitative conclusions are robust to focusing only on this subsample of students.

graders in the second city to perform well relative to sixth and ninth graders in that city, and ninth graders to perform slightly better than sixth graders. In the first city, we would expect no such pattern. By differencing out the mean test scores by grade from the first city, we could isolate the effects of television using grade patterns in the second city.

A simple way to implement this strategy would be to run a regression of test scores on the number of preschool years that television broadcasts were available in a student's city, controlling for grade and city fixed effects. Cities where television availability did not vary across grades would identify the grade fixed effects; sixth graders, for whom television was available throughout childhood in essentially all cities, would identify the city fixed effects. The remaining variation in the grade pattern of test scores between cities would identify the parameter on years of availability.

Note that the interpretation of these results—denominated in years of television broadcast availability rather than years a child actually had a television in his or her home—would differ greatly depending on the speed at which television ownership diffused. A given effect of a year of television availability could reflect a large effect of exposure if few households actually adopted, or a much smaller effect if adoption was widespread.

In order to make the magnitudes of our coefficients more directly interpretable, we therefore wish to scale our estimates using data on television exposure, constructed as described in the previous section. One way to do this would be to simply replace availability with exposure on the right-hand side. But these results would be identified in part by variation in television purchase decisions—likely to be strongly correlated with county-level unobservables—rather than by variation in the timing of television's introduction. Instead, we adopt a two-stage least squares (2SLS) approach. We include exposure on the right-hand side but instrument for it using data on the year in which television was first introduced. This means that the model will be identified solely by variation in the timing of television's introduction, but the magnitudes will be interpretable as the effect of a year of actual television exposure.²⁰

20. It also means that our estimates (and standard errors) will be consistent even if we measure exposure with error, provided that the error is classical, in the sense of being independent across DMAs.

To state this approach more formally, let y_{gc} be the average test scores of students in grade g in location c , measured as of 1965. Given the geographic information included in the Coleman data, we will define a location to be either an SMSA or a county (for areas not in SMSAs). Let TV_{gc} be the number of years of preschool television exposure of the average student in grade g and location c , constructed as described in Section III.C above. We can write

$$(1) \quad y_{gc} = \beta TV_{gc} + \phi_g W_c + \delta_c + \gamma_g + \varepsilon_{gc},$$

where δ_c and γ_g are location and grade fixed effects, respectively, and ε_{gc} is a city-grade level error term, possibly correlated across grades within a city.²¹

The term $\phi_g W_c$ represents the DMA-level log population and log income of a location W_c multiplied by a grade-specific coefficient vector ϕ_g , where the population and income figures are taken from the 1960 Census. As discussed above, an examination of the historical record suggests that DMA population and income were the most important observable predictors of the timing of television's introduction. Although our identification strategy will rely only on changes across cohorts within a given market (rather than differences across markets), including income and population controls (interacted with grade) will limit the chance that our results will be confounded by unobserved differences in cohort or time trends across markets of different size or wealth.²²

We instrument for TV_{gc} with interactions between grade dummies and dummies for whether the city was an early, middle, or late adopter of television. The first stage of this model can be written as

$$(2) \quad TV_{gc} = \beta_g^0 ADOPT_c + \phi_g^0 W_c + \delta_c^0 + \gamma_g^0 + \varepsilon_{gc}^0,$$

where $ADOPT_c$ is a vector of dummies indicating whether location c was an early, middle, or late adopter of television and β_g^0 is a separate vector of parameters for each grade g . The instruments

21. Note that this model assumes that the effects of television depend only on television ownership, not on the number of viewing hours. Available data on children in television households show no obvious time trend in television hours during the early years of television (see Section II above). This suggests that our instrument may not have had a first-order effect on hours watched, making specification in equation (1) a reasonable approximation.

22. Consistent with evidence that population and income capture the key dimensions of endogeneity in television timing, including additional DMA-level controls interacted with grade (educational attainment, racial composition, median age, and urbanization) does not meaningfully change our results.

$\beta_g^0 \text{ADOPT}_c$ capture the critical cross-city-cross-grade variation in the availability of television that will identify the effect of exposure. The crucial identifying assumption in this model is that, conditional on the controls, the interaction between the timing of television introduction and the birth cohort of the student is orthogonal to the error term in equation (1). Under this assumption, our estimate of the parameter β in equation (1) will be interpretable as the causal effect of an additional year of preschool television exposure on test scores.

Although our model can be estimated with aggregate data alone, we wish to take advantage of the availability of the individual-level data in the Coleman sample. This will allow us to include tighter controls for geography, in particular permitting the use of school rather than location fixed effects. It will also allow us to control for characteristics of individual households and students that might affect exam performance. Both types of information would be expected to improve precision. Of course, because the timing of television introduction is measured at the DMA level, in moving to microdata we must be careful to avoid aggregation bias (Moulton 1990). We will therefore cluster our standard errors at the DMA level, which will also account for any serial correlation across different grades within the same DMA (Bertrand, Duflo, and Mullainathan 2004).²³

In the next subsection, we present OLS estimates of the first-stage equation (2) and of the reduced-form second stage. In Section V we present 2SLS estimates of equation (1). We note that the latter estimates are necessarily local to the students whose exposure to television was affected by the introduction of television (Angrist 2004) so that students in households whose decision to adopt television was more responsive to broadcast availability would implicitly receive more weight in our estimation. In Section VI, we provide evidence on the heterogeneity in treatment effects in the student population and discuss how this heterogeneity is related to television viewership rates.

IV.B. First-Stage and Reduced-Form Estimates

Before estimating model (1) formally, it will be helpful to examine the variation that will identify it. In Figure II, we plot the

23. As expected, specifications in which we aggregate the Coleman test score data to the DMA level and then estimate our model at the aggregate level return similar point estimates with larger standard errors.

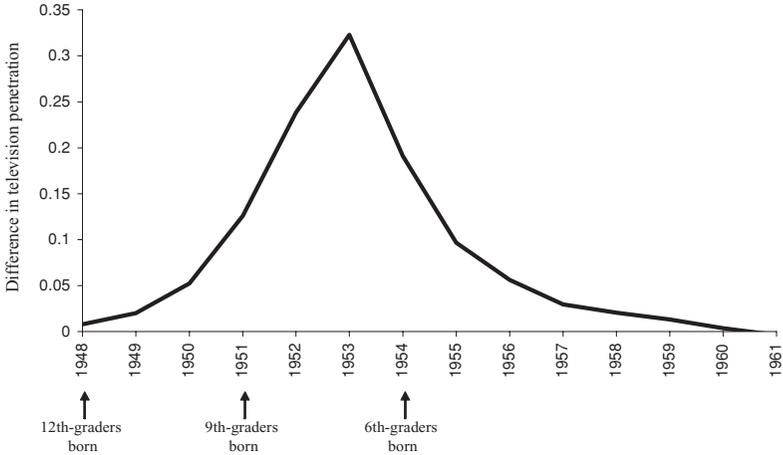


FIGURE II

Difference in Television Penetration between Early/Middle and Late Adopters

Source. Authors' calculations.

Notes. Figure is based on separate year-by-year regressions of television penetration on a dummy for early/middle television adoption (television introduced 1951 or earlier), log income, and log population. The values plotted are the coefficients on the early/middle adoption dummy. The values thus represent the difference in television penetration between early/middle adopters and late adopters in each year, adjusted for differences in log income and log population.

coefficients from year-by-year regressions of DMA television penetration on a dummy for having received a television station before 1952 (controlling for log population and log income). The figure thus shows how pre-1952 television introduction's impact on penetration changes over time. During the period from 1948 to 1954, when the twelfth graders in the Coleman sample were of preschool age, television penetration was substantially higher in early- and middle-adopting DMAs than in late-adopting DMAs. By contrast, in the post-1954 period, when the sixth graders in the sample were preschoolers, the late adopters (most of which received television by 1954) had largely caught up to the early and middle adopters. In other words, differences in adoption dates across DMAs had the largest impact on television exposure for the twelfth graders in the sample, a smaller impact on the ninth graders, and only a minimal impact on the sixth graders. This interaction between a student's grade and the impact of the timing of television introduction is what will allow us to estimate the effect of television exposure on test scores.

Turning to formal estimation, column (1) of Table I presents estimates of the first stage of our model, regressing TV_{gc} on interactions between grade dummies and dummies for whether the city was an early, middle, or late adopter of television. Observe first that, for a given grade, television exposure was lower the later television was introduced to the student's city. So, for example, students in grade nine whose DMAs adopted late were exposed to television for about 0.8 years less than ninth graders whose DMAs adopted early, and about 0.5 years less than those whose DMAs were middle adopters. A similar pattern is present for students in grade 12.

Next, note that, holding constant the timing of television's introduction to a market, twelfth graders on average had less preschool television exposure (between the ages of 2 and 6) than ninth graders, and much less than sixth graders (the omitted category). For example, twelfth graders in late-adopting DMAs had television in their homes for about 1.1 years less than sixth graders in these same DMAs, and about 0.3 years less than ninth graders. This is what we would expect because twelfth graders were born in 1948, ninth graders were born in 1951, and sixth graders were born in 1954. So in cities receiving television after 1948, ninth graders were more likely than twelfth graders to spend their preschool years in a city in which a television signal was available, and sixth graders were almost certain to have grown up with a television in the household.

These findings complement the evidence in Figure I in showing that the timing of broadcast availability had a substantial impact on television penetration and hence on students' exposure to television as young children. Each of the grade-timing interaction terms is strongly individually significant, and the F -test presented in Table I definitively rejects the null hypothesis that these interactions have no impact on exposure.²⁴

In column (2), we present a reduced-form second-stage estimate of the effect of our instruments on test scores. We use as our dependent variable the average of the student's (standardized) scores on the math, reading, verbal, and spatial reasoning tests. If television exposure exerted a negative long-term effect on cognitive skills, we would expect the coefficients on the grade-timing interactions in column (2) to move inversely with

24. The F -statistic in this first-stage model is sufficient to rule out any sizable weak instruments bias (Stock and Yogo 2002).

TABLE I
REDUCED-FORM ESTIMATES OF THE EFFECT OF PRESCHOOL TELEVISION EXPOSURE ON ADOLESCENT TEST SCORES

	Standardized average				Current viewing hours (5)
	Number of years of television exposure (1)	(2)	test score (3)	(4)	
DMA is middle adopter (TV in 1949-1951)					
× Grade 9	-0.2624 (0.0840)	-0.0229 (0.0439)	-0.0002 (0.0296)	-0.0076 (0.0290)	0.0527 (0.0372)
× Grade 12	-0.7137 (0.1626)	-0.0480 (0.0477)	-0.0312 (0.0343)	-0.0347 (0.0319)	0.0853 (0.0560)
DMA is late adopter (TV in 1952 or later)					
× Grade 9	-0.7641 (0.1069)	-0.0066 (0.0443)	0.0033 (0.0333)	-0.0077 (0.0338)	-0.0081 (0.0438)
× Grade 12	-1.0862 (0.2831)	0.0017 (0.0493)	0.0001 (0.0393)	-0.0140 (0.0371)	0.0536 (0.0687)
School and grade fixed effects?	YES	YES	YES	YES	YES
DMA demographics × grade?	YES	YES	YES	YES	YES
Student demographics?	NO	NO	NO	YES	YES
Demographics × grade?	NO	NO	NO	YES	YES
F(4, 135) stat on instruments	16.58	0.91	0.79	0.61	1.44
(p-value)	(<.0001)	.4628	.5334	.6537	.2229
Number of observations	346,562	346,562	346,562	346,562	335,981
Number of schools	800	800	800	800	800
Number of DMAs	136	136	136	136	136

Source: Authors' calculations based on Coleman Study data.
 Notes: Standard errors in parentheses are adjusted for clustering on DMA. DMA demographics include log(DMA population) in 1960 and log(DMA total income) in 1959. Student demographics includes controls for gender, English spoken at home, father's education, mother's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a refrigerator, a vacuum cleaner, or a car. Dummies are included to indicate missing values for demographic controls. The dependent variable in column (5) is the reported average number of daily television viewing hours at the time of the Coleman study (1965).

the coefficients in column (1). In other words, we would expect the students who had relatively less childhood television exposure to perform better on standardized tests. As the column shows, however, we do not see such a pattern. Although students from middle-adopting DMAs perform slightly better than those from late-adopting DMAs, these students perform worse than those from early-adopting DMAs. Additionally, among students from middle-adopting DMAs, twelfth graders perform worse than ninth graders and sixth graders, despite having spent more time without television in their households.

An F -test of the null hypothesis that the grade–timing interactions had no effect on test scores fails to reject at conventional significance levels. Adding demographic controls in columns (3) and (4) improves the precision of our estimates by explaining a larger share of the variation in test scores. These more precise estimates show even less evidence of a negative effect of television. In column (4), where our standard errors are lowest, we find small point estimates on nearly all interaction terms, and the differences among these coefficients do not support the hypothesis of a negative effect of television on test scores.

Finally, in column (5), we present reduced-form second-stage estimates of the effect of our instruments on the number of hours of contemporaneous (1965) television viewing. The grade–timing interactions are both individually and jointly insignificant. This confirms that our estimates will capture the effect of lagged rather than contemporaneous exposure.

V. TELEVISION AND COGNITIVE DEVELOPMENT

V.A. *Two-Stage Least Squares (2SLS) Estimates*

In Table II, we present estimates of equation (1) computed using 2SLS. Coefficients in these models can be interpreted as the causal effect of a year of preschool television exposure on test scores.

We present results for the average test score as well as for each individual component score. For each test, we present baseline estimates, estimates with demographic controls, and estimates with demographics interacted with a student's grade. Adding controls should improve the precision of our estimates by leaving a smaller share of the overall variation in test scores unexplained.

TABLE II
STRUCTURAL (2SLS) ESTIMATES OF THE EFFECT OF PRESCHOOL TELEVISION EXPOSURE ON ADOLESCENT TEST SCORES

	Standardized average test score	Standardized component				
		Math	Spatial reasoning	Verbal	Reading	General knowledge
Effect of number of years of preschool television exposure						
Baseline	0.0129 (0.0384)	-0.0192 (0.0440)	-0.0071 (0.0480)	0.0268 (0.0404)	0.0470 (0.0347)	0.0466 (0.0491)
Baseline + demographics	0.0140 (0.0308)	-0.0154 (0.0376)	-0.0046 (0.0457)	0.0254 (0.0294)	0.0461 (0.0295)	0.0677 (0.0408)
Baseline + demographics × grade	0.0225 (0.0279)	-0.0179 (0.0378)	0.0028 (0.0387)	0.0294 (0.0289)	0.0557 (0.0302)	0.0672 (0.0410)
Number of observations	346,562	346,562	346,562	346,562	346,562	226,487
Number of schools	800	800	800	800	800	705
Number of DMAs	136	136	136	136	136	134

Source. Authors' calculations based on Coleman Study data.
 Notes. Estimates are from 2SLS models with interactions between grade and category of television introduction year used as instruments for years of television exposure. Standard errors in parentheses are adjusted for clustering on DMA. All dependent measures are standardized to have a mean of zero and a standard deviation of unity within each grade. Baseline includes fixed effects for school and grade and interactions between grade and log(DMA population) in 1980 and log(DMA total income) in 1959. Demographics includes controls for gender, English spoken at home, father's education, mother's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a record player, a refrigerator, a vacuum cleaner, or a car. Dummies are included to indicate missing values for demographic controls. General knowledge test scores are available only for students in grades 9 and 12.

The first column shows our estimates of the effect of an additional year of television exposure on the student's average test score, expressed in units of standard deviations (by grade). In general, we find small, statistically insignificant, and positive estimates, suggesting that, if anything, childhood television exposure improves a student's test scores. Adding controls tends to increase the point estimates and, consistent with expectations, decrease the standard errors of these estimates. In the final specification with demographic controls interacted with grade dummies, we are able to reject negative effects of television larger than about 0.03 standard deviations per year of exposure.²⁵

The remaining columns present the estimated effect of television on test scores in each subject separately. In no case do we see clear evidence for a negative effect of television. Column (2) shows that the effects on mathematics and spatial reasoning range from slightly negative to slightly positive and are in all cases statistically insignificant. With our largest set of controls, we find point estimates of -0.018 and 0.003 standard deviations per year of television exposure for mathematics and spatial reasoning respectively. Our point estimates on verbal and reading scores are always positive, with the effect on reading scores a marginally statistically significant 0.06 standard deviations in the final specification ($p = .065$). This in turn means that we can rule out even very small negative effects—our confidence interval in this specification excludes a negative effect on reading scores of about 0.004 standard deviations. Finally, the preferred point estimate for the effect on general knowledge is a positive effect of about 0.07 standard deviations per year of television exposure.

Although we are reluctant to draw firm conclusions from the comparison of coefficients across test scores, we note that the pattern of relatively positive effects on verbal, reading, and general knowledge scores is consistent with a variety of existing evidence suggesting that children can learn language-based skills from television. For example, Rice (1983) argues that the presentation of verbal information on television is especially conducive to learning by young children. Rice and Woodsmall (1988) present laboratory evidence that children aged three and five can learn

25. Because we have multiple instruments, we can perform a test of over-identifying restrictions as an additional check on the validity of the instruments. A test using Hansen's J-statistic (Hansen 1982; Hoxby and Paserman 1998; Baum, Schaffer, and Stillman 2003) cannot reject the null hypothesis that the instruments are uncorrelated with the error term ($J = 1.928, p = .5874$).

unfamiliar words from watching television. The effect on general knowledge scores might also reflect the fact that television also exposes young children to a large number of facts, some of which might be retained into adolescence.²⁶

V.B. Interpretation of Magnitudes

To provide a better sense of the magnitudes of our coefficients and standard errors, we can contrast them with experimental findings in which children exposed to an intervention as preschoolers are followed into adolescence. Perhaps the two best-known instances of such experiments are the Perry Preschool Study and the Carolina Abecedarian Project (Campbell and Ramey 1995; Schweinhart et al. 2005).²⁷ Both studies focused on children from relatively poor families. The Perry study enrolled an intervention group in a two-year part-day preschool education program during ages three and four. The Abecedarian project enrolled children in a five-year full-day day care program through age five. In both cases, children were randomly assigned to intervention and control conditions, and both sets of children were followed into adolescence. In the Perry program, children in the intervention group scored one- to two-thirds of a standard deviation higher on achievement tests at age fourteen (the average age of students in our Coleman sample), with an overall effect of about one-half of a standard deviation. In the Abecedarian program, effect sizes on achievement at age fifteen were on the order of one-third of a standard deviation. Norming these effects for the differences in treatment duration between the studies, the Perry program had an impact on achievement of approximately 0.25 standard deviations per year of intervention, and the Abecedarian program had an impact of approximately 0.07 to 0.08 standard deviations

26. The fact that television exposure improves factual knowledge may also partly explain its effect on reading scores because some evidence indicates that background knowledge can improve reading comprehension (Langer 1984), at least if it is consistent with the information in the test passage (Alvermann, Smith, and Rendence 1985).

27. These are the only two randomized studies receiving detailed attention in Cunha et al.'s (2006) review. Currie's (2001) review identifies two other randomized preschool interventions with long-term follow-up data: the Milwaukee Project and the Early Training Project. The Milwaukee Project offered a five-year full-day day care program through age 5, along with job and academic training for mothers. As of grade 8, the study had an effect of about two-thirds of a standard deviation on IQ (more than 0.1 standard deviations per intervention year) but no statistically significant effect on achievement test scores (Barnett 1995). The Early Training Project, which involved a much less intensive intervention, substantially reduced special education participation in the long term, and had positive, though not statistically significant, long-term effects on student achievement (Currie 2001).

per year.²⁸ The long-term effects of these preschool interventions therefore tend to exceed effects on the order of the low end of our main confidence interval (about 0.03 standard deviations per year of television).

V.C. Specification Checks

Are the Instruments Correlated with Student Characteristics?

The models presented above are valid under the assumption that our instruments—interactions between the timing of television introduction and grade—are orthogonal to the error term. Of course, it is by definition impossible to test this assumption. Some relevant information, however, can be obtained by asking whether television exposure is correlated with trends in observable demographic characteristics. Although the absence of such a correlation is not proof of the identifying assumption, it does provide some confidence that unobserved heterogeneity is unlikely to bias our estimates of the effect of television.

There are two related possibilities we wish to test for. The first is that, in 1965, cross-grade differences in the household characteristics of students within an area are correlated with the timing of television's introduction. The second is that demographic trends during the 1950s are correlated with the timing of television's introduction, resulting in a relationship between a student's preschool television exposure and the local circumstances during his or her upbringing.

To conduct a test for the first possibility, we use the first-stage model (2) to create a predicted number of years of television exposure for each student. By regressing this predicted value on a set of demographic characteristics, we can test whether the variation in television exposure that is due to the timing of television introduction is correlated with cross-grade differences in observable student characteristics that might be expected to affect test scores.

The results of this test are presented in Table III. None of the demographics has a statistically significant correlation with predicted television exposure. Additionally, an *F*-test of the joint hypothesis that none of the demographic characteristics is

28. Additional calculations based on program details imply that the Perry and Abecedarian programs had effects of approximately 0.22 and 0.01 standard deviations per hour-year, respectively. A similar calculation assuming average early-childhood viewing of 1.5 hours per day puts the top end of the confidence interval for our television effects at about 0.02 standard deviations per hour-year.

TABLE III
IS PREDICTED TELEVISION EXPOSURE CORRELATED WITH OBSERVABLES?

	Predicted years of television exposure (1)	Standardized average test score (2)
Male	-0.00087 (0.00090)	-0.0589 (0.0077)
English not spoken at home	0.00152 (0.00100)	-0.2046 (0.0088)
Father's education (index)	-0.00024 (0.00031)	0.0622 (0.0018)
Mother's education (index)	-0.00009 (0.00029)	0.0551 (0.0017)
White	0.00184 (0.00209)	0.4400 (0.0334)
Lives with biological father	-0.00104 (0.00063)	0.0862 (0.0065)
Lives with biological mother	0.00003 (0.00092)	0.1992 (0.0068)
Family has		
Telephone	0.00198 (0.00203)	0.1215 (0.0090)
Record player	0.00082 (0.00077)	0.0246 (0.0050)
Refrigerator	-0.00074 (0.00206)	0.4555 (0.0161)
Vacuum	0.00105 (0.00094)	0.0696 (0.0104)
Car	0.00062 (0.00081)	0.0677 (0.0105)
$F(12, 135)$	1.41	705.63
(p -value)	(.1704)	(<.0001)
Number of observations	346,562	346,562
Number of schools	800	800
Number of DMAs	136	136

Source. Authors' calculations based on Coleman Study data.

Notes. Standard errors in parentheses are adjusted for clustering on DMA. Average test score is standardized to have a mean of zero and a standard deviation of unity within each grade. All regressions include fixed effects for school and grade and interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959. Dummies are included to indicate missing values for demographic controls.

correlated with years of television exposure fails to reject ($p = .170$). Thus we find no evidence of a correlation between the local availability of television and observable characteristics, once we control for DMA-level population and income. This is true despite the fact that, as Table III also shows, these demographic

characteristics are in most cases strong predictors of test scores.²⁹

To test for a bias from differences in time trends in demographics, we have also tested for a relationship between the timing of the introduction of television and changes in income, population density, and adult schooling levels by DMA in the 1950s (see the online appendix for details). We find no statistically significant relationship and no consistent direction of correlation, lending further support to the validity of our identifying assumption.

Are the Instruments Correlated with Teacher Characteristics or School Resources?

Another possibility is that differences in school resources or teacher quality across cohorts are correlated with television entry. Here, again, we must check both for differences in resources at the time of the Coleman Study and different trends in school quality over time.

To check whether contemporaneous (1965) differences in teacher characteristics across grades are correlated with the year of introduction of television, we take advantage of the fact that the Coleman Study collected a set of teacher surveys in addition to student surveys and test scores. Using these, we estimate a regression of predicted television exposure by DMA-grade on the average characteristics of teachers who taught in that grade in 1965. (See the online appendix for details.) Only one of the teacher characteristics (number of subjects taught) is statistically significantly related to predicted television exposure in that grade ($p = .040$). An F -test of the joint significance of the twelve teacher characteristics fails to reject at conventional significance levels ($p = .111$). Additionally, the signs of the coefficients suggest no clear pattern of more resources being associated with greater or lesser television exposure, again supporting the view that there were no systematic cross-grade trends in teacher quality that were correlated with the timing of the introduction of television.³⁰

29. Results are quite similar when we conduct the test on collapsed, DMA-grade-level data: an F -test does not reject the null hypothesis that predicted exposure is uncorrelated with student characteristics ($p = .371$). We have also conducted a parallel exercise (in the spirit of Altonji, Elder, and Taber 2005) in which we predict each student's average test score using her demographics and then use this predicted measure as the dependent variable in 2SLS analysis paralleling Table II. In this case, we again find no evidence of any correlation between our instruments and the demographic predictors of test scores.

30. Consistent with the conclusion that cross-grade variation in teacher characteristics is unrelated to the timing of television's introduction to local areas, we

To check for correlated time trends in school quality, we estimate a regression of the year of television introduction by U.S. state on cohort changes in schooling investments, as measured by Card and Krueger (1992).³¹ (See online appendix for details.) An *F*-test of the joint significance of the changes in schooling investments fails to reject at conventional significance levels ($p = .101$), and the pattern of coefficients does not suggest any consistent relationship between school quality trends and the timing of television introduction.

Are the Results Affected by Mobility?

In our calculations thus far we have implicitly assumed that the students in our sample grew up in the DMA where they currently reside. The Coleman Study provides data on where students report having grown up.³² Roughly 72% of students report having spent most of their lives in their current locality, suggesting that for the bulk of the sample our assignment to DMA will be highly accurate. Another 13% report having spent most of their lives in the same state but in a different city or town, whereas most of the remainder moved at some point from a different state. Because DMAs often include a large fraction of a state's population, many of the 13% who moved from a different town will still be assigned correctly. The assignment of the remaining students will be noisier, but since nearly half of DMAs (and thus television markets) spill across state boundaries, a student's current residence may still contain some information about his or her childhood DMA.

We have estimated our main specification separately for students who grew up in their current state and those who grew up outside of their current state. As expected, our coefficients are generally stronger (more positive) for those in the former group. (See online appendix for details.)

Does Television Exposure Drive Sample Selection?

There are two possible sources of endogenous selection bias in our estimates. The first is that preschool television exposure affects the rate of high school completion and thus affects the

also find that including the full range of teacher characteristics as controls in our 2SLS models yields results very similar to those in Table II (results not shown).

31. The expansion of kindergartens, another important trend in schooling investment, occurred after the television introduction period we study and is therefore not likely to be a confound in our analysis (Cascio 2004).

32. The survey question was either "Where have you spent most of your life?" (grades 9 and 12) or "Where were you born?" (grade 6). Follow-up data collected for a limited subsample suggest that students' responses to this question were accurate in 88% to 98% of cases. See appendix section 9.7 of Coleman (1966).

composition of students who appear in the twelfth grade portion of the Coleman sample. The second is that television exposure affects participation in the Coleman study conditional on being enrolled in school, say, because of effects on attendance.

The evidence in Table III (discussed above) speaks to both of these concerns. It shows that observable correlates of test scores appear to be balanced with respect to preschool television exposure. If television exposure changed the distribution of test scores conditional on selecting into the Coleman sample, we would expect it to affect the conditional distribution of other observable characteristics as well. For example, if exposure caused more low-achieving students to drop out of school between the ninth and twelfth grades, we would expect to see relatively fewer twelfth graders with low test scores in high-exposure areas. However, we would expect to see relatively fewer twelfth graders with low family income and parental education as well. The fact that we do not see this pattern suggests that selection is unlikely to be biasing the results.

To more directly address the possibility that television exposure affects dropout rates, we use Census microdata to study the effect of television exposure on high school completion (see online appendix for details). There is no evidence that preschool television exposure affects the likelihood of having completed high school as of adulthood although we note that the precision of these estimates is lower than the precision of estimates based on the Coleman data.

We have also reestimated our models excluding twelfth graders who are most likely to have selectively dropped out of school prior to surveying (results not shown). As expected, the standard errors of our models increase due to the exclusion of a large portion of the data, but the resulting regressions continue to show no evidence of negative effects of television.

Finally, to investigate effects of television on selection into the pool of Coleman test takers, we have compared the number of students in the Coleman sample with the number we would predict based on principals' reports of school sizes (in the spirit of Jencks [1972]). We find no evidence that television exposure affected rates of inclusion in the Coleman sample.

V.D. Television and Noncognitive Outcomes

The analysis above addresses the effect of television viewing on cognitive development. But it may be that many of television's

TABLE IV
EFFECT OF PRESCHOOL TELEVISION EXPOSURE ON ADOLESCENT SOCIAL AND
BEHAVIORAL OUTCOMES

Dependent variable	Effect of one year of television exposure	<i>N</i>
Number of hours spent on homework each day	0.0148 (0.0422)	334,717
Number of books read during summer (standardized)	-0.0760 (0.0422)	336,127
Student sometimes feels like (s)he "just can't learn"	-0.0040 (0.0208)	327,281
Highest grade student wants to finish in school (standardized)	-0.0265 (0.0221)	333,570
Share of membership organizations	0.0170 (0.0398)	217,392

Source. Authors' calculations based on Coleman Study data.

Notes. Estimates are from 2SLS models with interactions between grade and category of television introduction year used as instruments for years of television exposure. Standard errors in parentheses are adjusted for clustering on DMA. Standardized measures have a mean of zero and a standard deviation of unity within each grade. Baseline includes fixed effects for school and grade and interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959. All regressions include controls for gender, English spoken at home, father's education, mother's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a record player, a refrigerator, a vacuum cleaner, or a car, as well as for interactions between grade dummies and these controls. Dummies are included to indicate missing values for demographic controls. Share of membership organizations is number of the following organizations that the student belongs to, divided by the total number of organizations for which the student provides a response: sports team, Student Council, debate team, and hobby club. Participation in membership organizations is available only for students in grades 9 and 12.

most important effects are on noncognitive traits.³³ We can use the Coleman data to estimate the effect of early childhood television exposure on several social and behavioral outcomes in later years. We note that, as with the previous analysis, our data do not permit us to say anything about the effect of television on contemporaneous noncognitive outcomes. Table IV reports 2SLS estimates of the effect of preschool television exposure on several adolescent attitudinal and behavioral outcomes. The effects are mostly small, negative, and statistically insignificant. The main exception is a marginally statistically significant negative effect on the number of books a student reads during the summer. We also find a statistically insignificant and small positive effect on the number of hours the student spends on homework each day.

We have also used data from the Integrated Public Use Microdata Series (Ruggles et al. 2004) to test for effects of television on long-run labor market outcomes. We extracted information on

33. See, for example, Heckman and Rubinstein (2001) for evidence of the economic value of noncognitive skills.

schooling attainment and wages for cohorts born in 1948, 1951, and 1954 from the 1970, 1980, 1990, and 2000 1% samples of the Census. Information on state of birth provides a coarse measure of the geographic area in which individuals lived in early childhood, allowing us to apply a similar identification strategy to estimate the causal impact of television. The results show no evidence of a negative effect of television, although the coarseness of the geographic identifiers means that the precision of these estimates is limited. Details of this exercise are available in our online appendix.

VI. HETEROGENEITY IN THE EFFECTS OF TELEVISION

Our results thus far focus on the effect of preschool television exposure on the test scores of the average student in our dataset. For many purposes, however, it will be important to know how the effects of television are distributed in the population, especially with respect to the socioeconomic status of the student's household. A body of evidence from developmental psychology shows that the in-home learning environment is richer in higher-socioeconomic-status households, especially with respect to language and vocabulary (Hart and Risley 1995, 1999). Embedded in a simple time-allocation framework (Becker 1965), this evidence would lead one to expect that television is more beneficial to children from more disadvantaged backgrounds, because for such children the activities crowded out by television are likely to be less cognitively stimulating.³⁴ Of course, this prediction could change if richer or more educated parents were better able to select educational programming for their children to watch.³⁵

In this section, we offer evidence on the question of which children benefit the most (or are harmed the least) from television exposure. On the whole, our findings support the prediction that television is most beneficial to children in households with the least parental human capital.

34. The distinction between the direct effect of television content on the viewer and the indirect effect working through displacement of other activities is discussed by Gaddy (1986) and Beentjes and Van der Voort (1988), among others.

35. The evidence we are aware of does not consistently support the view that more educated parents invest more in controlling programming. Rossiter and Robertson (1975) find no evidence of greater supervision by more educated parents, and Pinon, Huston, and Wright (1989) find no relationship between parental education and children's viewing of educational television. One possible explanation is that education also raises parents' opportunity cost of time.

In Table V, we present a first look at heterogeneous effects by splitting the sample along several salient demographic dimensions. The first two columns repeat our basic 2SLS specification for students whose mothers do and do not have a high school education.³⁶ The estimated effect of a year of television exposure on the average test score is 0.04 for students whose mothers have less than a high school education, and 0.01 for students whose mothers have a high school degree. A similar pattern is present for individual test scores.

The next two columns compare households where English was and was not the primary language. The estimated effects of television on verbal, reading, and general knowledge scores for students in non-English-speaking households are positive and nontrivial in magnitude. For the sample of students whose family members primarily speak English, the point estimates are still positive but are much smaller. The point estimates for math and spatial reasoning also suggest more positive effects for students in non-English-speaking households.

In the final two columns, we present results for white and nonwhite students. We find that nonwhite students benefit considerably more from television exposure than do white students. The point estimate of the effect on average test scores is more than 0.05 for nonwhite students, as compared to less than 0.01 for white students.

To combine the information from these various subsample comparisons, we take advantage of a question in the Coleman Study survey that asks students how often they were read to at home prior to starting school. The possible responses range from "never" to "regularly." We anticipate that this measure will be correlated with the overall amount of "quality" time parents spend with their children, and so we treat it as a broad proxy for parental investment. Because parental reading may be directly affected by the availability of television and may also be measured with error,³⁷ we will not use the measure directly but instead use the predicted level as a function of demographics. This will combine variation in parental education, English knowledge, race, and so forth into a single summary measure of parental investment.

36. We obtain similar results using father's education to split the sample rather than mother's education.

37. Response agreement between children and their parents on the question of preschool reading ranged from 60% to 80%, depending on the student's grade (Coleman 1966).

TABLE V
 HETEROGENEITY IN THE EFFECTS OF PRESCHOOL TELEVISION EXPOSURE ON ADOLESCENT TEST SCORES

	Mother has HS degree?		English at home?		White?	
	Yes (1)	No (2)	Yes (3)	No (4)	Yes (5)	No (6)
Effect of number of years of preschool television exposure on						
Average score	0.0135 (0.0297)	0.0418 (0.0334)	0.0157 (0.0277)	0.0766 (0.0479)	0.0026 (0.0262)	0.0526 (0.0489)
Math	-0.0477 (0.0422)	0.0289 (0.0354)	-0.0283 (0.0390)	0.0352 (0.0476)	-0.0293 (0.0405)	-0.0055 (0.0445)
Spatial reasoning	0.0125 (0.0395)	-0.0342 (0.0446)	-0.0050 (0.0362)	0.0592 (0.0608)	0.0225 (0.0277)	-0.0345 (0.0577)
Verbal	0.0308 (0.0294)	0.0482 (0.0398)	0.0252 (0.0285)	0.0740 (0.0529)	-0.0191 (0.0189)	0.1167 (0.0471)
Reading	0.0450 (0.0316)	0.0942 (0.0454)	0.0528 (0.0332)	0.0927 (0.0453)	0.0273 (0.0386)	0.0922 (0.0543)
General knowledge	0.0074 (0.0529)	0.0657 (0.0487)	0.0260 (0.0378)	0.2123 (0.1551)	-0.0945 (0.0423)	0.1401 (0.0786)
Number of observations						
All grades	157,788	112,849	280,455	57,514	211,580	126,886
Grades 9 and 12	110,805	82,874	186,126	36,196	142,328	79,713

Source. Authors' calculations based on Coleman Study data.

Notes. Estimates are from 2SLS models with interactions between grade and category of television introduction year used as instruments for years of television exposure. Standard errors in parentheses are adjusted for clustering on DMA. All dependent measures are standardized to have a mean of zero and a standard deviation of unity within each grade. All specifications include fixed effects for school and grade, interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959, controls for gender, English spoken at home, father's education, mother's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a record player, a refrigerator, a vacuum cleaner, or a car, and demographic controls interacted with grade dummies. Controls for variable on which the sample is split are excluded. Dummies are included to indicate missing values for demographic controls. General knowledge test scores are available only for students in grades 9 and 12.

Formally, let r_i be an index of preschool reading, normalized to have a mean of zero and a standard deviation of one. We will estimate a model of the form

$$(3) \quad y_{gci} = \beta_1(\text{TV}_{gc} \times r_i) + \beta_2\text{TV}_{gc} + X_i\lambda + \phi_g W_c + \delta_c + \gamma_g + \varepsilon_{gci},$$

where we now index individuals by i and explicitly include a vector of individual demographics X_i . We will instrument for the vector $(\text{TV}_{gc}, \text{TV}_{gc} \times r_i)$ with a vector of our television introduction instruments and the instruments interacted with demographic characteristics X_i .

Table VI presents estimates of equation (3). The first row shows the coefficient on the interaction term $(\text{TV}_{gc} \times r_i)$. If the effects of television come mostly through displacement of other activities, we would expect the coefficient β_1 to be negative. The results show that this is indeed the case. With average test scores as the dependent variable, we find that a one-standard-deviation decrease in the parental reading index increases the marginal effect of a year of television exposure by 0.037 standard deviations, and the interaction is statistically significant ($p = .021$). The interactions are of similar magnitudes in the regressions of individual test scores. The interaction effects on math, spatial reasoning, and reading are significant at the 5% level, the interaction effect on verbal is significant at the 10% level, and the interaction effect on general knowledge is not significant.

The bottom two rows of the table show the implied marginal effects at the tenth and ninetieth percentiles of the distribution of the parental reading index. For children with parental investment at the bottom of the distribution, the effects of television are large and positive both on average and for individual test scores. The effect on average test scores is equal to 0.09 standard deviations per year of exposure, and the coefficient is marginally statistically significant ($p = .061$). This point estimate is large, but we note that it is not out of keeping with the effects of other preschool interventions discussed in subsection V.B above. For children with the highest parental investment, the effect of television is negative on average, though none of the point estimates are significantly different from zero.

These findings provide support for the hypothesis that children whose home environments were more conducive to learning were more negatively impacted by television. A possible concern is that the heterogeneity we identify is driven by differences across

TABLE VI
 PARENTAL INVESTMENTS AND THE EFFECTS OF PRESCHOOL TELEVISION EXPOSURE ON ADOLESCENT TEST SCORES

	Standardized component					
	Standardized average test score	Math	Spatial reasoning	Verbal	Reading	General knowledge
Years of television exposure	-0.0371	-0.0314	-0.0364	-0.0312	-0.0300	-0.0269
× Parental reading index	(0.0159)	(0.0158)	(0.0116)	(0.0176)	(0.0151)	(0.0178)
Years of television exposure	0.0174	-0.0123	0.0095	0.0199	0.0361	0.0374
	(0.0296)	(0.0354)	(0.0360)	(0.0309)	(0.0323)	(0.0413)
Implied effect of television exposure by percentile of parental reading index						
Tenth percentile	0.0872	0.0467	0.0781	0.0786	0.0927	0.0880
	(0.0462)	(0.0472)	(0.0447)	(0.0531)	(0.0482)	(0.0576)
Ninetieth percentile	-0.0260	-0.0491	-0.0331	-0.0167	0.0010	0.0060
	(0.0316)	(0.0393)	(0.0365)	(0.0301)	(0.0327)	(0.0428)
Number of observations	263,854	263,854	263,854	263,854	263,854	175,414

Source. Authors' calculations based on Coleman Study data.

Notes. Estimates are from 2SLS models with interactions between grade and category of television introduction year and interactions between these variables and the full set of student demographics, used as instruments for years of television exposure and years of television exposure interacted with the reported frequency with which a child's parents read to him or her in early childhood. Standard errors in parentheses are adjusted for clustering on DMA. Parental reading index standardized to have a mean of zero and a standard deviation of unity. All dependent measures are standardized to have a mean of zero and a standard deviation of unity within each grade. All specifications include fixed effects for school and grade, interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959, controls for gender, English spoken at home, father's education, mother's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a record player, a refrigerator, a vacuum cleaner, or a car, and demographic controls interacted with grade dummies. Dummies are included to indicate missing values for demographic controls. General knowledge test scores are available only for students in grades 9 and 12.

demographic groups in either television penetration or preschool viewing hours rather than differences in the effect of television per se. In Appendix II, we use a limited amount of data on penetration and hours to get a rough estimate of the importance of these confounds. We find that children with lower parental investment are slightly less likely to own televisions but also watch slightly more, conditional on owning. The combined effect of these forces is such that correcting for heterogeneity in penetration and viewership leaves the results essentially unchanged. Although data limitations make these results far from definitive, they suggest that much of the heterogeneity we identify reflects real variation in the effect of television.

The fact that households in which the benefits of television are largest are also those in which children watch the most is an interesting result in its own right. For a more structured test of this hypothesis, we have computed each student's predicted number of television viewing hours per day by regressing reported television viewing on our standard vector of demographics. Among students whose predicted television viewing is above average for their grade, the estimated effect of television exposure is to raise test scores by 0.05 standard deviations. Among those whose predicted viewing is below average, the estimated effect is almost exactly 0. Although by no means conclusive, this pattern is broadly consistent with a model in which television viewing hours are chosen optimally in response to variation across households in the cognitive benefits (or costs) of television exposure.

VII. CONCLUSIONS

In this paper we show that the introduction of television in the 1940s and 1950s had, if anything, small positive effects on the achievement of students exposed to television as preschoolers. Our findings suggest that much of the recent correlational evidence attributing negative developmental effects to childhood television viewing may require reevaluation.

As discussed in the introduction, there are important caveats to these results. First, our data speak only to the effects of early childhood television on academic achievement in adolescence. They do not provide evidence on contemporaneous effects, nor do they provide direct evidence on the effects of television on older children. Second, it is possible that the type and variety

of television content has changed over time in a way that would alter its effects on cognitive development.

We note, however, that there is no obvious reason to presume that changes over time in content would make television's effect more negative. Indeed, Johnson (2005) argues that television programs today are more cognitively demanding than programs in earlier decades, and the most popular shows among the children of 2003 (*The Simpsons*, *American Idol*, *Malcolm in the Middle*) do not seem obviously more or less intellectually rich than those most popular among the children of 1953 (*I Love Lucy*, *Superman*, *The Red Buttons Variety Show*).³⁸ Finally, a large number of well-known educational television programs have been introduced since our sample period, many of which have been linked to improvements in early childhood development (Kaiser Family Foundation 2005).

APPENDIX I: CROSS-SECTIONAL CORRELATION BETWEEN TELEVISION AND TEST SCORES

In this section, we consider what results we might have obtained had we followed the approach of most previous literature on the effects of television: looking at the cross-sectional correlation between television viewership and test scores. The results are informative about the direction and magnitude of biases that may arise in studies that take this approach.

Table A.2 presents regressions of both average test scores and self-reported hours of (contemporaneous) television viewing on demographics. The first half of the table shows coefficients on family background variables, such as race and education. In almost all of these cases, the effects of these demographic characteristics on television hours are statistically significant and in the *opposite* direction from their effects on average test scores. Therefore, we would expect any unobserved variation in these characteristics to tend to bias an OLS regression of test scores on television viewing towards finding negative effects of television. The second half of the table shows that measures of durables ownership—a proxy for family income or wealth—tend to have positive effects on *both* television viewing and test scores,

38. Data on 1953 viewing patterns are from a survey of elementary pupils' "favorite" programs reported in *Television* magazine, April 1955, p. 84. Data on 2003 viewing patterns are from Nielsen audience data for children ages 2–11.

TABLE A.1
SUMMARY STATISTICS BY GRADE

	Grade 6	Grade 9	Grade 12
Years of television exposure	3.966 (0.762)	3.358 (1.210)	1.765 (1.194)
Year of television introduction			
Early (1948 or earlier)	39%	40%	38%
Middle (1948–1951)	25%	22%	23%
Late (1952 or later)	36%	38%	39%
Father's education (1–5)	4.041 (1.992)	3.898 (1.846)	3.793 (1.859)
Mother's education (1–5)	3.986 (1.762)	3.820 (1.532)	3.785 (1.529)
Male	51%	51%	49%
English not spoken at home	18%	17%	15%
White	59%	62%	67%
Lives with biological father	74%	74%	76%
Lives with biological mother	89%	90%	91%
Family has			
Telephone	75%	79%	85%
Record player	81%	87%	89%
Refrigerator	93%	97%	98%
Vacuum	70%	73%	78%
Car	83%	84%	88%
Number of observations	120,075	131,074	95,413

Source. Authors' calculations based on Coleman Study data.

Notes. Standard deviations in parentheses. Table displays the mean value of each variable excluding observations for which the variable is missing. Percentages refer to the share of students with the given characteristics within each grade.

controlling for family background. This finding is not surprising because these proxies for wealth are highly correlated with television ownership and are probably also highly related to the quality of the television set available in the household. So an OLS regression of test scores on television viewing that did not control carefully for family income might find that television has a *positive* effect on student performance. This type of bias seems especially likely in contexts where television ownership is not universal or where quality of sets or programming is likely to be highly variable with income.

These estimates suggest that OLS regressions of test scores on television viewership can easily be subject to upward or downward bias, depending on which household characteristics are measured well and which are measured poorly by the econometrician. OLS regressions of average test scores on self-reported viewing

TABLE A.2
THE CORRELATES OF TELEVISION VIEWING INTENSITY

	Number of hours of television viewing (1)	Standardized average test score (2)
Male	0.0415 (0.0080)	-0.0589 (0.0077)
English not spoken at home	0.0009 (0.0104)	-0.2046 (0.0088)
Father's education (index)	-0.0509 (0.0034)	0.0622 (0.0018)
Mother's education (index)	-0.0306 (0.0043)	0.0551 (0.0017)
White	-0.1750 (0.0241)	0.4400 (0.0334)
Lives with biological father	-0.0537 (0.0071)	0.0862 (0.0065)
Lives with biological mother	0.1029 (0.0107)	0.1992 (0.0068)
Family has		
Telephone	0.1199 (0.0147)	0.1215 (0.0090)
Record player	0.1716 (0.0083)	0.0246 (0.0050)
Refrigerator	0.3482 (0.0320)	0.4555 (0.0161)
Vacuum	-0.0085 (0.0114)	0.0696 (0.0104)
Car	0.0446 (0.0137)	0.0677 (0.0105)
<i>F</i> (12, 135)	236.09	705.63
(<i>p</i> -value)	(<.0001)	(<.0001)
Number of observations	335,981	346,562
Number of schools	800	800
Number of DMAs	136	136

Source. Authors' calculations based on Coleman Study data.

Notes. Standard errors in parentheses are adjusted for clustering on DMA. Average test score is standardized to have a mean of zero and a standard deviation of unity within each grade. All regressions include fixed effects for school and grade and interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959. Dummies are included to indicate missing values for demographic controls.

hours confirm this expectation. When we control for family background measures such as race and education, but not for our wealth proxies, we find an average effect of television viewing that is positive and highly statistically significant. In contrast, when we include wealth proxies but not family background controls, the estimated effect becomes negative and statistically significant.

This finding may help to explain why correlational studies of the effects of television reach highly variable conclusions (Strasburger 1986). Because these studies are only as good as the controls they employ, and because we find that omitted variables problems could lead either to an upward or downward bias of the effects of television, it is not surprising that different academic studies employing different econometric specifications reach radically different conclusions. In a study that controls carefully for family background but not for income, we would expect to find positive effects of television. By contrast, controlling carefully for income or wealth but not for parental education and other background characteristics will lead to a downward bias and findings of deleterious effects of television. We note, however, that we would expect (and preliminary data analysis confirms) that the correlation between household wealth and television viewing hours tends to be negative in more recent data, suggesting an unambiguous downward bias in correlational estimates of the effect of television viewing on test scores.

APPENDIX II: SEPARATING HETEROGENEITY IN THE DIFFUSION, VIEWERSHIP, AND EFFECT OF TELEVISION

The first portion of Table A.3 repeats the estimates in columns (1) and (2) of Table V, which indicate that students whose mothers did not have a high school degree benefited more from television than students whose mothers had a high school degree. These differences may be confounded by the fact that parental education is correlated with both the likelihood of television ownership and children's viewing time. In this section, we show that simple corrections for these confounds have small effects on our substantive conclusions. Indeed, because children with more educated parents were more likely to live in a house with a television set but tended to watch less conditional on ownership, biases from differences in ownership and viewing intensity tend to pull in opposite directions, and therefore to partially cancel one another out.

In the second portion of Table A.3 we adjust our estimates for differences in rates of television penetration across households. We compute average television penetration from 1949 to 1955 for both high-school-educated and non-high-school-educated Gallup poll respondents (Roper Center for Public Opinion Research

TABLE A.3
HETEROGENEITY IN TELEVISION EFFECTS ADJUSTED FOR PENETRATION AND HOURS

	Standardized average test score	Standardized component				
		Math	Spatial reasoning	Verbal	Reading	General knowledge
Unadjusted estimates						
Does mother have high school degree?						
Yes	0.0135 (0.0297)	-0.0477 (0.0422)	0.0125 (0.0395)	0.0308 (0.0294)	0.0450 (0.0316)	0.0074 (0.0529)
No	0.0418 (0.0334)	0.0289 (0.0354)	-0.0342 (0.0446)	0.0482 (0.0398)	0.0942 (0.0454)	0.0657 (0.0487)
Adjusted for differences in average television penetration, 1949-1955						
Does mother have high school degree?						
Yes	0.0117 (0.0257)	-0.0414 (0.0366)	0.0109 (0.0342)	0.0267 (0.0255)	0.0390 (0.0274)	0.0064 (0.0459)
No	0.0474 (0.0378)	0.0328 (0.0401)	-0.0387 (0.0505)	0.0546 (0.0451)	0.1068 (0.0514)	0.0744 (0.0551)
Adjusted for differences in average daily preschool viewing hours and television penetration						
Does mother have high school degree?						
Yes	0.0122 (0.0269)	-0.0432 (0.0382)	0.0114 (0.0358)	0.0279 (0.0267)	0.0408 (0.0286)	0.0067 (0.0479)
No	0.0461 (0.0368)	0.0319 (0.0390)	-0.0376 (0.0491)	0.0531 (0.0439)	0.1038 (0.0500)	0.0724 (0.0536)

Source. Authors' calculations based on Coleman Study data.

Notes. Estimates are from 2SLS models with samples split by mother's education. Standard errors in parentheses are adjusted for clustering on DMA. All dependent measures are standardized to have a mean of zero and a standard deviation of unity within each grade. All specifications include fixed effects for school and grade, interactions between grade and log(DMA population) in 1960 and log(DMA total income) in 1959, controls for gender, English spoken at home, father's education, race, lives with biological father, lives with biological mother, and separate dummies for whether student's family has a telephone, a record player, a refrigerator, a vacuum cleaner, or a car, and demographic controls interacted with grade dummies. Dummies are included to indicate missing values for demographic controls. General knowledge test scores are only available for students in grades 9 and 12.

1949–1955).³⁹ Using these averages, we then compute the ratio of each group's penetration to overall television penetration during this period and scale each coefficient accordingly. Since high-school-educated respondents to the Gallup poll tended to be about 15% more likely to own televisions than the average respondent during this period, we divide the coefficient (and standard error) on television exposure by 1.15 for students whose mothers have high school degrees. Similarly, since Gallup respondents who did not complete high school were about 12% less likely to own a television than the average respondent, we divide the figures for students whose mothers did not complete high school by 0.88. As the second portion of Table A.3 shows, taking these adjustments into account makes little difference for our qualitative conclusions.

In the third portion of Table A.3, we further adjust our estimates to allow for differences in viewing intensity by parental education. We estimate preschool viewing hours for each respondent in the Coleman sample by scaling reported hours of current (1965) daily viewership to reflect the difference in viewing intensity between preschoolers and adolescents.⁴⁰ We then rescale the coefficients for the high- and low-education groups by the ratio of the group's average daily preschool viewing hours to the overall average. Again, this adjustment does not make a substantial difference.

UNIVERSITY OF CHICAGO AND NBER
UNIVERSITY OF CHICAGO AND NBER

REFERENCES

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113 (1) (2005), 151–184.

39. Another way to avoid bias from different penetration rates would be to ask whether television's effect differs in *counties* with either high or low average education levels. The fact that lower education counties might also have less penetration is already corrected for in the estimates because our exposure measure is built from county-level penetration data. We do not report these results here, but they show a similar pattern: in counties with lower-than-median rates of high-school completion, we estimate larger positive television effects on reading, verbal, and general knowledge scores than in counties with above-median education. The effect on reading scores in low-education counties is statistically significant at the 5% level.

40. We scale each student's reported hours of television viewing proportionally so that the average predicted preschool viewing in each grade is equal to Schramm, Lyle, and Parker's (1961) estimates of preschool viewing intensity in the 1950s.

- Alvermann, Donna E., Lynn C. Smith, and John E. Readence, "Prior Knowledge Activation and the Comprehension of Compatible and Incompatible Text," *Reading Research Quarterly*, 20 (4) (1985), 420–436.
- American Academy of Pediatrics, "American Academy of Pediatrics: Children, Adolescents, and Television," *Pediatrics*, 107 (2) (2001), 423–426.
- Angrist, Joshua D. "Treatment Effect Heterogeneity in Theory and Practice," *Economic Journal*, 114 (494) (2004), 83.
- Barnett, W. Steven, "Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes," *The Future of Children*, 5 (3) (1995), 25–50.
- Barnouw, Erik, *Tube of Plenty: The Evolution of American Television*. 2nd rev. ed. (New York: Oxford University Press, 1990).
- Baum, Christopher F., Mark E. Schaffer, and Steven Stillman, "Instrumental Variables and GMM: Estimation and Testing," *Stata Journal*, 3 (1) (2003), 1–31.
- Becker, Gary S. "A Theory of the Allocation of Time," *Economic Journal*, 75 (299) (1965), 493–517.
- Beentjes, Johannes W. J., and Tom H. A. Van der Voort, "Television's Impact on Children's Reading Skills: A Review of Research," *Reading Research Quarterly*, 23 (4) (1988), 389–413.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119 (1) (2004), 249.
- Bishop, John H. "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review*, 79 (1) (1989), 178–197.
- Campbell, Frances A., and Craig T. Ramey, "Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention," *American Educational Research Journal*, 32 (4) (1995), 743–772.
- Card, David, and Alan B. Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100 (1) (1992), 1–40.
- Cascio, Elizabeth U. "Schooling Attainment and the Introduction of Kindergartens into Public Schools," mimeo, University of California, Davis, April 2004.
- Christakis, Dimitri A., Frederick J. Zimmerman, David L. DiGiuseppe, and Carolyn A. McCarty, "Early Television Exposure and Subsequent Attentional Problems in Children," *Pediatrics*, 113 (4) (2004), 708–713.
- Coleman, James Samuel, and United States Office of Education and National Center for Educational Statistics, *Equality of Educational Opportunity* (Washington, DC: U.S. Department of Health Education and Welfare Office of Education, 1966).
- Cook, Thomas D., Hilary Appleton, Ross F. Conner, Ann Shaffer, Gary Tamkin, and Stephen J. Weber, "Sesame Street" Revisited (New York: Russell Sage Foundation, 1975).
- Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov, "Interpreting the Evidence on Life Cycle Skill Formation," E. Hanushek and F. Welch, eds., in *Handbook of the Economics of Education*, 1st ed., Handbooks in Economics, Vol. 1 (Amsterdam: North-Holland, 2006).
- Currie, Janet, "Early Childhood Education Programs," *Journal of Economic Perspectives*, 15 (2) (2001), 213–238.
- Dahl, Gordon, and Stefano DellaVigna, "Does Movie Violence Increase Violent Crime?" mimeo, UC Berkeley, June 2006.
- DellaVigna, Stefano, and Ethan Kaplan, "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122 (3) (2007).
- Diaz-Guerrero, Rogelio, Isabel Reyes-Lagunes, Donald B. Witzke, and Wayne H. Holtzman, "Plaza Sesamo in Mexico: An Evaluation," *Journal of Communication*, 26 (2) (1976), 145–154.
- Djankov, Simeon, Caralee McLiesh, Tatiana Nenova, and Andrei Shleifer, "Who Owns the Media?" *Journal of Law and Economics*, 46 (2) (2003), 341–381.
- Ehrenberg, Ronald G., and Dominic J. Brewer, "Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited," *Economics of Education Review*, 14 (1) (1995), 1.
- Eisenberg, Azriel Louis, *Children and Radio Programs: A Study of More than Three Thousand Children in the New York Metropolitan Area* (New York: Columbia University Press, 1936).

- Fox Meadow School PTA, "Radio for Children—Parents Listen In," *Child Study Magazine*, 11 (7) (1933).
- Gaddy, Gary D., "Television's Impact on High School Achievement," *Public Opinion Quarterly*, 50 (3) (1986), 340–359.
- Gentile, D. A., C. Oberg, N. E. Sherwood, M. Story, D. A. Walsh, and M. Hogan, "Well-Child Visits in the Video Age: Pediatricians and the American Academy of Pediatrics' Guidelines for Children's Media Use," *Pediatrics*, 114 (5) (2004), 1235–1241.
- Gentzkow, Matthew, "Television and Voter Turnout," *Quarterly Journal of Economics*, 121 (3) (2006), 931–972.
- Gentzkow, Matthew, and Jesse M. Shapiro, "Media Bias and Reputation," *Journal of Political Economy*, 114 (2) (2006), 280–316.
- "Media, Education, and Anti-Americanism in the Muslim World," *Journal of Economic Perspectives*, 18 (3) (2004), 117–133.
- Glenn, Norval D., "Television Watching, Newspaper Reading, and Cohort Differences in Verbal Ability," *Sociology of Education*, 67 (3) (1994), 216–230.
- Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, *Journal of the Econometric Society*, 25 (4) (1957), 501–522.
- Griliches, Zvi, and William M. Mason, "Education, Income, and Ability," *Journal of Political Economy*, 80 (3) (1972), S74–S103.
- Hancox, Robert J., Barry J. Milne, and Richie Poulton, "Association of Television Viewing during Childhood with Poor Educational Achievement," *Archive of Pediatrics & Adolescent Medicine*, 159 (7) (2005), 614–618.
- Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (4) (1982), 1029–1054.
- Hanushek, Eric A., and John F. Kain, "On the Value of Equality of Educational Opportunity as a Guide to Public Policy," in *On Equality of Educational Opportunity*, Frederick Mosteller and Daniel P. Moynihan, eds. (New York: Random House, 1972).
- Harrison, Linda Faye, and Tannis MacBeth Williams, "Television and Cognitive Development," in *The Impact of Television: A Natural Experiment in Three Communities*, Tannis MacBeth Williams, ed. (London: Academic Press, 1986).
- Hart, Betty, and Todd R. Risley, *Meaningful Differences in the Everyday Experience of Young American Children* (Baltimore, MD: Paul H. Brookes Publishing Co., 1995).
- *The Social World of Children Learning to Talk* (Baltimore, MD: Paul H. Brookes Publishing Co., 1999).
- Heckman, James J., "Randomization as an Instrumental Variable," *Review of Economics and Statistics*, 78 (2) (1996), 336–341.
- Heckman, James J., and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 91 (2) (2001), 145–149.
- Heckman, James J., and Guilherme Sedlacek, "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection into the Labor Market," *Journal of Political Economy*, 93 (6) (1985), 1077–1125.
- Hennigan, Karen M., Linda Heath, J. D. Wharton, M. L. Delrosario, T. D. Cook, and B. J. Calder, "Impact of the Introduction of Television on Crime in the United States: Empirical-Findings and Theoretical Implications," *Journal of Personality and Social Psychology*, 42 (3) (1982), 461–477.
- Hoxby, Caroline M., and M. Daniele Paserman, "Overidentification Tests with Grouped Data," NBER Technical Working Paper 223 (February 1998).
- Jencks, Christopher S., "The Quality of the Data Collected by the Equality of Educational Opportunity Survey," in *On Equality of Educational Opportunity: Papers Deriving from the Harvard University Faculty Seminar on the Coleman Report*, Frederick Mosteller and Daniel P. Moynihan, eds. (New York: Vintage Books, 1972).
- Johnson, Steven, *Everything Bad Is Good for You: How Today's Popular Culture Is Actually Making Us Smarter* (New York: Riverhead Hardcover, 2005).
- Kaiser Family Foundation, "The Effects of Electronic Media on Children Ages Zero to Six: A History of Research," 2005.

- Langer, Judith A., "Examining Background Knowledge and Text Comprehension," *Reading Research Quarterly*, 19 (4) (1984), 468–481.
- Maccoby, Eleanor E., "Television: Its Impact on School Children," *Public Opinion Quarterly*, 15 (3) (1951), 421–444.
- Moulton, Brent R., "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit," *Review of Economics and Statistics*, 72 (2) (1990), 334.
- Olken, Benjamin A., "Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages," NBER Working Paper No. 12561, October, 2006.
- Parker, Edwin B., "The Effects of Television on Public Library Circulation," *Public Opinion Quarterly*, 27 (4) (1963), 578–589.
- Pinon, Marites F., Aletha C. Huston, and John C. Wright, "Family Ecology and Child Characteristics that Predict Young Children's Educational Television Viewing," *Child Development*, 60 (4) (1989), 846–856.
- Rice, Mabel L., "The Role of Television in Language Acquisition," *Developmental Review*, 3 (1983), 221–224.
- Rice, Mabel L., and Linda Woodsmall, "Lessons from Television: Children's Word Learning When Viewing," *Child Development*, 59 (2) (1988), 420–429.
- Rideout, Victoria J., Elizabeth A. Vandewater, and Ellen A. Wartella, "Zero to Six: Electronic Media in the Lives of Infants, Toddlers, and Preschoolers," Tech. rept., Kailer Family Foundation, 2003.
- Roper Center for Public Opinion Research, *Gallup Polls*, Vols. USAIPO1949-0441, USAIPO1950-0460, USAIPO1951-0470, USAIPO1953-0513, USAIPO1953-0514, USAIPO1953-0522, USAIPO1954-0529, USAIPO1954-0537, USAIPO1955-0546, USAIPO1955-0549 (Mansfield, CT: Gallup Organization, 1949–1955).
- Roslow, Sydney, "Programming Trends," *Television*, 9 (4) (1952), 22–23.
- Rossiter, John R., and Thomas S. Robertson, "Children's Television Viewing: An Examination of Parent-Child Consensus," *Sociometry*, 38 (3) (1975), 308–326.
- Roy, A. D. "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (2) (1951), 135–146.
- Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander, *Integrated Public Use Microdata Series: Version 3.0* (www.ipums.org, 2004).
- Schramm, Wilbur, Jack Lyle, and Edwin B. Parker, *Television in the Lives of Our Children* (Stanford, CA: Stanford University Press, 1961).
- Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High / Scope Perry Preschool Study Through Age 40*, Monographs of the High/Scope Educational Research Foundation (Ypsilanti, MI: High/Scope Press, 2005).
- Slotten, Hugh Richard, *Radio and Television Regulation: Broadcast Technology in the United States, 1920–1960* (Baltimore: Johns Hopkins University Press, 2000).
- Sterling, Christopher H., and John M. Kittross, *Stay Tuned: A History of American Broadcasting*, 3rd ed., LEA's Communication Series (Mahwah, NJ: Lawrence Erlbaum Associates, 2001).
- Stock, James H., and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," NBER Technical Working Paper 284, 2002.
- Strasburger, V. C., "Does Television Affect Learning and School Performance?" *Pediatrician*, 13 (2-3) (1986), 141–7.
- Stromberg, David, "Radio's Impact on Public Spending," *Quarterly Journal of Economics*, 119 (1) (2004), 189–221.
- Television Bureau of Advertising, *TV Basics 2003: A Report on the Growth and Scope of Television* (<http://www.tvb.org/rcentral/mediatrendstracks/tvbasics>, October 19, 2003).
- Television Digest, *Television Factbook* (Washington, DC: Television Factbook, Inc., various years).
- Television Magazine, "Programming," *Television*, 5 (1) (1948), 41–43.
- "Four Hundred Thousand Dollars a Week for the Children's Market," *Television*, 8 (8) (1951), 21–22.
- "The Child Audience," *Television*, 12 (4) (1955), 83–84.
- *Market Book* (New York: Frederick Kugel Company, various years).

- Van Evra, Judith Page, *Television and Child Development* (Mahwah, NJ: Lawrence Erlbaum Associates, 1998).
- Winn, Marie, *The Plug-In Drug: Television, Computers, and Family Life* (New York: Penguin, 2002).
- Wirtz, Willard, Harold Howe II, Bernard C. Watson, Ralph W. Tyler, Ledyard R. Tucker, Vivian H. T. Tom, Robert L. Thorndike, Barbara Thompson, Thomas W. F. Strand, Rosedith Sitgreaves, Wilbur Schramm, Katherine P. Layton, Owen B. Kiernan, H. Thomas James, Matina S. Horner, Edythe J. Gaines, Frank W. Erwin, Bruce K. Eckland, Luis C. Cortez, Sandra A. Clark, and Benjamin S. Bloom, *On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline* (New York: College Entrance Examination Board, 1977).
- Zavodny, Madeline, "Does Watching Television Rot Your Mind? Evidence of the Effect on Test Scores," *Economics of Education Review*, 25 (5) (2006), 565–573.
- Zimmerman, Frederick J., and Dimitri A. Christakis, "Children's Television Viewing and Cognitive Outcomes: A Longitudinal Analysis of National Data," *Archive of Pediatrics & Adolescent Medicine*, 159 (7) (2005), 619–625.