Available online at www.sciencedirect.com
SCIENCE DIRECT®
Economics of Education Review 25 (2006) 565–573
Economics of Education Review
www.elsevier.com/locate/econedurev

Does watching television rot your mind? Estimates of the effect on test scores

Madeline Zavodny*

Box 1189, Agnes Scott College, 141 E. College Ave., Decatur, GA 30030, USA

Received 26 February 2004; accepted 8 August 2005

Abstract

This study examines whether the number of hours of television watched by young adults is associated with performance on standardized exams and whether any such relationship is causal. Data from the National Longitudinal Survey of Youth, the High School and Beyond survey and the National Education Longitudinal Study all indicate a negative cross-sectional relationship between hours of television viewing and test scores, even after controlling for a variety of socioeconomic characteristics. However, endogeneity bias may underlie this negative relationship. Models that include individual or family fixed effects to partially control for endogeneity suggest that television viewing does not negatively affect performance on standardized exams.

JEL classification: J22; I20

Keywords: Human capital; Educational economics

1. Introduction

Television is one of the most omnipresent features of Americans' lives. The average American adult watches about 15 h of television per week, accounting for almost one-half of free time (Robinson & Godbey, 1997). Television viewing accounts for the second largest amount of waking hours—after attending school—among US children (Juster & Stafford, 1991). The substantial amount of time that most individuals spend watching television makes it important to examine its effects on society, including human capital accumulation and academic achievement. This study uses scores on standardized tests as a proxy for these variables and estimates the relationship between television viewing and test scores among young adults. The cross-sectional results here indicate that watching more television is negatively associated with contemporaneous scores on achievement tests, but the results from individual and family fixed effects models do not indicate an adverse effect.

There are several reasons why watching television might affect test scores and, more generally, human capital.[1] Television viewing may displace other

[1] It should be noted that there is substantial disagreement about whether standardized test scores are accurate measures of human capital, either innate or acquired. Critics contend that standardized tests such as the AFQT contain racial, socioeconomic and

*Tel.: +1 404 471 6377; fax: +1 404 471 5478.
E-mail address: mzavodny@agnesscott.edu.

0272-7757/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.econedurev.2005.08.003

activities that contribute to children's human capital, such as studying, reading and participating in extracurricular activities. In addition, critics often charge that watching television shortens youths' attention spans, causing them to learn less in school and hampering their performance on exams. Alternatively, television viewing may be beneficial, exposing people to cultural events and ideas they would not otherwise encounter.

Television viewing also may have no causal effect on outcomes such as test scores. Instead, the number of hours of television an individual watches may merely reflect other individual characteristics that are associated with test scores, such as ability, socioeconomic status or parental involvement. A common set of factors—some of which may be unobservable—may affect both television viewing and test scores.[2] In addition, young adults who watch more television may choose to do so because they have relatively low returns to acquiring more human capital through studying or other such activities that are substitutes for watching television; heavy television viewers may have a relatively low opportunity cost of watching television. This suggests that television viewing is endogenous in human capital models—such as models of the determinants of test scores—because the amount of television individuals watch depends on their return to human capital. The analysis here focuses on partially controlling for endogeneity in estimates of the relationship between television viewing and test scores by controlling for a wide variety of individual and family characteristics and by including individual or family fixed effects in regression models.

Previous findings on the relationship between television viewing and test scores are mixed. A study of math achievement test scores using panel data from the National Education Longitudinal Study (NELS) reported a negative association between the number of hours of television watched on weekdays and test scores (Aksoy & Link, 2000). However, studies using panel data from the High School and Beyond (HSB) survey and the National Health Examination Survey did not find an association between the amount of television viewed and test scores, particularly after controlling for other socioeconomic factors (Gaddy, 1986; Gortmaker, Salter, Walker, & Dietz, 1990). Cross-sectional studies typically have found a negative relationship between television viewing and test scores (e.g., Hornik, 1981; Keith, Reimers, Fehrmann, Pottebaum, & Aubey, 1986). This negative cross-sectional relationship tends to persist, particularly for reading test scores, when controlling for IQ (Morgan & Gross, 1980).

A few previous studies explicitly addressed whether endogeneity underlies the negative relationship between television viewing and test scores found by some researchers. Consistent with the hypothesis that individuals with relatively low opportunity costs of watching television spend more hours doing so, Keith et al. (1986) reported that low-ability young adults watch more television than their more able peers. In addition, they found that the estimated relationship between hours viewed and test scores declines by three-fourths when controlling for family background and other factors in cross-sectional data, suggesting that only about one-fourth of the association between television viewing and test scores is causal. Also consistent with the endogeneity hypothesis, Gortmaker et al. (1990) reported that children with lower test scores during childhood tend to watch more television during adolescence.

This study makes several contributions to our understanding of the effect of watching television on test scores. Although there is considerable research on the relationship between television viewing and test scores, most previous studies treated television viewing as exogenous and controlled for a limited set of covariates. Earlier research that used panel data techniques to partially control for endogeneity reached conflicting results. This study uses several methods to control for endogeneity bias: a number of variables are included to control for observable characteristics;

(*footnote continued*)
cultural biases (e.g., Maume, Cancio, & Evans, 1996; Rodgers & Spriggs, 1996). Concerns about such biases in test scores are not relevant for the relationship between television viewing and test scores focused on here unless watching television affects test scores differentially across groups, such as Blacks and Whites. If so, then stratifying the data by race or interacting television viewing with race would control for this effect.

[2]For example, parental involvement may affect how many hours of television a child watches and also influence academic achievement, including performance on standardized tests. Regressions that fail to control for such common factors would mistakenly attribute the effect of these factors on test scores to hours of television watched. The direction of this bias is negative if these factors are negatively associated with the amount of television watched but positively associated with test scores. Such regressions therefore overestimate any negative relationship between television viewing and test scores.

Table 1
Summary of samples from data sets

	NLSY	HSB	NELS
Number of individuals	2477	14,988	6255
Type of data	Cross-section	Panel	Panel
Time period	1981	1980, 1982	1988, 1990, 1992
Age of respondents	16–19	13–22	12–20
Sibling structure	239 sibling pairs	432 twin pairs	None
Test scores	AFQT	Vocabulary, reading, math	Reading, math
TV variable	Hours per week	Hours per weekday	Hours per weekday, weekend day
Categorical or linear	Linear	Categorical (7)	Categorical (6 or 7)

Notes: The number of observations is the number of individuals from each data set included in the samples used here, not the total number available in the raw data. One of the pairs of twins in the HSB consists of three triplets. The twins data in the HSB are a cross-section in 1980.

individual fixed effects are included to control for unobservable, time-invariant characteristics; comparisons between siblings, including twins, are made to control for unobservable family characteristics; and the same regressions are estimated using three different data sets. Previous research has not used sibling comparisons to estimate the effect of watching television on test scores nor estimated panel data models for different data sets and a common set of covariates.

2. Data

This analysis uses three data sets to examine the relationship between television viewing and test scores: the National Longitudinal Survey of Youth 1979 (NLSY), the HSB survey and the NELS. Each survey includes test scores and a question about the number of hours of television watched by young adults.[3] The tests and the wording of the television question differ across the three studies. The HSB survey and the NELS have multiple observations on test scores and television viewing for some participants, which allows for the inclusion of individual fixed effects. The HSB has a sample of twins while the NLSY has a sample of siblings, allowing for family fixed effects to be included. The data used here from the NLSY and the HSB are from the early 1980s while the NELS data are from the late 1980s and early 1990s. Table 1 provides a summary of the data sets.

[3]None of the surveys asks about television program content, such as news and educational programs versus sports and sitcoms. Studies that focus on effects of program content are largely limited to young children.

2.1. NLSY

The NLSY began in 1979 with a survey of young adults who were born between 1957 and 1964. These young adults were periodically resurveyed. The survey design includes a nationally representative cross-sectional sample, an oversample of Blacks, Hispanics and low-income Whites and a military sample.[4] Only individuals enrolled in high school at the time of the 1981 survey are included here in order to make the NLSY sample somewhat comparable in age to the other data sets.[5] The NLSY sample used here has 2477 individuals who have valid test scores. This sample includes 239 families with at least two survey participants who are biological siblings.[6]

The NLSY contains two items that are particularly relevant for this study: a self-report of the number of hours of television watched during the week prior to the 1981 survey, and individuals' scores on a standardized test administered during the summer of 1980. The television variable is a linear variable topcoded at 96 h. The average person in the NLSY sample reported watching about 14 h of television during the previous seven days. The

[4]The oversample is an additional sample of Blacks, Hispanics and low-income Whites added to the NLSY's nationally representative sample to ensure adequate sample sizes for those groups. The NLSY weights used here make the entire sample nationally representative.

[5]Using a sample largely young enough to be subject to compulsory schooling laws also avoids the problem of years of educational attainment potentially affecting AFQT scores (Neal & Johnson, 1996).

[6]Sibling sets that indicated that they are step-siblings are not included.

standardized test is the Armed Forces Vocational Aptitude Battery (ASVAB), which consists of 10 parts. The NLSY reports individuals' scores as a percentile based on the four ASVAB parts that focus on verbal and math skills. This percentile score is commonly referred to as the AFQT, the Armed Forces Qualifying Test.[7]

2.2. HSB

The HSB is a longitudinal study that began in 1980 with students who were then high school sophomores and seniors. An attempt was made to reinterview all individuals in 1982. Survey participants were given a set of achievement tests in 1980, and the sophomore sample was retested in 1982 (when most were seniors). The tests include a vocabulary test, a reading test and a math test. The HSB reports individuals' test scores normalized within the samples by original grade, and the average of the normalized scores on the vocabulary, reading and math tests is used here as the primary test measure. Only sophomores are included in the main (non-twin) HSB sample used here so that two test scores, for 1980 and 1982, are available for each individual. This sample includes 14,988 individuals who were still in school in 1982.

The HSB study included a special sample of individuals who had a twin (or triplet) in the main sample. These twins were given the achievement tests in 1980. The special sample of twins was not resurveyed later, so the twins data are cross-sectional. The twins data used here has 431 sets of twins and one set of triplets.[8]

The HSB asked individuals in 1980 and 1982 how many hours of television they watched during a weekday within seven intervals. The weekday measure in the HSB survey is multiplied by seven and treated as a linear variable in this analysis to make it roughly comparable to the NLSY weekly television viewing variable.[9] By this measure, the average person in the HSB sample watched almost 21 h of television per week.

2.3. NELS

The NELS is a longitudinal study patterned after the HSB study. A sample of eighth graders was surveyed in 1988, and this sample was resurveyed at 2-year intervals through 1994. Participants were given achievement tests in a variety of subjects in 1988, 1990 and 1992. The NELS reports these test scores, including a composite test score for each year based on the reading and math test scores, as percentiles. This analysis uses these composite test scores. The NELS sample used here includes 6255 individuals who were surveyed in 1988, 1990 and 1992, have valid test scores each survey wave and remained in school through 1992 (when most were seniors).[10]

The first three waves of the NELS asked about the number of television hours per day usually watched on weekdays and, separately, on weekend days. The first two survey waves had seven intervals for the television questions, and the third wave six intervals.[11] The measure of weekly television viewing used here for each survey year is average weekday hours times five plus average weekend-day hours times two. By this measure, the average NELS respondent in the sample watched almost 19 h of television a week.

A common set of factors that are likely to affect both test scores and television watching was created from the NLSY, HSB and NELS data sets.[12] These factors can be divided into three categories: individual characteristics, family background and other uses of time. The set of individual characteristics includes age, sex, race (Black), ethnicity (Hispanic) and immigrant

[7]This study uses the 1989 revision of the AFQT scores.

[8]Seniors are not included in the main HSB sample (but are included in the twins sample) because the aptitude tests were not readministered to 1980 seniors in 1982. Differences between siblings who are a sophomore and a senior in 1980 are not examined here because the HSB reports test scores normed within grades.

[9]The seven intervals are none, less than 1 h, at least 1 h but less than 2 h, 2 h but less than 3 h, etc., up to 5 h or more hours. Intervals other than none were converted to the middle of the interval (e.g., 1.5 h for at least 1 h but less than 2 h) and multiplied by seven. The highest interval is converted here to 5.5 h (multiplied by seven). All results are qualitatively similar if a variable ranging from zero to six is included in the regressions instead of the weekly approximation or if other numbers are used to construct the highest value (6.5 h, 7.5 h, etc.).

[10]The NELS added new respondents in 1990 and 1992; these individuals are not included here. The NELS did not report a composite test score for 1992, so the average of the reading and math test score percentiles is used as the composite test score. The data used here are from the Base Year through Third Follow-Up version of the data.

[11]The 1988 and 1990 NELS intervals are the same as the HSB intervals. The 1992 NELS has 3–5 h as a single interval. As with the HSB data, the middle of the interval was used to construct the television variable.

[12]An appendix table that presents descriptive statistics for the variables is available from the author on request.

status.[13] The set of family background characteristics is the total number of siblings, the availability of reading materials in the household (measured here as two indicator variables for no reading materials and numerous reading materials, with the in-between group as the omitted category) and parents' education (four indicator variables for whether each parent is a high school graduate or a college graduate).[14] The two measures of other uses of time are how much time individuals spend reading and doing homework.[15]

3. Methods

The basic regression model for the cross-sectional models of test scores is

$$\begin{aligned}\text{Test Score}_{it} = {} & \alpha + \upsilon \text{Hours of TV}_{it} \\ & + \beta \text{Individual Characteristics}_{it} \\ & + \gamma \text{Family Background}_{it} \\ & + \delta \text{Other Uses of Time}_{\text{it}} + \tau T_{it} + \varepsilon_{it},\end{aligned} \tag{1}$$

where i indexes individuals and t indexes survey years. The left-hand side variable, *Test Score*, is normalized to a sample mean of zero and standard deviation of one.[16] The estimated coefficients therefore give the relationship between a change in a covariate—such as watching an additional hour of television per week—and the change in test scores relative to the standard deviation.

All versions of Eq. (1) estimated here include the measure of television viewing. The individual characteristics, family background and other uses of time vectors are added consecutively to the model in order to examine how the coefficient of the television variable changes as the regression controls for other factors that may affect both television viewing and test scores. As discussed above, not controlling for common factors is likely to negatively bias the television coefficient, so it should become more positive as the regression controls for those factors. Because the HSB and NELS surveys were administered to multiple students at the same school, some specifications using those data also include school fixed effects to control for school-level factors that may influence test scores. The vector T in Eq. (1) represents survey year fixed effects and is included when the model is estimated using data from the HSB or NELS.

Another way to partially control for endogeneity besides adding more control variables is to include individual or family fixed effects. The individual fixed effects control for unobservable, time-invariant individual characteristics that influence test scores and also may be associated with television viewing, such as ability. The individual fixed effects regressions, which are estimated using the HSB and NELS panel data, give the relationship between changes in television viewing and changes in test scores over time. The family fixed effects control for unobservable factors that are common across siblings, such as parental characteristics. The siblings and twins regressions with family fixed effects, which are estimated using NLSY and HSB cross-sectional data, respectively, give the relationship between differences in television viewing and differences in test scores across siblings.

The basic regression model for the fixed effects models of test scores is

$$\begin{aligned}\text{Test Score}_{ift} = {} & \alpha + \upsilon \text{Hours of TV}_{ift} \\ & + \sigma \text{FE}_{if} + \tau T_{it} + \varepsilon_{ift},\end{aligned} \tag{2}$$

where i indexes individuals, f indexes families and t indexes survey years. The family subscript applies to the family fixed effects regressions and the time subscript applies to the individual fixed effects regressions; none of the data sets used here has

[13] The regressions do not control for educational attainment for two reasons. First, educational attainment may be endogenous in test score regressions. In addition, the distribution of educational attainment is narrow in the NLSY sample, which is limited to individuals enrolled in high school in 1981. Education does not vary in the HSB and NELS base year samples and varies in later years only because of students skipping grades or being held back.

[14] For the NLSY sample, no reading materials is defined as the household not having a subscription to a newspaper, not having a subscription to a magazine and not having a library card when the respondent was 14 years old, and numerous reading material as having all three. For the HSB and NELS samples, no reading materials is defined as the household not receiving a newspaper and the household not having more than 50 books at the time of the survey, and numerous reading materials is defined as having both. Indicator variables for mother's education and father's education not reported are included in the regressions.

[15] For the NLSY data, these variables are scaled as hours per week. The HSB survey has ordinal response intervals for time spent reading for pleasure (four intervals, ranging from rarely or never to every day or almost every day) and on homework (seven intervals), which are used here as linear variables. The NELS has six (eight) response intervals for the reading (homework) question in 1988 and eight (nine) in 1990 and 1992, which are collapsed into the same six (eight) intervals as in 1988.

[16] The test scores are not corrected for age; instead, age is included directly in the regressions. The results are similar if the test scores are first regressed on age and then the normalized residuals are used as the dependent variable.

multiple observations on siblings over time. The fixed effects regressions do not include the controls for individual characteristics, family background or other uses of time because there is relatively little or no variation over time or within families in those variables.

All regressions are estimated using ordinary least squares (OLS). Separate regressions are estimated for each data set. The cross-sectional test score regressions are estimated separately by sex and by racial/ethnic group as well as for the entire samples to investigate whether the relationship between test scores and television viewing differs across groups.[17] Standard errors are White-corrected for heteroscedasticity and clustered on the individual (on the family in the family fixed effects regressions). The NLSY data are weighted using the 1981 survey weights. The HSB and NELS data are weighted using the panel data test weights, which account for attrition and failure to take the tests.

3.1. Methodological issues

The main methodological question that arises from the OLS and fixed effects approaches used here to partially control for the endogeneity of television viewing in models of test scores is, why not use instrumental variables or simultaneous equations methods instead? Identification in these alternate methods requires a variable that affects television viewing but does not also influence test scores (and vice versa for the simultaneous equations approach). There is no such obvious variable in the data. Geographic variation in television viewing is a possibility—perhaps because average temperature or time zone affects television viewing—but geographic variation at the MSA, county or state level accounts for little of the variation in television viewing across individuals in the NLSY. In addition, geographic variation in access to cable television or in the number of broadcast channels was not significantly associated with geographic variation in hours of television watched for the NLSY sample. Thus, although instrumental variables or simultaneous equations models would be a better way to control for endogeneity, it is not possible to identify those models with these data sets.

4. Results

4.1. Cross-sectional estimates

The cross-sectional results indicate that young adults who watch more hours of television have significantly lower test scores, but the estimates are small in magnitude and—consistent with endogeneity bias—become smaller as more controls are added. As Table 2 shows, in regressions that pool sexes and racial/ethnic groups and include only the measure of television viewing, an additional hour of television per week is associated with a decline in test scores of about .01–.02 standard deviations (column 1, specification A). Controlling for age, sex, race/ethnicity and immigrant status reduces these already small estimates by about 15–30 percent (specification B). Adding controls for the number of siblings, parents' education and availability of reading materials further reduces the estimates (specification C). Adding measures of the amount of time spent doing homework and reading has little effect on the estimates (specification D), as does including school fixed effects in the HSB and NELS data (specification E). Controlling for all these factors, the estimates indicate that an additional hour of television viewing is associated with a decline in test scores of about .003–.01 standard deviations, estimates that are statistically significant but miniscule.

The relationship between most other factors and test scores is considerably larger than the estimates for television viewing. In results not shown here, having an additional sibling is associated with a decline in test scores of about .02–.04 standard deviations, more than twice as large as the estimates for an additional hour of television per week.[18] Young adults who spend more time reading or doing homework have significantly higher test

[17] Chow tests indicate that the coefficients are jointly statistically different between males and females and between Whites and Blacks, so separate regressions are estimated instead of merely including interaction terms between television viewing and sex and race/ethnicity in a single regression. Whites and non-Black, non-Hispanic others are pooled as a racial/ethnic group because the NLSY does not clearly distinguish between these groups.

[18] An appendix table that presents the full results of specification *D* is available from the author on request. A discussion of the robustness of the results to including a quadratic measure of hours of television watched or a spline at 10 h; to dropping individuals who report not watching television is also available on request; and to examining scores on specific tests instead of on composite results. In general, the results shown here are robust to those changes.

Table 2
Relationship between hours of television watched and test scores, cross-sectional estimates

	All	Male	Female	Black	Hispanic	White/Other
NLSY: AFQT score						
(A) TV only	−.010**	−.007**	−.012**	.002	−.004	−.014**
	(.001)	(.002)	(.002)	(.002)	(.003)	(.002)
(B) A + individual	−.007**	−.007**	−.008**	.002	−.005	−.014**
characteristics	(.001)	(.002)	(.002)	(.002)	(.003)	(.002)
(C) B + family background	−.003**	−.002	−.005**	.002	−.002	−.007**
	(.001)	(.002)	(.001)	(.002)	(.003)	(.002)
(D) C + other uses of time	−.003*	−.002	−.004**	.003†	−.001	−.006**
	(.001)	(.002)	(.002)	(.002)	(.003)	(.002)
Number of observations	2477	1309	1168	719	415	1343
HSB: Composite test score						
(A) TV only	−.019**	−.018**	−.020**	−.006*	−.008**	−.017**
	(.001)	(.001)	(.001)	(.002)	(.002)	(.001)
(B) A + individual	−.015**	−.015**	−.016**	−.006**	−.009**	−.017**
characteristics	(.001)	(.001)	(.001)	(.002)	(.002)	(.001)
(C) B + family background	−.011**	−.011**	−.012**	−.005**	−.006**	−.012**
	(.001)	(.001)	(.001)	(.002)	(.002)	(.001)
(D) C + other uses of time	−.009*	−.009**	−.009**	−.004**	−.006**	−.010**
	(.001)	(.001)	(.001)	(.002)	(.002)	(.001)
(E) D + school fixed effects	−.008*	−.007**	−.008**	−.003*	−.003†	−.009**
	(.001)	(.001)	(.001)	(.002)	(.001)	(.001)
Number of observations	29,976	13,548	16,428	2792	4134	23,050
NELS: Composite test score						
(A) TV only	−.019**	−.017**	−.022**	−.007†	−.010*	−.018**
	(.001)	(.002)	(.002)	(.004)	(.005)	(.002)
(B) A + individual	−.016**	−.013**	−.019**	−.007†	−.012*	−.018**
characteristics	(.001)	(.002)	(.002)	(.004)	(.005)	(.002)
(C) B + family background	−.011**	−.009**	−.013**	−.006†	−.008†	−.012**
	(.001)	(.002)	(.002)	(.004)	(.005)	(.001)
(D) C + other uses of time	−.010*	−.008**	−.012**	−.005	−.006	−.011**
	(.001)	(.002)	(.001)	(.004)	(.004)	(.001)
(E) D + school fixed effects	−.009*	−.006**	−.011**	−.002	.000	−.008**
	(.001)	(.002)	(.001)	(.004)	(.003)	(.001)
Number of observations	18,765	8871	9894	1275	1707	15,783

**$p<.01$; *$p<.05$; †$p<.10$.

Notes: The dependent variable is normalized test scores. Shown are estimated coefficients of the number of hours of television watched per week (weekday hours times seven for the HSB sample and weekday hours times five plus weekend day hours times two for the NELS sample). Regressions include a constant and other variables as indicated (see text and Appendix Table 1 for details); HSB and NELS regressions also include year fixed effects for the survey waves. Observations are weighted using survey weights, and standard errors (in parentheses) are White-corrected for heteroscedasticity and clustered on the individual. "White/Other" is non-black, non-Hispanic individuals.

scores. In the NLSY results, for example, young adults who spend an additional hour reading per week have test scores about .01 standard deviations above the mean while young adults who do an additional hour of homework have test scores .02 standard deviations higher, compared with .003 standard deviations lower for young adults who watch an additional hour of television per week.

The relationship between television viewing and test scores varies somewhat across sexes and across racial/ethnic groups. The estimates tend to be smaller in magnitude for men than for women, particularly in the NELS data. The estimates are also smaller for Blacks and Hispanics than for Whites and other groups; most of the results fail to indicate a significant negative relationship for Blacks and Hispanics. Interestingly, at all ages, males and minorities tend to watch more television than females and Whites (Searls, Mead, & Ward, 1985). The results are consistent with the hypothesis

Table 3
Relationship between hours of television watched and test scores, individual and family fixed effects estimates

	Individual HSB	Individual NELS	Siblings NLSY	Twins HSB
TV	.001	.0001	−.002	−.001
	(.001)	(.0008)	(.005)	(.005)
Individual fixed effects	Yes	Yes	No	No
Family fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	No	No
Number of observations	29,976	18,765	490	865
Correlation in television viewing	.431	.527	.200	.445
Correlation in test scores	.871	.884	.571	.695

$^{**}p<.01$; $^{*}p<.05$; $^{\dagger}p<.10$.
Notes: The dependent variable is normalized test scores. Shown are estimated coefficients of the number of hours of television watched per week (weekday hours times seven for the HSB samples) and simple correlations among siblings at a point in time or within individuals over time. All regressions include fixed effects as indicated and a constant. Observations are weighted using survey weights, and standard errors (in parentheses) are White-corrected for heteroscedasticity and clustered on the individual in columns 1–2 and on the family in columns 3–4.

that groups that experience more detrimental effects of television on human capital watch less of it, on average.

4.2. Fixed effects estimates

Whereas most of the cross-sectional results indicate a negative relationship between test scores and television scores, the fixed effects estimates suggest that television viewing does not affect test scores. As indicated by the simple correlations reported in the last two rows of Table 3, television viewing is highly correlated within individuals over time and among siblings at a point in time, as are test scores. However, the regression results in the first row of Table 3 indicate that changes in individuals' test scores over time are not related to changes in television viewing in the HSB and the NELS. Similarly, differences in test scores across siblings in the NLSY and across twins in the HSB are not significantly related to differences in television viewing.

There are several possible interpretations of the insignificant relationship between television viewing and test scores in the individual and family fixed effect estimates. The difference between the cross-sectional and fixed effects estimates may indicate that factors that are time-invariant within individuals or common across siblings and that are not controlled for in the cross-sectional regressions underlie the negative estimates in the cross-sectional results. If so, then the fixed effects estimates correctly indicate that there is not a causal relationship between hours of television watched and test scores.

Alternatively, measurement error may underlie much of the changes in individuals' responses and the differences between siblings' reports of hours of television watched.[19] For example, although about one-third of twins in the HSB report the same interval for the number of hours usually watched on a weekday, almost another one-third indicate that they watch at least 2 h more (or less) per weekday than their twin. In the NELS, about one-quarter of respondents report changes in hours of television watched per week in excess of 10 h—a substantial change relative to the sample mean of less than 19 h per week—across survey waves. However, excluding siblings sets or individuals with large differences (in excess of 10 h per week) yields estimates similar to those shown in Table 3, suggesting that measurement error does not underlie the difference between the cross-sectional and fixed effects results.

Another possible reconciliation between the cross-sectional and fixed effects estimates is that the negative effect on test scores occurs as a result of watching television at younger ages. If young adults who are heavy television viewers were also heavy television viewers as children, then the cross-sectional estimates might reflect a detrimental effect of heavy television viewing at younger ages that

[19]Self-reports of time spent watching television are positively correlated with time diary measures (Price, Ritchie, Roberts, & Lieberman, 1986), but the extent of measurement error in self-reports is unknown.

carries over to test scores during later adolescence and adulthood. If the effect occurs at younger ages and persists, then changes in television viewing at older ages might have little impact on test scores.

It is possible that levels of television viewing are related to changes in test scores even though changes in television viewing are not. In other words, individuals who watch more television when younger may experience less improvement in test scores over time. However, if the change in test scores is regressed on weekly television viewing in the base year of the HSB and the NELS, the results are either insignificant or positive (not shown). Regressing the change in test scores in the NELS on the amount of television viewing reported in the previous wave also yields a positive result, even when controlling for other factors. These results provide no evidence that eighth graders or high school sophomores who watch more television experience declines in test scores over time relative to their peers who watch less television.

5. Conclusion

Everything from the decline in hours worked among US males (Juster & Stafford, 1991) to the decline in average SAT scores (Wirtz, 1977) has been attributed, in part, to a sectoral increase in hours spent watching television. This study's examination of the relationship between television viewing and test scores suggests that such concerns may be misplaced. In cross-sectional data on young adults from the 1980s and early 1990s, test scores on vocabulary, reading and math tests are negatively associated with contemporaneous reports of hours of television watched. However, the magnitude of the estimates is small, particularly in comparison with the effects of time spent reading and time spent doing homework. The estimated effect of television becomes smaller when controlling for other factors, such as parental education. Further, comparisons across siblings, including twins, at a point in time and within individuals over time do not indicate a negative relationship between test scores and television viewing. Changes in time spent watching television as a young adult therefore do not appear to adversely affect test scores. Instead, endogeneity likely underlies the negative relationship between test scores and television viewing reported in several previous studies.

Acknowledgements

The author thanks two anonymous referees and Mike Hilmer for helpful comments. Work on this study began while the author was affiliated with Occidental College. Any views expressed here are solely those of the author.

References

Aksoy, T., & Link, C. R. (2000). A panel analysis of student mathematics achievement in the US in the 1990s. *Economics of Education Review*, *19*(3), 261–277.

Gaddy, G. D. (1986). Television's impact on high school achievement. *Public Opinion Quarterly*, *50*(3), 340–359.

Gortmaker, S. L., Salter, C. A., Walker, D. K., & Dietz, W. H. (1990). The impact of television viewing on mental aptitude and achievement. *Public Opinion Quarterly*, *54*(4), 594–604.

Hornik, R. C. (1981). Out-of-school television and schooling. *Review of Educational Research*, *52*, 193–214.

Juster, F. T., & Stafford, F. P. (1991). The allocation of time: Empirical findings, behavioral models, and problems of measurement. *Journal of Economic Literature*, *29*(2), 471–522.

Keith, T. Z., Reimers, T. M., Fehrmann, P. G., Pottebaum, S. M., & Aubey, L. W. (1986). Parental involvement, homework, and TV time: Direct and indirect effects on high school achievement. *Journal of Educational Psychology*, *78*(5), 373–380.

Maume, D. J., Jr., Cancio, A. S., & Evans, T. D. (1996). Cognitive skills and racial wage inequality: Reply to Farkas and Vicknair. *American Sociological Review*, *61*(4), 561–564.

Morgan, M., & Gross, L. (1980). Television viewing, IQ and academic achievement. *Journal of Broadcasting*, *24*(2), 117–133.

Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in Black–White wage differences. *Journal of Political Economy*, *105*(5), 869–895.

Price, V. L., Ritchie, D., Roberts, D. F., & Lieberman, D. (1986). *The Stanford reading and television study: A progress report*. Stanford, CA: Institute for Communication Research.

Robinson, J. P., & Godbey, G. (1997). *Time for life: The surprising ways Americans use their time* (2nd ed.). University Park, PA: Pennsylvania State University Press.

Rodgers, W. M., III, & Spriggs, W. E. (1996). What does the AFQT really measure. *Review of Black Political Economy*, *24*(4), 13–46.

Searls, D. T., Mead, N. A., & Ward, B. (1985). The relationship of students' reading skills to TV watching, leisure time reading, and homework. *Journal of Reading*, *29*, 158–162.

Wirtz, W. (1977). *On further examination; Report of the advisory panel on the Scholastic Aptitude Test score decline*. New York: College Entrance Examination Board.