Learning and Instruction 47 (2017) 1–12

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Income differences among grade skippers and non-grade skippers across genders in the Terman sample, 1936–1976

Russell T. Warne*, Jonathan K Liu [1]

Department of Behavioral Science, Utah Valley University, United States

ARTICLE INFO

Article history:
Received 21 May 2016
Received in revised form
27 August 2016
Accepted 14 October 2016
Available online 24 October 2016

Keywords:
Grade skipping
Full-grade acceleration
Gifted education
Longitudinal studies
Income

ABSTRACT

Full-grade acceleration, also called *grade skipping*, is a widely supported practice among gifted education experts. Yet, the impacts of grade skipping in adulthood are unclear. Using data from Terman's longitudinal study of gifted children, we examined income differences from 1936 to 1976 between grade skippers and non-grade skippers after controlling for birth year, IQ, home environment, personality, and intellectual, social, and activity interests via propensity score modeling. After also controlling for adult education attainment, men who had skipped a grade earned an average of 3.63%–9.35% more annually than non-grade skipping men. The impact for grade skipping women was much smaller: −2.02%–0.42% annually. These results indicate no association between full-grade acceleration and income for women in this historic dataset, but suggest a slight relationship between the two variables for men (though whether this relationship is causal is unknown). Additionally, income gaps between accelerated and non-accelerated students did not narrow until the subjects were nearing the end of their careers. We discuss these findings in the context of gifted education policy and other research on academic acceleration.

One of the oldest academic interventions for gifted children is full-grade acceleration, which entails permitting a child to skip a grade in order to attend a grade one year earlier than their age peers. Leaders from gifted education's past (e.g., Hollingworth, 1926, 1942; Stanley, 1976; Terman, 1954; Terman & Oden, 1947) recognized the potential benefits of full-grade acceleration. These early opinions are still mainstream among gifted education experts, who often find that accelerated gifted children outperform their non-accelerated age peers on academic, social, and self-esteem measures (Assouline, Colangelo, VanTassel-Baska, & Lupkowski-Shoplik, 2015; Rogers, 2007).

In the 21st century research on full-grade acceleration continues. Recently researchers studying full-grade acceleration have found that accelerated gifted children outperform their (older) classmates on nearly every academic outcome, including high school and college grades, standardized tests, and advanced degree attainment (Cronbach, 1996; McClarty, 2015a; Park, Lubinski, & Benbow, 2013). These academic benefits usually do not come at a cost to social or emotional development (Gagné & Gagnier, 2004; Lee, Olszewski-Kubilius, & Peternel, 2010; Rogers, 2007). The only exception to these results that we were able to find was a Dutch study in which accelerated students' (older) peers rated them less positively as the students who had not been accelerated—especially if the accelerated students were male. However, accelerated students in this study had higher academic self-concepts than their older classmates (Hoogeveen, van Hell, & Verhoeven, 2009).

Despite the long history of interest in full-grade acceleration among gifted education researchers, few studies have examined long-term adult outcomes of children who skipped a grade. The limited research is mostly focused on academic outcomes (usually in college), social outcomes, and emotional outcomes of full-grade acceleration (e.g., Cronbach, 1996; McClarty, 2015a, 2015b; Park et al., 2013). Although this research is useful, there has been almost no research on financial outcomes of full-grade acceleration. The few researchers who have investigated economic outcomes (i.e., Cronbach, 1996; McClarty, 2015b) have not reported effect sizes, a violation of reporting standards that reduces the usefulness of their research (American Educational Research Association, 2006; American Psychological Association, 2010).

Therefore, teachers, administrators, parents, and advocates of gifted children have little information about the economic consequences of full-grade acceleration—a gap we hope to fill. Given the

* Corresponding author. Department of Behavioral Science, Utah Valley University, 800 W. University Parkway MC 115, Orem, UT 84058 United States.
E-mail address: rwarne@uvu.edu (R.T. Warne).
[1] Jonathan K Liu is now affiliated with the Department of Counseling Psychology & Special Education, Brigham Young University.

http://dx.doi.org/10.1016/j.learninstruc.2016.10.004
0959-4752/

strong support for acceleration among gifted education experts, new information about financial outcomes of acceleration may inform advocacy and scholarly work related to acceleration. Investigating the potential impact of acceleration on incomes is important because higher incomes are associated with a wide array of positive life outcomes (e.g., good health, longevity). Additionally, income is a nearly universal consequence of employment that permits comparisons across individuals. Other measures of career success are often less applicable to a wide population of adults. For example, Lubinski, Benbow, and Kell (2014) measured career success by counting the number of patents, peer review publications, or career awards that subjects earned in their longitudinal study. Although these may be good measures of success within some careers, these variables do not apply to many occupations.

1. Why might grade skippers have higher incomes?

For some readers, the connection between full-grade acceleration and adult income seems unclear. However, there are reasons why one would expect people who skip a grade to earn more income. First, in other samples students who finish high school earlier were more likely to earn a graduate degree in adulthood (e.g., Wai, 2015), which is correlated with higher incomes in adulthood. The causal relationship among these variables is not clear. It is possible that the characteristics that make a child experience acceleration are also the traits that make people more likely to pursue advanced education. Or possibly the greater academic challenge in childhood fosters an interest in learning and education that persists into adulthood (a possibility raised by Lubinski, Webb, Morelock, & Benbow, 2001). Regardless of the causal mechanisms at work, it would not be surprising if grade skippers later were more likely to obtain high levels of education, which then led to greater incomes.

Another possible explanation for the connection between acceleration in childhood and adult income is that “time is money.” For most people, acquiring expertise in a field requires learning new knowledge and developing new skills (Ericsson, Roring, & Nandagopal, 2007). It is likely that becoming an expert in many fields—especially a highly paid expert—takes time. By embarking on higher education and their careers earlier, grade skippers may earn higher incomes simply because they are further along in their careers and have developed their skills more fully. This extra time may also help them build a professional network or obtain the human capital needed to receive a high paying job.

Notwithstanding the theoretically plausible relationship between full-grade acceleration and adult income, it is important to recognize that other variables have relationships with adult income. One well-known correlate with income is gender, with men earning higher incomes than women both in the general population (Blau & Kahn, 2007) and in high ability populations (Lubinski et al., 2014). Another well established predictor of income is educational attainment, with better educated individuals generally earning higher incomes (Herrnstein & Murray, 1996; Nyborg & Jensen, 2001). Similarly, students with higher academic achievement tend to grow up to earn higher incomes (Strenze, 2007).

Some psychological traits are also positively correlated with income in adulthood. Motivation (Long, 1995; Lubinski et al., 2014) and intelligence (Jensen, 1998; Strenze, 2007; Warne, 2016) are well known examples of psychological variables with robust positive correlations with adult income. Lesser known is that among the “Big Five” personality traits, openness and conscientiousness are positively correlated with income, while neuroticism correlates negatively with income (O'Connell & Sheikh, 2011). It is possible that some of these variables are correlated with grade skipping. Therefore, any researchers who conduct a nonexperimental study on the economic impacts of grade skipping must attempt to control for these variables and thereby reduce the degree to which they could confound the results.

2. Research on adult income: two critical prior studies

Research on these issues is still in its infancy. Indeed, there have only been two prior studies in which researchers investigated the economic impact of grade skipping (Cronbach, 1996; McClarty, 2015b). Both of these studies produced results showing that children who experienced full-grade acceleration earned higher incomes as adults. However, both studies have shortcomings that make further research on the issue necessary, and it is not entirely clear that there even is a link between grade skipping and adult income. In this section of the article we will discuss these two studies and explain the need for our research.

Cronbach (1996)—using data from Terman's (1926) longitudinal study—first compared the adult income of gifted men who skipped at least one grade with a matched group of non-accelerated men. He found that income was higher in the accelerated group, but only among sample members without an advanced degree. Among sample members with an advanced degree, there was no difference between incomes in the two groups.

Cronbach's (1996) study was the first study on the adult incomes of grade skippers, but it has shortcomings. First, Cronbach did not report an effect size or any other statistic that would indicate the magnitude of the income differences in Terman's sample. Therefore, it is not clear how much of a financial advantage accelerated students could gain in their adult years. Second, Cronbach only matched the grade skippers and the non-grade skippers in the Terman sample on a limited number of variables: high school graduation year, final adult education status (both through weighting the two groups until they were equivalent), and gender (by only analyzing data from male subjects).

The characteristics of McClarty's (2015b) study are similar to those of Cronbach's (1996) study. Using data from the NELS:88 sample, she compared grade skippers with similar non-grade skippers and found that accelerated students held more prestigious jobs and higher incomes, but their job satisfaction did not differ. McClarty did provide annual income differences between the two groups (ranging from $920 approximately five years after high school graduation to $5112 approximately eight years after graduation). However, no standard deviations were reported, which makes calculating an effect size impossible. Like Cronbach, McClarty (2015b) also controlled for a small number of confounding covariates—gender, race, socioeconomic status, and eighth grade achievement—though she controlled for these covariates through the more sophisticated Coursened Exact Matching (CEM) method (see Iacus, King, & Porro, 2011).

Although neither Cronbach (1996) nor McClarty (2015b) made methodological errors when attempting to control for pre-existing group differences, both CEM and Cronbach's weighting methods have been surpassed by other methods of simulating the causal impact of a treatment in a non-experimental setting, namely propensity score modeling (Guo & Fraser, 2010). Propensity score modeling is an improvement over weighting and CEM because propensity score modeling permits researchers to control for a much larger number of variables than these methods (e.g., Warne, Larsen, Anderson, & Odasso, 2015). We designed this study to build upon the previous efforts of Cronbach and McClarty to examine the long-term economic impacts of full-grade acceleration. Specifically, the study is designed to answer three research questions:

1. After controlling for childhood covariates and adult education level, what is the size of the income gap between full-grade

accelerated gifted students and non-accelerated gifted students in their adult years?
2. If there is a difference between accelerated and non-accelerated gifted students' adult incomes, is the relative size of this difference stable, or does its size vary in adulthood?
3. Is there a difference between the financial impact of full-grade acceleration for men and women?

Cronbach (1996) and McClarty (2015b) addressed Questions 1 and 2 in their studies. Question 3 will be newly addressed in our work. Based on these prior studies, our hypothesis was that accelerated individuals would have higher incomes in adulthood, though we did not postulate an effect size quantifying the size of this gap. Additionally, we did not have an *a priori* hypothesis about the stability of the income gap or whether there would be gender differences in the size of the income difference between accelerated and non-accelerated students. Thus, this study was mostly exploratory in nature.

3. Methods

3.1. Data source and variables

To answer our research questions, we decided to re-analyze the data from Terman's longitudinal study of gifted children (Terman et al., 1922-1991). Supplemental File 1 has a full description of the data and includes how the variables were prepared for analysis. This section of the article gives a brief description of that process. Although Terman's subjects lived long ago (their average birth year was 1910), we chose to use the Terman sample for two reasons. First, this study is the only American dataset that follows a large number of grade skippers from childhood through much of their lives (in this case from 1921 through 1999, with income data available from 1936 through 1976). Second, full-grade acceleration was much more common when the study's sample members were children than in later decades (Terman, 1954; Warren, Hoffman, & Andrew, 2014), and we were confident that Terman's data would contain a large proportion of grade skippers.

After downloading the data, we selected covariates that prior research had shown were associated with the probability that a student would be accelerated a full grade (Cronbach, 1996; McClarty, 2015b; Wells, Lohman, & Marron, 2009). We only selected covariates that were measured in the first wave of data collection in 1922 and were available for both male and female subjects. These variables that were selected are displayed in Table 1. We were able to find variables in the first wave of data collection that could serve as proxies for academic aptitude, personality traits, and socioeconomic status—all of which could influence whether a child was grade skipped or were associated with adult income levels, according to previous research (i.e., Cronbach, 1996; McClarty, 2015b; Wells et al., 2009). We also included childhood interests in our propensity score model because it was plausible that these ratings—which related to a subject's interest in social life, intellectual pursuits, and activities—would be correlated with the probability that a child would be accelerated a grade. Descriptions of all of these variables are available in Supplemental File 1.

Grade skippers in the Terman data were identified through the birthdate method in which students who were younger than the typical age at high school graduation are labeled as having experienced full-grade acceleration during their K-12 career. This method has been used in other studies to identify accelerated students (e.g., Cronbach, 1996; McClarty, 2015b; Wells et al., 2009). Application of the birthdate method is particularly conservative for the Terman sample, and only subjects who had graduated from high school before their 17th birthday were identified as grade skippers. For details, see Supplemental File 1.

Table 1 also shows the income variables used in the analysis. Terman and his successors collected income data for 22 different years between 1936 and 1976, of which 15 are publicly available. All of the income variables consisted of the individual's earned income for the year; the income of the person's spouse and unearned income (e.g., pensions, social security payments royalties) were not included. See Supplemental File 1 for more information.

Finally, we believed it was important to also control for adult education attainment because some prior research had shown that accelerated students were more likely to earn an advanced degree (Cronbach, 1996; Park et al., 2013; Wai, 2015). This difference in educational attainment can make a simple mean income comparison between accelerated and non-accelerated subjects deceptive because any differences may be due to differing education levels in adulthood, rather than their different K-12 experiences. Therefore, we also used the subjects' year that they attained a bachelor's, master's, or doctorate degree in our statistical analysis.

3.2. Sample description

The Terman sample consisted of 1528 individuals (856 males and 672 females) who were identified as gifted because of their performance on cognitive or intelligence tests. Terman and his colleagues searched extensively for individuals who earned scores equivalent to an IQ of 135 or higher (see Burks, Jensen, & Terman,

Table 1
Variables included in the analysis.

Childhood covariates controlled for in the propensity score Analysis	
Variable(s) code	Variable description
B050	Birth year
B044 or B004	IQ score
B22768	Home environment rating (proxy for socioeconomic status)
B020	Emotional stability (i.e., personality)
B021	Intellectual interests
B022	Social interests
B023	Activity interests
B23118	Amount of reading compared to average child
Variable used to identify grade skippers	
ED018	Age at high school graduation
Adult education variables	
ED007	Year of high school graduation
ED008	Year of bachelor's degree
ED009	Year of master's degree
ED010	Year of doctorate/professional degree
Adult income variables	
B36072 – B36075	1936 income
B40080	1940 income
B50012	1946 income
B50013	1947 income
B50014	1948 income
B50015	1949 income
B55003	1954 income
B60005	1956 income
B60006	1957 income
B60007	1958 income
B60008	1959 income
B72040	1960–1969 income (median of 1965 used in analysis)
B72038	1970 income
B72039	1971 income
B77128	1976 income

Note. For details concerning these variables, see Supplemental File 1.

1930; Terman, 1926, for selection details). Income data were collected at various times from 1936 through 1976, and 782 men (91.4%) and 532 women (79.2%) reported income in at least once during that time period.

The men were in a wide variety of careers spanning almost the entire social scale, though from 1940 through the 1970's about half of men were classified as professionals. During the same time period, men working in business gradually increased (from 31.0% to 43.6%), while those in clerical or sales jobs decreased (from 18.3% to 3.5%). Notwithstanding these changes, Holahan and Sears (1995, p. 48) noted that, "Overall, the group's career lines were fairly stable once established." For the women in the Terman study, the two most common income producing jobs for women were teaching and secretarial work. It is important to note, though, that the Terman women's work history was often characterized by repeated entry into and exit from the work force (Holahan & Sears, 1995).

All sample members were born between 1900 and 1926 (mean birth year = 1910) and lived in California during part or all of their childhood. The subjects' childhood homes were privileged by the standards of the 1920's. Twenty percent of mothers and over one-quarter of fathers had a bachelor's degree, and the living fathers were more likely to be in a high status occupation than the general urban population of California at the time (Terman, 1926).

Classifying the subjects into modern racial groups is difficult. Yet, it is apparent that almost all of Terman's subjects were white. Approximately 80% of their fathers and 83% of their mothers were born in the United States (Holahan & Sears, 1995, p. 12). However, at least six individuals were multiracial (Shurkin, 1992; Terman, 1926). There were also small number of individuals in the study with Mexican, "Indian" (probably Native American), "Spanish" (which probably included non-Mexican Hispanics, in addition to Iberian groups), and Syrian heritage. In total, we estimate that there were between 20 and 50 individuals with partial or total non-European ancestry.

3.3. Missing data

With over 1500 subjects and over 4000 variables, there is quite a bit of missing data in the Terman study dataset. These missing data happened for a variety of reasons. For example, in 1922 some subjects who would later be added to the sample had not been found yet; Terman would continue to add subjects to his study until 1928 (see Burks et al., 1930; Oden, 1968). Other individuals were missing data files, which would occur if a subject or an informant (e.g., a parent, teacher, spouse) did not respond to requests for data.

Subject attrition was also a factor in missing data: By 1977 (the year that income data were collected for the last time) the participation rate had dropped to 62.8% of living subjects and 46.0% of the original sample (Holahan & Sears, 1995, pp. 35-36). In Terman's data subject attrition was the result of several processes, including death, incapacitation, explicitly withdrawing from the study, or quietly failing to return surveys (Cronbach, 1995). Modern methods can be adept at compensating for missingness (Enders, 2001). However, using these methods in the Terman dataset is problematic, mostly because the patterns of missingness and the relationship between that missingness and the observed variables will not be the same across different types of attriters. These obstacles are explained in more detail in Supplemental File 1.

Therefore, we used two simpler methods of handling missing data. First, when data were missing from the childhood covariates (see Table 1), we used mean imputation to replace their missing data. Although this reduces the variability of the data and may produce biased parameter estimates (Brown, 1994), we thought it was better than alternate options because it did not reduce the sample size. For missing income data, we used pairwise deletion in the analyses, a common procedure for analyzing the Terman longitudinal data (e.g., Holahan & Sears, 1995).

3.4. Statistical models

We analyzed the data for men and women in the Terman sample separately because the career trajectories of the two sex groups were very different (Holahan & Sears, 1995). Additionally, analyzing the data from males and females separately has been a long-standing procedure with the Terman data from the earliest days of the study, and maintaining this practice makes our research comparable with prior studies using this sample (e.g., Cronbach, 1996; Holahan & Sears, 1995; Oden, 1968; Terman & Oden, 1959).

3.4.1. Propensity score model

To match the subjects on the childhood covariates, we inputted the childhood covariates into a logistic regression model where the dependent variable was the subjects' grade skipping status (0 = non-skipper, 1 = skipper). The predicted probability that each subject would have accelerated a grade was calculated, and this was the propensity score. A small number of subjects (17 men and 18 women) were eliminated from further analysis because their propensity scores were outside of the range of common support, and there were no counterfactual subjects available for them (Fan & Nowell, 2011; Guo & Fraser, 2010). The remaining individuals were ordered by their propensity scores (from highest to lowest) and then divided the sample into five strata equally sized strata on the basis of their propensity scores (in accordance with recommendations from Guo & Fraser, 2010; Fan & Nowell, 2011). As a result, the top stratum consisted of individuals with the highest 20% of propensity scores, the second stratum consisted of individuals with the next highest 20% of propensity scores, and so on until the bottom stratum consisted of individuals with the lowest 20% of propensity scores. After subjects were eliminated for missing too much data or for not being useful in creating the propensity score model, there were 810 identified accelerated subjects (432 males and 378 females) and 651 non-accelerated subjects (386 males and 265 females).

3.4.2. Path analysis model

Another potentially confounding variable that needed controlling was adult educational attainment. However, this variable cannot be included in the propensity score model because covariates used to calculate propensity score should be variables that are present during or before the time subjects are assigned to groups (Guo & Fraser, 2010). To eliminate this confounding we performed a path analysis that separates the direct impact of acceleration on adult income levels from the indirect impact that acceleration may have on income via education level (as measured in years of postsecondary education) as a mediator variable. This model is shown in Fig. 1.

We used MPlus to perform the path analysis within each stratum. The three unstandardized paths in each stratum's model were analyzed separately. See Supplemental File 1 for details. The direct path parameter estimate, calculated across all strata within each year, provided an estimate for the annual income difference (expressed in dollars) between accelerated and non-accelerated groups after matching for all the childhood covariates and controlling for adult educational attainment. These values were then converted into both a percentage and an effect size to indicate the difference in income between grade skippers and non-grade skippers; these conversions standardized the results to control for the changing value of the dollar across the years of the study. For both measures, positive numbers indicated a higher income for the accelerated subjects.

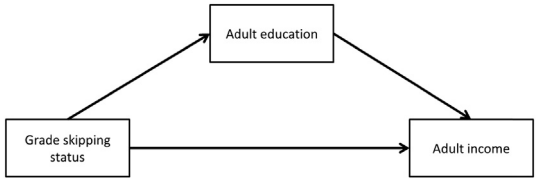

Fig. 1. Path analysis model for examining the direct and indirect effects of full-grade acceleration on adult income.

3.4.3. Alternative analysis

Because the longitudinal nature of this study violates some of the assumptions of meta-analytic methods, some readers may object to our analysis. To complement this study, we have provided an alternate analysis of the Terman data in Supplemental File 2. The supplemental file details a set of longitudinal hierarchical linear models that we created to investigate the typical income trajectory of accelerated and non-accelerated subjects. Both analyses use the same set of childhood covariates and the same income data in adulthood, though they rely on different assumptions about data dependence and missing data.

4. Results

Table 2 shows descriptive statistics for the sample as a whole and for the two groups. Our initial propensity score models did not successfully balance the covariates within each stratum. For men, the birth year ($\eta^2 = 0.0156$, $p = 0.001$), IQ ($\eta^2 = 0.0027$, $p = 0.025$), and social interest variables ($\eta^2 = 0.0133$, $p = 0.024$) were unbalanced, meaning that these covariates were not fully controlled in the propensity score creation process. In an effort to balance these covariates and create a propensity score model that better estimated the economic impact of acceleration, we created polynomial terms for these covariates by mean centering them and then squaring the covariates to create a new covariate and including these polynomial terms to the propensity score model. This resulted in balancing the IQ covariate, though a slight imbalance remained for the birth year ($\eta^2 = 0.018$, $p = 0.001$) and social interest variables ($\eta^2 = 0.014$, $p = 0.017$). We judged these differences to be small enough that we could proceed with the analysis. For women, only the birth year covariate was unbalanced in the original propensity score model. For the sake of comparability with the men's data, we created the same three polynomial terms and added them to a new propensity score model. This succeeded in balancing all of the covariates in propensity score model for the female subjects.

4.1. Results for men

Across the 15 time points in which adult income were available, the unweighted mean difference in incomes was 9.35%, with men who had experienced full-grade acceleration earning more than their non-accelerated peers, after controlling for childhood variables and adult educational attainment. This corresponds to an unweighted mean Cohen's *d* value of 0.102 for the effect sizes derived from these 15 data points. These values are shown in Table 3, which also provide the unstandardized path estimates for the other two paths in the model (which combine to provide an indirect path from grade skipping status to adult income mediated via adult education attainment).

Fig. 2 shows how the income differences between the two groups change over the subjects' adult years. A polynomial trendline (calculated as the ordinary least squares polynomial regression line of best fit for the unweighted data points) in the figure has been added for ease of interpretation and allows for the possibility of changing score gaps across time. This trendline shows that the income advantage that grade skippers have over non-grade skippers was maintained until the 1970's—when the average subject was in his early 60's. In that decade the group averages were smaller than at any other time in the subjects' adult years. Yet, the income gap between the two groups never closed completely.

4.2. Results for women

Table 3 also shows the income difference between women who had experienced full-grade acceleration and those who had not (after controlling for childhood variables and adult education attainment). On average, women who experienced full-grade acceleration earned incomes that were 0.42% higher than those who graduated from high school with their age peers. This group difference corresponds to an unweighted average Cohen's *d* of 0.002. Table 3 also shows the unstandardized path estimates for the indirect path between grade skipping and adult income mediated via adult education attainment.

The income differences between grade skipping and non-grade skipping women is shown in Fig. 3, which—like Fig. 2—includes a polynomial trendline (calculated as the ordinary least squares polynomial regression line of best fit for the unweighted data points) that aids in interpretation. The trendline in Fig. 3 shows that the income differences, though slight, are consistent through most of the adult years, with the only exceptions being in 1936 and 1947–1949, where women who had experienced full-grade acceleration earned less than those who had not experienced full-grade acceleration. Additionally, there is a noticeable negative trend towards the end of the study, with the last four observed income differences showing an advantage for non-accelerated women.

4.3. Comparing Men's and Women's results

Comparing the information in Table 3 and in Figs. 2 and 3 is enlightening. The most noticeable difference is that the income differences between accelerated and non-accelerated students were more pronounced in men than in women. Whether measured as a standardized effect size or a percentage difference in income, the men in the Terman sample earned higher incomes than non-accelerated men, but the accelerated women did not earn more money than their non-accelerated counterparts. Indeed, Table 3 shows that the relationship among all of the variables in the path analysis were weaker for women than for men.

Yet, the trendlines in the two figures are similar and show that the income differences are stable through most of the Terman sample's adult years, with declines not coming until the mid-1960's. The only difference between the trendlines' general shape is that women who had skipped a grade earned lower incomes in the early years of their careers. But, like the men's trendline, the trendline for the women took a downward trajectory in the 1960's. This downward slope towards the end of the subjects' careers may indicate that the relationship between full-grade acceleration and adult income may not persist throughout adults' working years.

5. Discussion

5.1. Research questions

We designed this research study to use the Terman data to answer three research questions:

1 After controlling for childhood covariates and adult education level, what is the size of the income gap between full-grade

Table 2
Descriptive statistics for variables included in the study.

Variable	Men: M (SD), *n*		Women: M (SD), *n*	
	Grade skippers	Non-skippers	Grade skippers	Non-skippers
Birth year	1910.7 (3.20), 432	1909.5 (4.15), 386	1911.2 (3.11), 378	1910.2 (3.97), 265
IQ	149.2 (10.27), 432	143.5 (6.83), 386	147.6 (9.88), 378	144.1 (8.44), 265
Home environment rating	4.6 (0.40), 432	4.6 (0.41), 386	4.6 (0.41), 378	4.6 (0.34), 265
Emotional stability	2.3 (0.36), 432	2.3 (0.32), 386	2.3 (0.39), 378	2.3 (0.39), 265
Intellectual interests	4.0 (0.47), 432	4.0 (0.41), 386	4.0 (0.47), 378	4.0 (0.43), 265
Social interests	3.8 (0.51), 432	3.8 (0.46), 386	3.7 (0.48), 378	3.7 (0.39), 265
Activity interests	3.3 (0.49), 432	3.3 (0.43), 386	3.3 (0.58), 378	3.2 (0.58), 265
Amount of reading compared to average child	4.3 (0.66), 432	4.3 (0.64), 386	4.3 (0.61), 378	4.3 (0.63), 265
Age graduated from high school	16.0 (0.63), 432	17.5 (0.61), 386	16.0 (0.61), 378	17.3 (0.43), 265
Year of high school graduation	1927.0 (3.3), 432	1927.2 (4.0), 386	1927.6 (3.1), 378	1927.7 (3.9), 265
Year of bachelor's degree	1932.4 (9.5), 328	1933.2 (9.5), 234	1932.4 (4.4), 255	1932.7 (6.0), 170
Year of master's degree	1935.1 (5.4), 113	1935.0 (5.8), 73	1936.4 (7.4), 101	1936.3 (6.5), 59
Year of doctorate or professional degree	1937.9 (7.0), 148	1937.0 (7.7), 83	1942.3 (10.1), 16	1943.2 (11.7), 13
1936 income	1322.81 (1023.29), 285	1356.90 (972.37), 232	742.73 (481.81), 220	840.31 (628.80), 129
1940 income	2761.79 (1818.51), 301	2664.14 (1815.18), 251	1626.32 (723.01), 152	1621.28 (784.92), 94
1946 income	7584.37 (10,239.71), 339	7665.90 (10,234.61), 305	3117.74 (1980.02), 124	3095.95 (3158.44), 74
1947 income	8337.39 (8756.351), 345	8342.12 (11,324.16), 311	3162.90 (2097.95), 124	3332.10 (3929.15), 81
1948 income	8973.93 (8899.33), 349	8609.12 (10,075.98), 307	3499.20 (2100.06), 125	5201.28 (10,760.42), 78
1949 income	10,140.17 (17,315.17), 351	9066.99 (10,165.98), 312	3664.39 (2096.79), 132	4053.09 (4614.11), 81
1954 income	14,369.12 (13,862.29), 353	12,086.33 (12,164.06), 300	4210.76 (2776.51), 158	4593.75 (3805.89), 96
1956 income	15,316.29 (12,239.35), 313	14,295.28 (13,283.00), 254	4881.94 (3348.10), 144	4955.56 (3600.60), 90
1957 income	16,180 (13,272.82), 315	15,054.90 (13,500.25), 255	10,919.25 (22,674.87), 161	9601.94 (20,410.18), 103
1958 income	17,561.90 (14,490.89), 315	15,435.29 (14,490.89), 255	10,578.62 (21,695.65), 159	9857.14 (20,224.43), 105
1959 income	18,176.10 (14,198.03), 318	16,727.63 (15,701.18), 257	10,377.91 (20,908.94), 172	10,154.55 (19,777.29), 110
1965 income	24,098.96 (17,893.10), 192	21,132.08 (15,059.00), 159	7450.98 (4580.69), 102	7646.15 (5206.46), 65
1970 income	27,066.35 (19,956.03), 211	25,607.36 (16,734.837), 163	9511.45 (5986.35), 131	11,527.78 (11,428.86), 72
1971 income	28,316.83 (24,599.90), 202	26,503.14 (17,640.53), 159	9858.27 (5825.13), 127	10,402.78 (6583.50), 72
1976 income	30,117.50 (27,813.52), 160	29,982.69 (40,005.71), 104	10,912.50 (7556.88), 72	12,192.86 (13,318.27), 42

accelerated gifted students and non-accelerated gifted students in their adult years?

2 If there is a difference between accelerated and non-accelerated gifted students' adult incomes, is the relative size of this difference stable, or does its size vary in adulthood?

3 Is there a difference between the financial impact of full-grade acceleration for men and women?

In regards to the first research question, we found that full-grade acceleration in the K-12 years was associated with a higher income as an adult. Men who skipped a grade earned an income that was—on average—9.35% higher than men who did not ($d = 0.102$), while the female grade skippers earned an average of 0.42% more money per year than females who did not experience full-grade acceleration ($d = 0.002$). The men's results supported our *a priori* hypothesis about accelerated subjects having higher incomes in adulthood, but the women's results did not.

When these income differences were examined across 40 years of data, we found that the advantage that male grade skippers had over male non-grade skippers was stable until the average subject in the Terman study was 55 years old. Starting at that point and continuing for the final 11 years that the income data were collected, the gaps between income groups decreased. Extrapolating from the trendline, we speculate that the gaps between the two groups would close for men in 1983, when the average subject was 73 years old. For women, the trendline shows that the income gaps closed in 1968, when the average female subject was 58 years old. This 15-year difference pertains to our third research question on differences in men's and women's trends in income for grade skippers and non-skippers.

Another difference we found to be important was that the relationship between full-grade acceleration and income was far stronger for men than for women. Despite these differences in the results for men and for women, we found that the trendlines were somewhat similar across both sex groups. For both men and women, once the income differences formed, they were stable until the mid-1960's when the differences started to disappear. However, it is important to note in Figs. 2 and 3 that there is substantial variability or "noise" in the results, with a small number of aberrant data points that are far from the trendline. This shows that researchers who investigate long-term outcomes of acceleration should adopt a longitudinal design (as we did in our alternate analysis in Supplemental File 2) so that outlier data points can be balanced out by the more consistent results from several other years.

Although both sex groups show a narrowing in income gaps between grade skippers and non-grade skippers in late middle age or early maturity, it is unclear whether this actually indicates that non-accelerated students were "catching up" to their peers. One possibility is that this finding may be an artifact of how the subjects entered retirement. About half of the Terman sample were subject to mandatory retirement, with age 65 being the most common age of forced retirement (Holahan & Sears, 1995, pp. 70; 117). This would correspond to a mean mandatory retirement year of 1975, with subjects beginning to turn 65 ten years prior and continuing until after Terman's successors stopped collecting income data. Yet, even when forced into retirement, many of Terman's male subjects (and his female subjects who had careers) kept working part-time. Because these older subjects were also the most likely to have been accelerated (Cronbach, 1996; see also Supplemental File 1), the narrowing income between accelerated and non-accelerated individuals may merely indicate that older grade skippers worked less—and therefore had lower incomes—than younger workers.

5.2. Comparison with the alternate analysis

Readers who compare the results in this article with the results of our alternate analysis in Supplemental File 2 will notice similarities. First, in both sets of analyses the male accelerated subjects had higher incomes than non-accelerated subjects. In the alternate

Table 3
Path analysis results.

Male subjects					
Year	Grade skipping → Adult education path	Adult education → Adult income path	Grade skipping → Adult income path		
	Unstandardized path	Unstandardized path	Unstandardized path	% Difference in incomes	Cohen's *d* value
1936	1.06	-$8.09	$184.90	13.8%	0.185
1940	1.17	$68.03	$236.92	8.71%	0.130
1946	0.99	$-21.55	$627.14	8.19%	0.061
1947	1.01	$44.40	$535.16	6.38%	0.053
1948	1.03	$246.14	$587.18	6.63%	0.062
1949	1.06	$84.71	$2143.48	22.12%	0.148
1954	1.12	$394.08	$1855.14	13.88%	0.141
1956	0.99	$449.45	$1066.59	6.74%	0.079
1957	0.97	$496.68	$1069.82	6.79%	0.080
1958	1.02	$480.32	$2277.88	13.64%	0.158
1959	0.91	$543.79	$1322.39	7.51%	0.089
1965	1.18	$419.35	$2903.03	12.67%	0.173
1970	1.08	$507.24	$1213.06	4.55%	0.065
1971	1.11	$321.31	$1572.23	5.66%	0.078
1976	1.59	$1595.33	$932.46	3.02%	0.028
Mean				9.35%	0.102
Female subjects					
Year	Grade skipping → Adult education path	Adult education → Adult income path	Grade skipping → Adult income path		
	Unstandardized path	Unstandardized path	Unstandardized path	% Difference in incomes	Cohen's *d* value
1936	0.67	$24.15	-$60.17	−7.73%	-0.111
1940	0.04	$104.21	$103.53	6.38%	0.139
1946	0.18	$142.99	$97.52	3.14%	0.039
1947	0.14	$165.76	$-82.79	−2.56%	-0.028
1948	0.13	$173.19	-$642.79	−15.50%	-0.093
1949	0.10	$234.74	-$62.53	−1.65%	-0.020
1954	0.16	$326.72	$78.07	1.80%	0.024
1956	0.12	$336.87	$285.21	5.80%	0.083
1957	0.10	$-36.98	$1179.77	11.40%	0.054
1958	0.08	$10.29	$689.49	6.73%	0.033
1959	0.10	$-230.96	$1129.81	11.03%	0.055
1965	−0.06	$539.37	-$68.92	−0.92%	-0.014
1970	−0.05	$665.90	-$465.99	−4.58%	-0.056
1971	−0.18	$586.77	-$24.80	−0.25%	-0.004
1976	0.05	$246.03	-$773.05	−6.78%	-0.077
Mean				0.42%	0.002

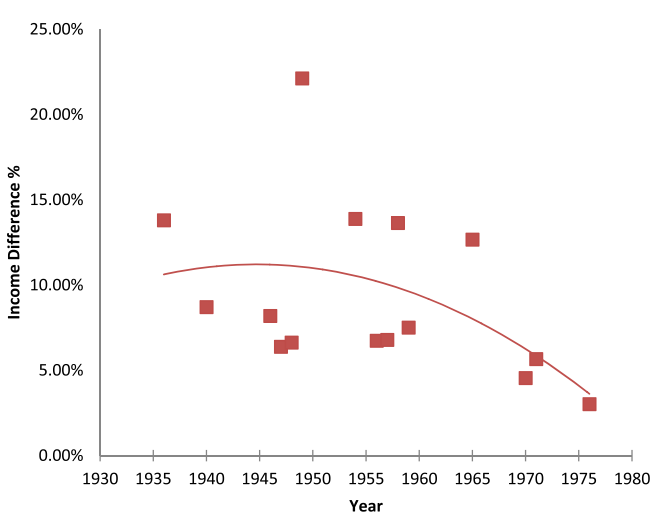

Fig. 2. Men's income differences between grade skippers and non-skippers over 40 years. Polynomial trendline added for ease of interpretation.

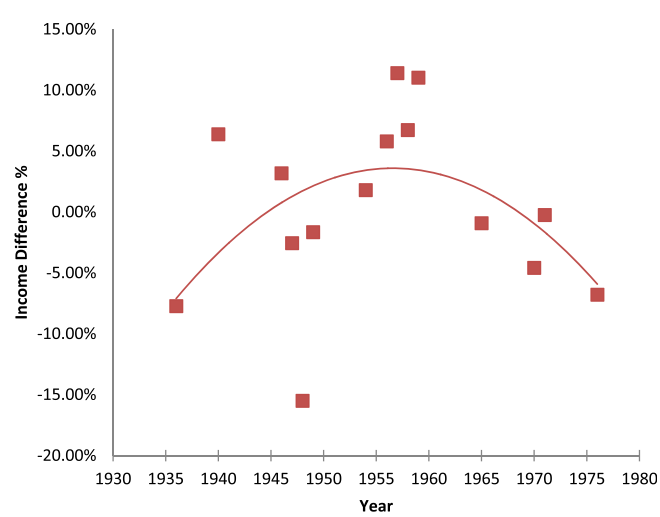

Fig. 3. Women's income differences between grade skippers and non-skippers over 40 years. Polynomial trendline added for ease of interpretation.

analysis, this income difference averaged 3.63% per year, which corresponded to an effect size of $\delta = 0.034$ (See Supplemental File 2 for an explanation of the Cohen's *d*-like effect size δ in a hierarchical linear modeling context.). These values are smaller than the results in this article, though still a noteworthy amount of money when considered as an annual difference compounded across an entire career. Second, both analyses found that academically accelerated females did not earn higher incomes than non-accelerated women.

In the analysis explained in Supplemental File 2 the difference was 2.02% (favoring non-accelerated women), or $\delta = -0.005$. Third, the nonlinear trend in this study (shown in Figs. 2 and 3) was also found in the alternate analysis, indicating that the income differences between groups are not constant throughout adulthood. However, the nonlinear nature of the trends was not equivalent, with income gaps never closing for men and the gap between groups narrowing for women as time progressed. We invite readers to consult Supplemental File 2 and determine which set of results they find more trustworthy.

5.3. General discussion

Our results agree generally with the results from Cronbach's (1996) analysis of the Terman data. Cronbach also found an income advantage for some accelerated male subjects in 1949, 1958, and the 1960's, though he did not state the magnitude of these income differences. Using a different method to control for confounding variables, we also found that male grade skippers had an income advantage over their non-skipping peers in these years. Yet, we observed that income differences narrowed in the 1970's—something that Cronbach (1996) missed by not examining all income variables in the dataset. Cronbach also did not examine women's income differences, and we found that these income differences also appear in the data from the Terman women, though the differences are much smaller.

Our results also generally agree with those from McClarty's (2015b) analysis of NELS:88 data, even though McClarty's subjects were born two generations after Terman's subjects. She found that between 1997 and 2000 students who had skipped a grade earned higher incomes than those who did not. However, McClarty (2015b) found that year after year the income differences between accelerated and non-accelerated increased, whereas we found that the income differences in Terman's sample were somewhat consistent across most of the adult working years (especially for men). This difference may be due to points in the lifespan covered by the two datasets. The NELS:88 dataset only contained income data for four consecutive years (1997–2000), ending when most subjects were 26 years old. Terman, though, did not collect income data until 1936, when the average subject was 26 years old. It is conceivable that McClarty (2015b) observed the growth of these income gaps in the mid-20's and that these gaps may stabilize soon after; our study may not detect these gaps until they have already formed and then continued to observe them for another 40 years. The possibility of a cohort effect or sample idiosyncracies cannot be ruled out either. Analysis with another longitudinal dataset with income data from the teenage years into middle age would aid in understanding the development and growth of income gaps across accelerated and non-accelerated students.

Beyond a comparison with similar studies, it is important to consider the results of this study in light of the wider research context. First, no prior study has examined the consequences of grade skipping into old age. Our study shows that there are potentially long-term benefits to full-grade acceleration. With many of the benefits of educational interventions fade out over time (e.g., Claessens, Engel, & Curran, 2014), the possible persistence of the financial benefits of full-grade acceleration for decades into adulthood is remarkable.

At first glance, the effect size values ($d = 0.102$ for men and $d = 0.002$ for women) are not impressive, especially to readers who use Cohen's (1988) standard of $d = 0.20$ as the threshold for a "small" effect size. For this reason, we join with statistical experts in urging interpretation of effect sizes in context (e.g., Thompson, 2001). In this case, the context is what these effect sizes measure: income differences, which is why we converted the income differences to percentages. In this context, the women's effect size of $d = 0.002$ still seems small: 0.42%, and it is hard to imagine how an income difference of less than half a percent annually could make much of a difference in women's lives (especially when one considers that the income advantage for accelerated women was so inconsistent; see Fig. 2). But when interpreted in this context, we see that the men's effect size of $d = 0.102$, though close to zero, could have an important impact on individuals' lives because it corresponds to a 9.35% annual income difference. Almost nobody would turn down a 9.35% pay raise, and most people would recognize that a pay boost of this size would be beneficial to their lifestyle. Thus, it is apparent that for the men in the Terman sample, the income differences between accelerated and nonaccelerated subjects were practically significant (see Thompson, 2002, for information about practical significance of statistical results).

Likewise, the dollar amounts in Table 3 seem small at first glance. Yet, converting these values to 2015 dollars using a tool from the Bureau of Labor Statistics (n.d.) shows that in modern terms, the financial benefits of grade skipping total to $10,304.52 per year (on average; median = $8831.19) for men and $1165.72 per year (on average; median = $691.65) for women. The value for men may seem high, but it is important to remember that the subjects in Terman's study earned higher incomes than the general population of the same sex at the time (Terman & Oden, 1947, 1959). An additional 9% increase of an income that is already higher than average quickly adds up to a large amount of money, especially as it compounds over time. Because intelligence is correlated with income (Gottfredson, 1997; Jensen, 1998; Nyborg & Jensen, 2001; Strenze, 2007), even slight increases in income for high-IQ adults can translate into large economic benefits.

5.4. Implications

Although these results are fascinating, we are hesitant to recommend policy changes concerning full-grade acceleration solely on the basis of this study. However, when combined with other research on the other benefits of full-grade acceleration, this study shows that the widespread support that gifted education experts give to full-grade acceleration (e.g., Assouline et al., 2015; Colangelo, Assouline, & Gross, 2004; Rogers, 2007, 2015) has an empirical basis. Yet, few students receive the potential financial benefits of full-grade acceleration. In the 21st century in Grades 1–8 approximately 0.25% of students per year in the United States skip a grade (Warren et al., 2014, p. 435). Full-grade acceleration is likewise very rare in most European nations (Hoogeveen, 2015) and in Australia (Young, Rogers, Hoekman, van Vliet, & Long, 2015). A 0.25% annual rate is likely lower than it was in the early 20th century when Terman's subjects were children or adolescents (e.g., Almack & Almack, 1921; Downes, 1913; Madsen, 1920). Indeed, near the end of his life Terman (1954) mourned the decreased popularity of full-grade acceleration.

However, we recognize that it is not clear how many students could be accelerated, and this study says nothing about whether acceleration is underutilized. Yet, there is evidence that many students are capable of schoolwork that is normally assigned in an older grade. One analysis of accountability testing data showed that 12% of students performed at least one grade level higher in mathematics, and 35% performed at least one grade higher in language arts (Makel, Matthews, Peters, Rambo-Hernandez, & Plucker, 2016). It's likely that a noticeable proportion of these students are good candidates for acceleration. While mainstream thought among gifted education experts is that full-grade acceleration is not appropriate for every gifted child (Assouline, Colangelo, Lupkowski-Shoplik, Lipscomb, & Forstadt, 2009; Feldhusen, Proctor, & Black, 2002), it is possible that the practice is appropriate for

more students than experience it now.

Recommendations concerning future research are much easier to make than policy recommendations. First, we recommend that other researchers replicate our findings with more modern samples. McClarty's (2015b) use of the NELS:88 sample is helpful in updating our research findings, for example. Second, we recommend that researchers in the future control for more confounding variables than we, Cronbach (1996), or McClarty (2015b) did. Although controlling for a small number of variables is better than controlling for none, the remaining unobserved covariates are still confounding the results of our study and others. This probably makes grade skipping appear to be a more beneficial intervention than it really is—an issue we discuss in the next section. Finally, we recommend that gifted education researchers undertake experimental longitudinal research on full-grade acceleration by randomly assigning eligible students to either skip a grade or remain in a grade with their age peers. Only with a randomized control group design could education scholars assess the causal impact of full-grade academic acceleration on financial, educational, emotional, and social outcomes without worrying about the potential influence of unobserved confounding variables.[2,3]

5.5. Limitations

Although we believe that our study is an important contribution to gifted education, we recognize that it has several shortcomings. The first—and most insurmountable—is the use of Terman's data. Terman's longitudinal study is a research project "…locked in time…" (Cravens, 1992, p. 184), and the subjects were born, grew up, and worked in a very different historical and cultural context than what 21st century gifted children experience. Most subjects lived through the Depression and World War II in the early phases of life, and this impacted their educational opportunities and life experiences (Holahan & Sears, 1995; Subotnik, Karp, & Morgan, 1989). The historical milieu of the Terman study was especially influential on the work history of the female subjects. Over 40% of women were homemakers as their primary career, and the two most common income producing jobs for women were secretarial work and teaching (Holahan & Sears, 1995, p. 87). These historical realities likely limited female subjects' career choices and opportunities for promotion and advancement; this is likely the reason why the parameter estimates for the path analysis in our study for women were weaker than they were for men. Career trajectories of a sample of gifted women today would be very different, and we doubt that the results for the analysis of the women's data would be applicable to modern students. Nevertheless, the results are interesting from a historical perspective.

Likewise, the geographic and cultural makeup of the sample imposes some limitations on the generalizability of the study. Through much of the Terman sample's working years the United States had a unique cultural, scientific, and economic role in the world, and California exercised a disproportionate influence on American affairs in that time. With some of Terman's sample working in Hollywood, academia, and big business (Shurkin, 1992; Terman & Oden, 1959), there were opportunities available to these individuals that would likely not be available to accelerated individuals in other nations. Yet, we believe that the study is still useful to readers in other nations because grade skipping is not a uniquely American intervention (e.g., Young et al., 2015). And given the ubiquity in many nations of organizing schoolchildren by age, data from an American sample like Terman's can provide clues into the possible outcomes of academic acceleration in other countries.

Second, in addition to typical cohort threats to external validity, Terman's study also has a unique shortcoming that originated with the man himself. Terman interfered in his subjects' lives in many ways, including writing letters of recommendation for employers, using his influence to get subjects admitted to college, making educational recommendations to schools and parents, and developing genuinely close friendships with some (Holahan & Sears, 1995; Leslie, 2000; Shurkin, 1992).[4] Indeed, at midlife 41.4% of men and 52.1% of women said that participating in Terman's study provided benefits in their life (Oden, 1968, p. 38). These facts decrease the external validity of a study that already has poor generalizability.

Third, Terman's dataset is missing variables that educational researchers today would include in almost any longitudinal study on academic and career outcomes. For example, there are no measures of motivation or self-esteem until the subjects reached middle age. Trying to find proxies for these characteristics in childhood in order to control for them was impossible. Other variables in the dataset are poor proxies for modern constructs. An example of this is the measure of personality, which was derived from an instrument designed to measure emotional instability and identify soldiers at risk for mental health problems (see Supplemental File 1). Given the primitive state of personality assessment at the time (see Gibby & Zickar, 2008, for a review), this was the best available instrument to Terman. However, using such an instrument on a juvenile sample today would be unacceptable. In the end, this study is limited—especially in regards to the childhood covariates—by the theory and practice of the 1920's.

Another important limitation is that Terman's data are correlational, which means they are not equipped to answer research

[2] Four of the five peer reviewers of this article had concerns about this recommendation. However, the randomized control trial design has been part of medicine for over 80 years (Emanuel & Miller, 2001) and is recognized as providing the strongest evidence for effectiveness in the educational and social sciences (Graesser, 2009; Skidmore & Thompson, 2012; Winch & Campbell, 1969). Thus, a randomized control trial would not—in and of itself—be unethical. The ethical quandary arises when a treatment that is likely beneficial is withheld from some participants for the purpose of creating a randomized control group. There are two experiences that a control group could receive: either subjects receive an inert treatment (i.e., a placebo), or an active treatment (i.e., a pre-existing therapy, often one which is already widely available). The grade skipping analogue to an inert placebo would be withdrawing the child from school for at least a temporary time period—a clearly unethical option (which would be illegal in many countries, to boot). The analogue to an active treatment would be the "business as usual" procedure of having a child advanced through the primary and secondary education system with his or her age peers. This active control "treatment" would be ethical because the vast majority of subjects would experience it anyway (see Gross, 2004; Hoogeveen, 2015; Warren et al., 2014). Therefore, if non-accelerated subjects in a study are assigned to attend the grade they are enrolled in with their age peers, there is no ethical problem with our proposed study. This viewpoint is in accordance with the latest version of the Helsinki Declaration.

[3] It is also relevant to mention that the magnitude of the benefits of full-grade acceleration (or even whether there are any benefits at all) is unknown until a randomized control trial is conducted. In fact, it could be unethical to recommend grade skipping until such a trial is done because these recommendations will not be based on the best possible evidence (Emanuel & Miller, 2001). In the eyes of some ethicists, advocating for a treatment which has unknown benefits is unethical because it is not possible to weigh the potential harm of the intervention against its benefits. Thus, a randomized control trial of full-grade acceleration may be an ethical imperative so that parents, school personnel, and researchers do not inadvertently harm gifted children by having them skip a grade.

[4] This interference was not always a bad thing. Shurkin (1992) recounted the story of Terman writing a letter to the American authorities on behalf of one of the subjects in his study that was of Japanese descent. In the letter Terman (a well-known psychologist and prominent college professor at the time) vouched for the family's loyalty to the United States, which may have kept them out of the internment camps that many American citizens in California with Japanese ancestry were unjustly confined to during the World War II. Although this is an example of interfering with the internal validity of his study, Terman's decision was ethically impeccable.

questions about causality. Without an experimental design with random assignment of subjects to grade skipping and non-grade skipping groups it is impossible to state whether grade skipping causes income difference in adulthood for gifted subjects. Propensity score modeling can control for confounding covariates, but if those models do not include every relevant covariate related to group assignment, then the model will produce inaccurate results about the impact of an intervention (Steiner, Cook, Shadish, & Clark, 2010). Because propensity score models that are missing relevant covariates usually produce positively biased results, we recommend that readers interpret our results as a maximum estimate of the economic benefits of grade skipping. Additionally, we recommend follow-up studies on the financial impacts of grade skipping with modern samples that control for more covariates and—if ethically and practically possible—a longitudinal study with random assignment to grade skipping or non-grade skipping groups. These types of studies can shed light onto the extent of the overestimate of the impact of acceleration in our study and improve the theortical understanding of why grade skippers might earn higher incomes in adulthood.

Despite these limitations, Terman's study is the *only* study that follows a sample of American gifted children through their entire working lives, which makes it uniquely suited to answer our research questions about the stability of income differences throughout the adult years. Until enough time has passed for other longitudinal studies to have data spanning most of the lifespan of a large number of grade skippers, Terman's study is the only option for answering our research questions.

Even if no other study encompasses such a wide range of the lifespan, some readers may still object to the use of such an old archival dataset. Yet, other researchers have found value in analyzing datasets that are of a similar age or even older. Fancher (1985) re-analyzed data from Spearman's (1904) landmark study on the general intelligence factor and found data anomalies. One group of researchers (Johnson et al., 1985) re-analyzed data from the 1880's and 1890's collected at one of Sir Francis Galton's anthropometric laboratories and found substantial correlations among siblings on many variables (including reaction time, hand grip strength, and visual acuity), indicating a likely genetic component to these traits. Johnson et al. (1985) also found a correlation in Galton's data between social class and several measures (with upper class males exceeding the performance of their lower class counterparts in visual acuity, hearing acuity, reaction time, breathing strength, and other variables). Likewise, Grigoriev, Lapteva, and Lynn (2016) analyzed data from late 19th century imperial Russia to find that regional literacy rates were positively correlated with height of military recruits and negatively correlated with regional infant mortality and fertility rates. Finally, there is a voluminous body of research on secular changes in height and other anthropometric measures from historic time periods—data that still provide valuable information to scholars, even though 21st century living conditions in industrial nations often bear little resemblance to the environment in which the data were originally collected. (See, for example, Komlos, Hau, & Bourguinat's, 2003, study of changes in height of French military recruits from 1666 to 1760). Clearly, archival data like Terman's dataset can be valuable in answering research questions of modern scholars interested in IQ, environmental characteristics, and life outcomes.

6. Conclusion

Although full-grade acceleration is widely supported (Assouline et al., 2015; Rogers, 2007), there is much that educational scholars do not know about the long-term consequences of this intervention. In this study we used data from Terman's (1926) longitudinal study of gifted children to investigate the financial consequences of full-grade acceleration on subjects' annual incomes as adults. We found that male subjects who skipped a grade had incomes that were an average of 9.35% higher than non-grade skippers after controlling for childhood covariates and adult educational attainment. The financial impact of full-grade acceleration for women was much smaller—an average of 0.42%. Our results also indicate that these income differences were stable until towards the end of the subjects' careers when income differences narrowed (for men) or disappeared completely (for women). An alternative analysis of the same data produced similar results (see Supplemental File 2), though with smaller income differences (a 3.63% mean annual income difference for men and a −2.02% mean annual income difference for women). However, the archival nature of the study and the lack of an experimental design in Terman's study means that we cannot say whether these income differences were caused by grade skipping.

We hope that this study, despite its limitations, provides vital information to educational scholars and practitioners about one of the benefits of full-grade acceleration. Although this study should not be used as the sole means of deciding whether to advance a child to a higher grade than his or her age peers, it nevertheless joins the literature on the benefits of full-grade acceleration. This study also provides a basis on which future educational researchers can build upon to learn more about the long-term impacts of grade skipping.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.learninstruc.2016.10.004.

References

Almack, J. C., & Almack, J. S. (1921). Gifted pupils in the high school. *School & Society, 14*, 227–228.

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*, 33–40. http://dx.doi.org/10.3102/0013189x035006033.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Assouline, S., Colangelo, N., Lupkowski-Shoplik, A., Lipscomb, J., & Forstadt, L. (2009). *Iowa acceleration scale manual* (3rd ed.). Scottsdale, AZ: Great Potential Press.

Assouline, S. G., Colangelo, N., VanTassel-Baska, J., & Lupkowski-Shoplik, A. (2015). *A nation empowered: Evidence trumps the excuses holding back America's brightest students*. Iowa City, IA: Belin-Blank Center.

Blau, F. D., & Kahn, L. M. (2007). The gender pay gap: Have women gone as far as they can? *The Academy of Management Perspectives, 21*(1), 7–23. http://dx.doi.org/10.5465/amp.2007.24286161.

Brown, R. L. (1994). Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of five methods. *Structural Equation Modeling, 1*, 287–316. http://dx.doi.org/10.1080/10705519409539983.

Bureau of Labor Statistics. (n.d.). CPI inflation calculator. Retrieved from http://data.bls.gov/cgi-bin/cpicalc.pl.

Burks, B. S., Jensen, D. W., & Terman, L. M. (1930). *Genetic studies of genius: Vol. III. The promise of youth: Follow-up studies of a thousand gifted children*. Stanford, CA: Stanford University Press.

Claessens, A., Engel, M., & Curran, F. C. (2014). Academic content, student learning, and the persistence of preschool effects. *American Educational Research Journal, 51*, 403–434. http://dx.doi.org/10.3102/0002831213513634.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Colangelo, N., Assouline, S. G., & Gross, M. U. M. (2004). *A nation deceived: How schools hold back America's brightest students* (Vol. 2). Iowa City, IA: University of Iowa.

Cravens, H. (1992). A scientific project locked in time: The Terman genetic studies of genius, 1920s-1950s. *American Psychologist, 47*, 183–189. http://dx.doi.org/10.1037/0003-066X.47.2.183.

Cronbach, L. J. (1995). Attrition: Causes and consequences. In C. K. Holahan, & R. R. Sears (Eds.), *The gifted group in later maturity* (pp. 316–331). Stanford, CA: Stanford University Press.

Cronbach, L. J. (1996). Acceleration among the Terman males: Correlates in midlife and after. In C. P. Benbow, & D. Lubinski (Eds.), *Intellectual talent: Psychometric*

and social issues (pp. 179–191). Baltimore, MD: Johns Hopkins University Press.
Downes, F. E. (1913). Seven years with unusually gifted pupils. *The Psychological Clinic, 6*, 13–17.
Emanuel, E. J., & Miller, F. G. (2001). The ethics of placebo-controlled trials — a middle ground. *New England Journal of Medicine, 345*, 915–919. http://dx.doi.org/10.1056/NEJM200109203451211.
Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*, 128–141. http://dx.doi.org/10.1207/s15328007sem0801_7.
Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies, 18*, 3–56. http://dx.doi.org/10.1080/13598130701350593.
Fancher, R. E. (1985). Spearman's original computation of *g*: A model for burt? *British Journal of Psychology, 76*, 341–352. http://dx.doi.org/10.1111/j.2044-8295.1985.tb01957.x.
Fan, X., & Nowell, D. L. (2011). Using propensity score matching in educational research. *Gifted Child Quarterly, 55*, 74–79. http://dx.doi.org/10.1177/0016986210390635.
Feldhusen, J. F., Proctor, T. B., & Black, K. N. (2002). Guidelines for grade advancement of precocious children. *Roeper Review, 24*, 169–171. http://dx.doi.org/10.1080/02783198609553000.
Gagné, F., & Gagnier, N. (2004). The socio-affective and academic impact of early entrance to school. *Roeper Review, 26*, 128–138. http://dx.doi.org/10.1080/02783190409554258.
Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology, 11*, 164–184. http://dx.doi.org/10.1037/a0013041.
Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24*, 79–132. http://dx.doi.org/10.1016/S0160-2896(97)90014-3.
Graesser, A. C. (2009). Inaugural editorial for Journal of Educational Psychology. *Journal of Educational Psychology, 101*, 269, 261. doi:1031037/a0014883.
Grigoriev, A., Lapteva, E., & Lynn, R. (2016). Regional differences in intelligence, infant mortality, stature and fertility in European Russia in the late nineteenth century. *Intelligence, 55*, 34–37. http://dx.doi.org/10.1016/j.intell.2016.01.003.
Gross, M. U. M. (2004). *Exceptionally gifted children* (2nd ed.). New York, NY: Routledge.
Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Thousand Oaks, CA: Sage.
Herrnstein, R. J., & Murray, C. (1996). *The bell curve: Intelligence and class structure in American life* (2nd ed.). New York, NY: Free Press.
Holahan, C. K., & Sears, R. R. (1995). *The gifted group in later maturity*. Stanford, CA: University of Stanford Press.
Hollingworth, L. S. (1926). Children who test above 180 IQ (Stanford-Binet). In *Gifted children: Their nature and nurture* (pp. 222–266). New York, NY: Macmillan.
Hollingworth, L. S. (1942). *Children above 180 IQ, Stanford-binet: Origin and development*. Yonkers-on-Hudson, NY: World Book.
Hoogeveen, L. (2015). Academic acceleration in Europe: A comparision of accelerative opportunities and activities. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 209–223). Iowa City, IA: Belin-Blank Center.
Hoogeveen, L., van Hell, J. G., & Verhoeven, L. (2009). Self-concept and social status of accelerated and nonaccelerated students in the first 2 years of secondary school in The Netherlands. *Gifted Child Quarterly, 53*, 50–67. http://dx.doi.org/10.1177/0016986208326556.
Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association, 106*, 345–361. http://dx.doi.org/10.1198/jasa.2011.tm09599.
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M., & Cole, R. E. (1985). Galton's data a century later. *American Psychologist, 40*, 875–892. http://dx.doi.org/10.1037/0003-066x.40.8.875.
Komlos, J., Hau, M., & Bourguinat, N. (2003). An anthropometric history of early-modern France. *European Review of Economic History, 7*, 159–189. http://dx.doi.org/10.1017/S1361491603000066.
Lee, S.-Y., Olszewski-Kubilius, P., & Peternel, G. (2010). The efficacy of academic acceleration for gifted minority students. *Gifted Child Quarterly, 54*, 189–208. http://dx.doi.org/10.1177/0016986210369256.
Leslie, M. (2000). The vexing legacy of Lewis Terman. *Stanford Magazine*. Retrieved from http://alumni.stanford.edu/get/page/magazine/article/?article_id=40678.
Long, J. E. (1995). The effects of tastes and motivation on individual income. *Industrial and Labor Relations Review, 48*, 338–351. http://dx.doi.org/10.2307/2524491.
Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science, 25*, 2217–2232. http://dx.doi.org/10.1177/0956797614551371.
Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: A 10-year follow-up of the profoundly gifted. *Journal of Applied Psychology, 86*, 718–729. http://dx.doi.org/10.1037/0021-9010.86.4.718.
Madsen, I. N. (1920). High-school students' intelligence ratings according to the Army alpha test. *School & Society, 11*, 298–300. http://dx.doi.org/10.2307/2940642.
Makel, M. C., Matthews, M. S., Peters, S. J., Rambo-Hernandez, K. E., & Plucker, J. A. (2016). *How can so many students be invisible? Large percentages of American students perform above grade level*. Johns Hopkins University Institute for Educational Policy. Retrieved from http://education.jhu.edu/edpolicy/comentary/IEPPolicyBriefFinalPLUCKER.pdf.
McClarty, K. L. (2015a). Life in the fast lane: Effects of early grade acceleration on high school and college outcomes. *Gifted Child Quarterly, 59*, 3–13. http://dx.doi.org/10.1177/0016986214559595.
McClarty, K. L. (2015b). Early to rise: The effects of acceleration on occupational prestige, earnings, and satisfaction. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 171–180). Iowa City, IA: Belin-Blank Center.
Nyborg, H., & Jensen, A. R. (2001). Occupation and income related to psychometric *g*. *Intelligence, 29*, 45–55. http://dx.doi.org/10.1016/S0160-2896(00)00042-8.
Oden, M. H. (1968). The fulfillment of promise: 40-year follow-up of the Terman gifted group. *Genetic Psychology Monographs, 77*, 3–93.
O'Connell, M., & Sheikh, H. (2011). 'Big Five' personality dimensions and social attainment: Evidence from beyond the campus. *Personality and Individual Differences, 50*, 828–833. http://dx.doi.org/10.1016/j.paid.2011.01.004.
Park, G., Lubinski, D., & Benbow, C. P. (2013). When less is more: Effects of grade skipping on adult STEM productivity among mathematically precocious adolescents. *Journal of Educational Psychology, 105*, 176–198. http://dx.doi.org/10.1037/a0029481.
Rogers, K. B. (2007). Lessons learned about educating the gifted and talented: A synthesis of the research on educational practice. *Gifted Child Quarterly, 51*, 382–396. http://dx.doi.org/10.1177/0016986207306324.
Rogers, K. B. (2015). The academic, socialization, and psychological effects of acceleration: Research synthesis. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 19–29). Iowa City, IA: Belin-Blank Center.
Shurkin, J. N. (1992). *Terman's kids: The groundbreaking study of how the gifted grow up*. Boston, MA: Little, Brown and Company.
Skidmore, S. T., & Thompson, B. (2012). Propagation of misinformation about frequencies of RFTs/RCTs in education. *Educational Researcher, 41*, 163–170. http://dx.doi.org/10.3102/0013189x12441998.
Spearman, C. (1904). "General intelligence", objectively determined and measured. *American Journal of Psychology, 15*, 201–293. http://dx.doi.org/10.2307/1412107.
Stanley, J. C. (1976). The case for extreme educational acceleration of intellectually brilliant youths. *Gifted Child Quarterly, 20*, 66–75. http://dx.doi.org/10.1177/001698627602000120.
Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods, 15*, 250–267. http://dx.doi.org/10.1037/a0018719.
Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401–426. http://dx.doi.org/10.1016/j.intell.2006.09.004.
Subotnik, R. F., Karp, D. E., & Morgan, E. R. (1989). High IQ children at midlife: An investigation into the generalizability of Terman's genetic studies of genius. *Roeper Review, 11*, 139–144. http://dx.doi.org/10.1080/02783198909553190.
Terman, L. M. (1926). *Genetic studies of genius: Vol. I. Mental and physical traits of a thousand gifted children* (2nd ed.). Stanford, CA: Stanford University Press.
Terman, L. M. (1954). The discovery and encouragement of exceptional talent. *American Psychologist, 9*, 221–230. http://dx.doi.org/10.1037/h0060516.
Terman, L., et al. (1922-1991). *Terman life-cycle study of children with high ability, 1922-1991*. ICPSR08092-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. http://dx.doi.org/10.3886/ICPR08092.ve.
Terman, L. M., & Oden, M. H. (1947). *Genetic studies of genius: Vol. IV. The gifted child grows up: Twenty-five years' follow-up of a superior group*. Stanford, CA: Stanford University Press.
Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius: Vol. V. The gifted group at mid-life: Thirty-five years' follow-up of the superior child*. Stanford, CA: Stanford University Press.
Thompson, B. (2001). Significance, effect sizes, stepwise methods, and other issues: Strong arguments move the field. *Journal of Experimental Education, 70*, 80–93. http://dx.doi.org/10.1080/00220970109599499.
Thompson, B. (2002). "Statistical","practical", and "clinical": How many kinds of significance do counselors need to consider? *Journal of Counseling & Development, 80*, 64–71. http://dx.doi.org/10.1002/j.1556-6678.2002.tb00167.x.
Wai, J. (2015). Long-term effects of educational acceleration. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 73–83). Iowa City, IA: Belin-Blank Center.
Warne, R. T. (2016). Five reasons to put the *g* back into giftedness: An argument for applying the Cattell–Horn–Carroll theory of intelligence to gifted education research and practice. *Gifted Child Quarterly, 60*, 3–15. http://dx.doi.org/10.1177/0016986215605360.
Warne, R. T., Larsen, R., Anderson, B., & Odasso, A. J. (2015). The impact of participation in the advanced placement program on students' college admissions test scores. *The Journal of Educational Research, 108*, 400–416. http://dx.doi.org/10.1080/00220671.2014.917253.
Warren, J. R., Hoffman, E., & Andrew, M. (2014). Patterns and trends in grade retention rates in the United States, 1995–2010. *Educational Researcher, 43*, 433–443. http://dx.doi.org/10.3102/0013189x14563599.
Wells, R., Lohman, D., & Marron, M. (2009). What factors are associated with grade acceleration? An analysis and comparison of two U.S. databases. *Journal of*

Advanced Academics, 20, 248–273. http://dx.doi.org/10.1177/1932202X0902000203.

Winch, R. F., & Campbell, D. T. (1969). Proof? No. Evidence? Yes. The significance of tests of significance. *The American Sociologist, 4*, 140–143. http://dx.doi.org/10.2307/27701483.

Young, M., Rogers, K., Hoekman, K., van Vliet, H., & Long, L. C. (2015). Acceleration in Australia: Flexible pacing opens the way for early university admission. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 225–240). Iowa City, IA: Belin-Blank Center.