

**School Size and Returns to Education:
Evidence from the Consolidation Movement, 1930-1970**

Christopher Berry
Harvard University
Department of Government
Program on Education Policy and Governance
cberry@fas.harvard.edu

Draft v1.3

In the middle of the twentieth century, a quiet revolution remade public education in the United States. As late as 1930, American schools were small, community-controlled institutions, most employing but a single teacher. From roughly 1930 to 1970, a rapid movement toward centralization and professionalization reduced the number of schools by more than 100,000, as nearly two-thirds of the schools that existed in the former year were eliminated through a process of consolidation. The average size of a school increased fivefold over this short period. In the process, school districts evolved into professionally run educational bureaucracies, some operating hundreds of schools and educating hundreds of thousands of students. It is difficult to imagine a more important change in the organization of public education in the past century.

Despite the dramatic scale and break-neck pace of these changes in the organization of public education, little is known about the consequences of consolidation. Has the quality of public education risen commensurately as schools have become larger and more professional, as consolidation's proponents promised? The small literature on the effects of school size on student outcomes has emerged relatively recently, and so speaks only obliquely to the pre-1970 consolidation movement.¹ A yet smaller literature addressed to the effects of school quality on student outcomes before 1970 has ignored issues school size (e.g., Card and Krueger, 1992a). Several case studies of consolidation (e.g., Reynolds, 1999) provide valuable historical details related to particular states or districts, but offer few general findings. The aim of this paper is to begin filling the gap in our understanding of the consequences of consolidation, focusing on the effects of changing school size on student outcomes, specifically wages.

The paper is organized as follows. The next section provides background information on the school consolidation related trends in the number districts and the state government's role in funding public education. Section 2 reviews the literature on the relationship between school size and student outcomes. Section 3 describes the estimation strategy and data. The main analytical results are provided in section 4. The final section evaluates the consolidation movement in light of these results, discusses potential contemporary policy implications, and suggests directions for future research on consolidation and issues of school size generally.

1. Background

The movement for school consolidation must be seen as part of a larger trend toward the professionalization of education that began in the late nineteenth century.² To the "administrative progressives" of the time, the concentration of authority over schooling in the hands of professional educators was seen as a cure for both the corruption of city school systems and the parochialism of rural systems. Consolidation came first to urban areas, where one of the cornerstones of the progressive attack on rule by political machines was the formal organization of schooling under the leadership of

¹ See Andrews, Duncombe, and Yinger (2002) and Cotton (1996) for reviews of the literature on school and district size.

² The discussion in this paragraph is based on Tyack (1974).

professional superintendents. Reformers then turned their attention to rural areas, where they decried the inefficient, unprofessional and “backward” practices of small community schools. In their vision of a professionally run school, reformers drew their inspiration from the modern corporation, with its principles of “scientific” management by experts.

At the center of this reform movement was a push for larger schools. The leading education reformer of the early twentieth century, Ellwood P. Cubberley, pressed three primary arguments in favor of school consolidation.³ First, in small schools, many of which had only one teacher, the ratio of administrators and school officials to teachers was unnecessarily high. Larger schools allowed for more efficient, centralized administration. Second, at a time when many small schools did not even divide children by grade level, consolidation held the promise of highly specialized instruction. Not only could teachers in large schools specialize by grade, but also by subject area. Third, a consolidated school could provide better facilities at lower cost. For instance, Cubberley’s plan for a model elementary school building included, in addition to classrooms, a manual training room, a library, an assembly hall, a domestic science room, and a science laboratory (1922, 253). Supporting these facilities required concentration of students and resources. In sum, consolidated schools, in the view of Cubberley and other reformers, provided economies of scale in administration, instruction, and facilities.

The desire to consolidate schools was linked directly with the necessity of consolidating districts. It was the general opinion of Cubberley and other reformers of the day that one consolidated school should be created in place of five to seven existing schools, on average (Cubberley, 1922, 227). The average school district at the time, however, had only two schools, and most districts in rural areas operated only a single school. Thus, consolidating five to seven schools usually would require consolidating school districts as well. In Cubberley’s view, therefore, the district system of school governance was “the real root of the matter.” He contended that, “To have a fully organized school board in every little school district in a county, a board endowed by law with important financial and educational powers, is wholly unnecessary from any business or educational point of view, and is more likely to prevent progressive action than to secure it” (Cubberley, 1922, 186). Indeed, in his review of consolidation efforts up to the time of his writing, Cubberley concluded that, “The stronger the district system was entrenched the greater the difficulty of securing consolidations. Both trustees and people seemed to unite to resist any change” (1922, 233). A central aim of the consolidation movement, therefore, was to change the governance of education, placing authority at higher levels, such as township, county, or state governments.

The school consolidation movement, therefore, was about more than professional management. As Strang (1987) argues, school consolidation meant the centralization of authority of public education, along two dimensions. First, consolidation removed day-to-day authority over education from the school community to more distant educational bureaucracies, which Tyack (1974, p. 25) has described as a “transfer of power from laymen to professionals.” That is, as school districts became larger and more complex, day-to-day authority over schools gradually shifted from elected school boards to

³ This discussion draws especially from chapter 10 of Cubberley (1922).

superintendents and administrators. Tellingly, over the period 1930-1970, 9 out of 10 school board members nationwide saw their positions disappear. Meanwhile, the number of superintendents increased and the number of principals ballooned, even as the number of schools and districts dwindled.⁴

Second, state governments took an active role in consolidation. Professional educators linked to state departments of education often spearheaded initiatives to consolidate local schools, as part of broader efforts to expand state control over public education. In other words, not only was local control over education weakened by the elimination of most locally elected school boards, but the authority of the remaining boards was eroded as state governments gradually extended their authority over previously locally controlled issues, such as accreditation, curriculum, and teacher certification (Strang, 1987). In short, as schools became larger school boards became more distant from local communities and lost authority relative to professional administrators and state government officials.

Local resistance to consolidation was often fierce, especially in rural areas where the school was the central institution of the community. As Reynolds observes, in the pre-consolidation era, the local school “was typically the key neighborhood institution binding neighbors and linking them to the larger social and cultural world around them” (1999, 61). Thus, consolidation of the local district, and in particular loss of the local school, often threatened a community’s social cohesion and economic vitality. For these reasons, the impetus for consolidation seldom appears to have come from local communities in the absence of initiatives by state officials. Rather, state governments often induced consolidation through fiscal incentives or forced consolidation by legislatively redrawing district boundaries (Hooker and Mueller, 1970). In response to state pressures, “defensive consolidation” was also common, in which districts rushed to consolidate in anticipation of an even less desired forced consolidation by the state government (Reynolds, 1999). In the long run, as we shall see, few communities withstood the financial and political pressures of the state government and professional educators for long.

1.1 Vanishing Act

After proliferating since the founding of the first colonies, American schools began a precipitous decline around 1920. Figure 1 shows that the number of public schools grew from 116,000 in 1869 to peak at 217,000 in 1920. The number of schools declined rapidly over the succeeding 50 years. The pace of decline slowed in the 1970s, and the number of schools reached a nadir in the late 1980s at around 83,000. Since then, approximately 10,000 schools have been added nationwide, in the first significant burst of (net) new school construction in over 60 years. Also notable over the period was a pronounced shift away from one-teacher schools. In the first year for which data on one-

⁴ The number of school districts declined by 90 percent from 1930 to 1970. Assuming a constant board size, 9 out of 10 elected board positions were eliminated. Strang (1987) shows that the number of superintendents increased slightly over this period, and the number of principals grew severalfold.

teacher schools are available, 1927, they composed 60 percent of all public schools. By 1970, one-teacher schools had all but died out; only about 400 remained as of 1999.

Coincident with the consolidation of schools, the number of districts also declined dramatically. The informal nature of school districts early in the twentieth century is evidenced by the fact that many states did not even keep a count of the number of districts prior to 1930.⁵ The 1931-32 edition of the *Biennial Survey of Education* was the first to report statistics of school districts in each state.⁶ As Figure 2 shows, the data soon revealed a picture of decline in the number of districts. The number of districts fell by half between 1931 and 1953, as over 60,000 districts were dissolved by consolidation. The number of districts declined by half again between 1953 and 1963, and roughly half again by 1973. The number of districts stabilized in the early 1970s and has not changed appreciably over the last 30 years.

At the same time that schools and districts were consolidating to form larger units, the number of pupils attending public schools was on the rise. Average daily attendance (ADA) in public elementary and secondary schools roughly doubled from 1929 to 1969, rising from approximately 21 to 42 million.⁷ These two forces, declining numbers of schools and districts coupled with rising attendance, produced substantially larger educational institutions over the course of the twentieth century. Figures 3 and 4 show the average size of schools and districts over the periods for which data are available. Over the period of rapid consolidation, 1930 to 1970, ADA per school increased from 87 to 440. At the same time, ADA per school district increased from approximately 170 to 2,300 students.⁸ In other words, the average school was five times larger in 1970 than in 1930, and the average school district was 14 times larger. Both schools and districts witnessed their most rapid burst of growth in the years from 1950 to 1970, as increasing attendance rates, the baby boom, and institutional consolidation coincided.

If the growing size of schools and districts represents one form of centralization of authority over public education, the second dimension of centralization was evident in the increasing role of state governments in what had traditionally been a strictly local matter. As demonstrated in Figure 5, the state share of funding for public education grew considerably from about 1930 to 1950, and made a smaller jump again in the late 1970s. The local share of revenue, meanwhile, declined from more than 80 percent early in the century to less than half by the mid-1970s. For about the last 25 years, state and local governments have contributed nearly equal shares of public education funding. The

⁵ Although data on the number of school districts are not available prior to 1931, if districts followed a trajectory comparable to schools, we can infer that the number of districts was at its apex around that time.

⁶ The *Biennial Survey of Education* was the federal government's first publication to systematically track statistics related to state and local education. The *Biennial Survey* began publication in 1869, changed titles to the *Digest of Education Statistics* in 1960, and continues publication under that name to this day.

⁷ Average daily attendance is a better indicator of size than is enrollment. Early in the century, there were often substantial discrepancies between the number of students nominally enrolled in schools and those who actually attended regularly. Today, the two are nearly identical. For a comparison of the average daily attendance and enrollment over time, see Heckman, Layne-Farrar, and Todd (1996).

⁸ From 1970 to 2000, average district size continued to increase, reaching 2,900 students in the latter year.

federal share of education funding has risen noticeably from its starting point of next to nothing in 1920, but still remains at less than 10 percent.

As discussed above, school consolidation must be seen as part of a broader movement of educational reform. Over the same period, 1930 to 1970, for example, the school term grew longer, class sizes grew smaller, and teachers became better paid. [I will add numbers for these trends.] The overall effect of these related reforms was to transform the small, informal, community controlled schools of the 19th century into centralized, professionally run educational bureaucracies. The American public school system as we know it today was born during this brief, tumultuous period.

2. Related Literature

There have been two identifiable waves of literature on school size (Howley, 1993). The first wave studies, appearing roughly from the 1920s through the 1970s, focused primarily on input measures of school quality.⁹ Larger schools were consistently found to be superior in this regard. For instance, larger schools were shown to have better facilities, more highly qualified teachers and administrators, a greater depth and variety of course offerings, and more extracurricular activities. The influential Conant Report represents the high point of this first wave of literature (Conant, 1967). James Conant, the former Harvard University president, studied questionnaires from over 2,000 high schools nationwide and concluded that large “comprehensive” high schools were more cost-efficient and provided higher quality schooling through a wider range of course offerings. Although the Conant Report received great publicity at the time and has been credited with spurring school consolidation, most consolidation that was to occur had already taken place before the report was released.¹⁰

Beginning in the 1980s, the school size literature broadly changed focus from school inputs to student outcomes and achievement. This ongoing second wave of studies has generally been less favorable to large schools. Of the seven studies of school size and student performance reviewed by Andrews, Duncombe and Yinger (2002), only one, Kenny (1982), found increasing returns to scale.¹¹ The remaining six studies found decreasing returns to scale. Four of these studies also identified constant returns to scale over at least some range of the data, suggesting that returns to scale in school size are non-linear (Andrews, Duncombe, and Yinger, 2002). Summers and Wolf (1977) find that African American students are particularly harmed by large school size, while Lee and Smith (1997) find that students of low socio-economic status do particularly poorly in large schools. Although the reasons for the superior performance of small schools have not been definitively identified, explanations have focused on non-academic factors,

⁹ This literature is reviewed by Stemnock (1974).

¹⁰ A little-noted irony of the Conant Report was that there were no small schools in Conant’s data set (Walberg, 1993). The smallest school he examined had 750 students, well above the average high school size at the time.

¹¹ Usefully, Andrews, Duncombe, and Yinger (2002) restrict their survey to studies that meet minimum standards of methodological rigor.

such greater sense of community belonging among students, closer interaction with adults, and greater parental involvement (e.g., Cotton, 1996).

The empirical literature on the effects of *district* size on student outcomes is smaller and less consistent in its findings. Walberg and Fowler (1987) and Ferguson (1991) find a negative relationship between student achievement and district size, controlling for student and teacher characteristics, in New Jersey and Texas, respectively. On the other hand, Sebold and Dato (1981) find increasing returns to district size for California high schools, while Ferguson and Ladd (1996) find increasing returns to district size for elementary schools in Alabama. Unfortunately, as each of these studies focuses on a different state, it is difficult to identify the reasons for the discrepancies in their conclusions. Hoxby (2000) takes a different approach, focusing on competition among districts rather than size per se. In a study of metropolitan areas nationwide, Hoxby (2000) finds a negative relationship between student achievement and the concentration of enrollment in a small number of school districts.

As Andrews, Duncombe and Yinger conclude their review of the literature, “Mixed results emerge from those studies estimating returns to size at the district level. The results from estimates of returns to size at the school-level are more consistent. Generally, larger schools are associated with lower student performance holding school and nonschool inputs constant” (p. 20).¹²

Disappointingly few studies have attempted to explain systematically the *causes* consolidation over time and across states. Indeed, the literature on the subject is sparse enough that it can be reviewed here in its entirety; all three studies. David Strang (1987) was the first to develop and test a model of district consolidation. Strang argues that consolidation came about primarily as a result of efforts by state-level politicians and professional educators to centralize and professionalize the administration of public education. These reformers were able to overcome widespread opposition from local communities and interest groups by providing fiscal inducements for district consolidation. Specifically, by expanding their share of funding for public education, state governments were able to gain leverage over local decision makers by awarding or withholding funds in response to local consolidation efforts. Empirically, Strang (1987) finds that school districts were consolidated more quickly and extensively where the state government’s share of education funding was higher. Indeed, according to his point estimates, a 50 percent increase in state revenue reduces the predicted number of districts by half (Strang, 1987 p.362).

Kenny and Schmidt (1994) were the next to examine the causes of school district consolidation, focusing on the period from 1950 to 1980. They argue that decisions about consolidation present a tradeoff between the benefits of having many districts in order to satisfy the diverse preferences of a heterogeneous population and the benefits of

¹² Cotton (1996) reaches a similar conclusion in her review of over 100 articles on the effects school size on student outcomes. Cotton’s review is more comprehensive than Andrews, Duncombe, and Yinger (2002), but less attuned to methodological issues. It is encouraging that the two surveys come to roughly similar conclusions.

having large districts in order to take advantage of cost savings arising from economies of scale. They found that the decline in farm employment, increasing population density, and falling costs of transporting students were among the most important factors contributing to the decline in the number of school districts. Interestingly, they also found that the rate of teacher unionization was positively associated with consolidation, although they were not able to distinguish whether increasing unionization was a cause or an effect of increasing consolidation.¹³ Like Strang (1987), Kenny and Schmidt (1994) find that the increasing role of state governments in school funding was positively associated with district consolidation. A few state governments (e.g., Florida, Maryland, and Nevada) influenced consolidation more directly by mandating the consolidation of school districts to conform to county boundaries. Finally, Kenny and Schmidt (1994) find that states with greater income heterogeneity experienced less consolidation, but that this relationship was tempered by state funding of public education. Presumably, state funding reduced the variation in quality among districts, and hence removed opportunities for local communities to tailor schools to local tastes (Kenny and Schmidt, 1994 pp. 9-10).

The final study to examine district consolidation is Alesina et al. (2000). This study is more general than the others, in that it examines all types of local governments, not only school districts, and attempts to explain the formation as well as consolidation of local jurisdictions. An additional advantage of this study is that it is conducted at the county level, thus possibly identifying additional sources of variation in the number of local jurisdictions. As with Kenny and Schmidt (1994), Alesina et al. (2000) study the period after 1950, meaning that they miss the era of greatest school district consolidation. Alesina et al. (2000) focus their attention on population heterogeneity as a possible explanation for jurisdictional creation and dissolution. They find that less consolidation took place in counties that were more racially, ethnically, or religiously diverse. Alesina et al. (2000) thus argue that citizens seek to avoid racial heterogeneity within local school districts. These results are generally concordant with Kenny and Schmidt's (1994) finding that *income* heterogeneity makes consolidation less likely, suggesting generally that population diversity has been one of the few significant barriers to the consolidation of local school districts.

Most of the existing literature speaks at best indirectly to questions of the effects of district and school consolidation from 1930 to 1970. Moreover, if returns to size in education are indeed nonlinear, then results of recent studies may differ substantially from what would be obtained from analysis of earlier data. For instance, if returns to size are increasing up to some threshold size, and decreasing thereafter, studies conducted after consolidation had eliminated the smallest districts and schools may only observe cases in the range of decreasing returns. Analysis of data from prior to 1970, when the range of observed values for school and district size was wider, could reveal a

¹³ Unfortunately, because Kenny and Schmidt (1994) examine consolidation after 1950, some of their findings may not apply to the wider period under study in the present paper. For example, teacher unionization was trivial prior to 1960 (see Peltzman, 1993), and so is an implausible explanation for district consolidation over most of the study period.

substantially different picture. The remainder of the paper explores the effects of consolidation using data from the era of consolidation.

3. Empirical Framework and Data

It is not possible to test the effects of consolidation on conventional measures of student achievement, such as standardized test scores, which were not in wide use until after consolidation had largely come to an end. Instead, using the Public-Use Micro-Sample of the U.S. census (PUMS), it is possible to relate aspects of consolidation to the returns to education in the labor market. I follow the methodology established by David Card and Alan Krueger in a pair of influential papers on school quality and earnings (Card and Krueger, 1992a, 1992b). A brief outline of the statistical model follows.¹⁴

Let y_{ijkc} represent the (natural logarithm of) weekly earnings for individual i , born in state j in cohort c and currently working in state k of region r . Let E_{ijkc} represent the years of education completed by individual i , who is assumed to have been education in the public school system of the same state in which he was born. Card and Krueger postulate a linear function of log weekly earnings of the form

$$y_{ijkc} = \delta_{jc} + \mu_{kc} + \beta_c \cdot X_{ijkc} + (\gamma_{jc} + \rho_{rc}) \cdot E_{ijkc} + \varepsilon_{ijkc}, \quad (1)$$

where δ_{jc} represents a cohort-specific fixed effect for each state of birth, μ_{kc} represents a cohort-specific fixed effect for each state of residence, and ε_{ijkc} is a stochastic error term assumed to be identically and independently distributed across individuals. X_{ijkc} is a set of demographic variables, including a marital status indicator, labor market experience, labor market experience squared, and an indicator of whether the individual lives in a metropolitan statistical area. The model thus allows a cohort-specific region of residence effect and a cohort-specific state of birth effect on the return to education. That is, returns to education may differ across different regional labor markets, and returns may differ for individuals educated in a particular state regardless of their labor market.

Because the model includes interactions between state of birth dummies and education, and a second set of interactions between region of residence dummies and education, the state of birth-specific component of the return to education is identified by individuals who are born in one state and move to another. These cohort and state of birth-specific rates of return to education, γ_{jc} , are the key parameters of the model, which Card and Krueger seek to explain through differences in school quality. Specifically, Card and Krueger allow the returns to education for each state of birth and cohort to depend on the characteristics of the public schools, as well as state of birth and cohort fixed effects:

¹⁴ A more detailed methodological discussion of can be found in Card and Krueger (1996) and Heckman, Layne-Farrar and Todd (1996a, 1996b).

$$\gamma_{jc} = a_j + a_c + \varphi Q_{jc}, \quad (2)$$

where a_j and a_c are state of birth and cohort fixed effects, respectively, and Q_{jc} is a set of characteristics of public schools in state j during the education of cohort c . Card and Krueger (1992a) use class size, term length, and relative teacher salaries as their measures of school quality. I aim simply to extend the model by incorporating variables related to consolidation, such as district and school size and the state share of education funding, into Q_{jc} .¹⁵

Several features of this modeling strategy are noteworthy. First, the model is designed to identify the effect of school characteristics on the *slope* of the return to education, φ ; that is, the increase in earnings associated with an additional year of schooling. In addition, permanent differences in the rate of return to education across states are absorbed by the state of birth dummies in the second stage. Second, the first-stage model controls for (1) variation across labor markets in the level of earnings via the state of residence intercepts, (2) regional variation in the rate of return to education via interactions between region of residence dummies and years of education, as well as (3) differences in the average earnings of individuals born in different states via state of birth dummies. These features of the model obviate some common criticisms of studies that attempt to estimate the relationship between school characteristics and student outcomes. For instance, to the extent that family background characteristics (or other omitted variables) affect the level of earnings, rather than the rate of return to education, the estimated rates of return are cleansed of the effects of family background (more on this below).

3. Data

The data used to estimate rates of return to education are from the PUMS A Sample of the 1980 census. Following Card and Krueger (1992a), cases are restricted to white men born in the 48 mainland states and the District of Columbia between 1920 and 1949. The sample is divided into three 10-year birth cohorts. Three separate cohort-specific first-stage regressions were then run, as per equation (1) above, to obtain 147 separate cohort by state of birth estimates of the rate of return to education. These 147 slope estimates become the dependent variable in the second-stage models reported below. In the second stage, the estimated rates of return to education were matched to state-level characteristics of schools at the time each cohort attended school. Data on the

¹⁵ Although this model can be estimated in one step, Card and Krueger argue in favor of two-step estimation, primarily for computational convenience. In the first step, equation (1) is estimated and the state of birth by cohort returns to education, γ_{jc} , are obtained. In the second stage, these estimated returns are used as the dependent variable and equation (2) is estimated by GLS. The two-step estimates are asymptotically unbiased and efficient if proper weights are used in the second stage; specifically, the second-stage observations are weighted by the inverse sampling variance from the first stage.

pupil-teacher ratio, term length, and relative teacher wages (i.e., normalized by the average state wage) were obtained from Card and Krueger (1992a). In addition, I added data on average daily attendance per school and per district, as well as the state government's share of funding for public education. All data on school characteristics, including those of Card and Krueger (1992a), were obtained from various issues of the *Biennial Survey of Education* and later the *Digest of Education Statistics*. Additional details on the data are contained in the Appendix.

Table 1 shows the Spearman rank correlations among the various measures of school quality used by Card and Krueger (1992a) and the three dimensions of consolidation introduced here: school size, district size, and the state's share of funding for public education. Unsurprisingly, as the motivation for district consolidation was often the desire to create larger schools, district size is highly correlated with school size for all cohorts. Interestingly, district size and school size are also highly correlated with the pupil-teacher ratio. The correlations also suggest that teacher salaries were higher in larger districts and, especially, in larger schools. While the state share of funding was positively and significantly correlated with school and district size for all cohorts, consistent with the findings of Strang (1987) and Kenny and Schmidt (1994), state funding was, unexpectedly, even more highly correlated with class size. In other words, the state's expanding role in the funding of public education is associated with students moving into larger districts, larger schools, and larger classes. The relatively high correlations among several of these variables suggest that multicollinearity is an issue warranting attention in the second-stage models. If anything, multicollinearity should inflate standard errors and lead to a bias against finding statistically significant results.

4. Consolidation and the Rate of Return to Education

I began by estimating equation (1) using the 1980 census PUMS data, following the specification of Card and Krueger (1992a). Rates of return to education were estimated using three cohort-specific regressions of log weekly earnings on a set of state-of-residence indicators, 49 state-of-birth indicators, nine region of residence indicators interacted with completed years of education, and 49 state-of-birth indicators interacted with completed years of education.¹⁶ The models also included controls for labor market experience and its square, an indicator for current residence in a metropolitan area, and an indicator for being married with a spouse present. The cohort-specific interactions between state of birth and years of education become the observations for the dependent variable in the second-stage models. Because my first-stage model results are virtually identical to those reported by Card and Krueger (1992a), I do not discuss them further here.¹⁷

Table 2 presents the results of a series of regression models representing versions of equation (2) based on the estimated rates of return generated in the first stage model.

¹⁶ There are 49 states of birth because Alaska and Hawaii are excluded (they were not states until 1959), and Washington, D.C. is included as a "state" of birth.

¹⁷ The data appendix provides additional details.

The first model reproduces the primary specification of Card and Krueger (1992a).¹⁸ Consistent with their results, increasing class size has a negative effect on the return to education, while increasing teacher salaries have a significant positive effect. Next, school size, district size, and the state's share of funding for public education are added to the model sequentially. Of these three consolidation-related variables, only school size shows a statistically significant relationship with the estimated returns to education. The results indicate that increasing school size is associated with a decline in the return to education. When all three variables are added to the model as a group, school size remains the only one to achieve statistical significance. For both district size and the state share of funding, the standard errors are larger than the point estimates. In light of concerns about multicollinearity raised above, it is worth noting that school size remains significant, while district size and the state share of funding remain insignificant, when models (2) to (5) are run without the three Card and Krueger school quality measures (not shown). Also notable is that school size appears to absorb some of the effect of teacher salary, as the latter variable falls to insignificance in all of the models that include school size. The effect of class size, on the other hand, is robust across all of the models.

As discussed above, several recent studies have identified non-linear effects of school and district size on student outcomes. In the context of consolidation, it may be the case that increasing school and district size has a positive effect on returns to education initially, but that once schools or districts become "too big," diseconomies of scale dominate and returns to education diminish. In order to explore this hypothesis, I add quadratic terms, school and district size squared, in model (6). Interestingly, with the addition of the quadratic term, district size becomes significant, suggestive of an inverted u-shaped pattern in the returns to district size. The quadratic for school size, on the other hand, is not significant, and in model (7) it is dropped. Model (7) thus shows that school size has a uniformly negative effect on the returns to education, while returns to district size are first increasing in district size and then decreasing.

The nonlinear effect of district size in model (7), however, is not robust. Sensitivity analysis revealed that the significance of the quadratic term for district size depends exclusively on the effect of Washington, D.C. (hereafter D.C.). With only one school district throughout the study period, D.C. has the largest average district size in every cohort.¹⁹ For example, the average size of the D.C. district is 79,000, 85,000, and 98,000 students for the 1920-29, 1930-39, and 1940-49 cohorts, respectively. The second-largest state average district size in each of the three cohorts was 11,000, 13,000,

¹⁸ Card and Krueger (1992a) report two version of their model. In one, the rates of return to education in the first stage are estimated to be linear in log earnings. In the other, they use an ad-hoc non-linear model in which the years of education variable is coded relative to the second percentile of the distribution for each state. They find little difference between the two models. Heckman, Layne-Farrar, and Todd (1996), in their replication of the Card and Krueger (1992a) model, ignore the two percent non-linear model. I do the same here and simply estimate the first stage as linear in log earnings. Thus, the second stage results presented here should be compared with the equivalent estimates of Card and Krueger (1992a), column 10 of Table 5. I am able to nearly, but not perfectly, replicate their results. See the Appendix for further discussion.

¹⁹ Recall that Hawaii, which also has only one district, is excluded from the analysis.

and 20,000, respectively.²⁰ Thus, D.C. appears to be a substantial outlier in district size for every cohort. However, D.C.'s estimated return to education is not proportionate to its size. D.C. has the second-highest estimated return for the 1920-29 cohort, approximately an average return for the 1930-39 cohort, and is just below average for the 1940-49 cohort. Model (8) shows that when D.C. is excluded, district size and district size squared are no longer significant. However, when the quadratic term is dropped in model (9), district size becomes significant and positive with D.C. excluded. Interestingly, although D.C. also has the largest average school size for every cohort, the estimated effect of school size is essentially unaffected when D.C. is dropped from the model. If D.C. is excluded as an outlier, model (9) is the preferred specification and the results indicate increasing returns to scale in district size, but decreasing returns in school size.²¹

The point estimates from model (9) reveal substantively consequential effects of consolidation on the returns to education. An increase of one standard deviation in school size (equivalent to about 100 students) is associated with a decrease of one-third of a standard deviation in the rate of return to education. On the other hand, an increase of one standard deviation in district size (about 2,800 students) is associated with about one-fifth of a standard deviation's increase in the rate of return to schooling. For comparison, note that a one-standard deviation increase in class size (4 students) is estimated in model (9) to increase the return to education by about one-quarter of a standard deviation. Put more directly, an increase of school size by 100 students is associated with a 3.7 percent decline in earnings for high school graduates (those with exactly 12 years of education). An increase of district size by 2,800 students is associated with a 2.5 percent increase in earnings for high school graduates.

4.1 Sensitivity Analysis

Heckman, Layne-Farrar, and Todd (HLT) (1996) provide a thoughtful critique of the Card and Krueger (1992a) study and its methodology.²² The most important issue raised by the Heckman team is that Card and Krueger are not able to separate the effects of school characteristics from family background or other early community influences. Card and Krueger merely posit that family background affects the level of earnings rather than the rate of return to education. HLT challenge this assumption as arbitrary and point out that if it does not hold, even Card and Krueger's state fixed effects strategy provides no guarantee that estimated effects of school quality are not merely proxies for effects of

²⁰ Maryland had the second-largest average district size in every cohort.

²¹ Another case for excluding D.C. is that it is not actually a state. Practically, however, the inclusion or exclusion of D.C. is of little consequence. For model (7), the quadratic in school size is maximized at 48,000 students, well over twice the average size for every state except D.C. Thus, even under this specification, returns to education are increasing in district size over the practically relevant range of the data. Nevertheless, the fundamental question of whether district size ever reaches a point of decreasing returns to scale is an important one. Analysis of more recent data, where there are more observations comparable in scale to D.C., may help to resolve this issue.

²² Card and Krueger (1996) provides a response to this critique.

other early environmental factors on earnings. Because my analysis follows the Card and Krueger empirical methodology, these issues warrant attention here.²³

As discussed in Section 2 above, the rapid increases in school and district size from 1930 to 1970 should be seen as part of a larger movement toward centralization and professionalization in American public education. Over this period, not only did schools and districts become larger, but classes became smaller, terms longer, teachers better paid, and state governments more involved in school governance. Of course, the timing and extent of these changes varied from state-to-state. If the pace of reform in a state was a function of the voters' valuation of educational quality, then we must be concerned that the observed effects of school quality using the Card and Krueger method merely proxy for unobserved early environmental characteristics. That is, if voters who placed a higher value on education were more likely to reform their school systems, then it will be difficult to disentangle the effects of those reforms from the effects of being raised in a community that places a high value on education. For instance, we have seen some evidence that returns to education are positively related to district size, positively related to relative teacher wages, and negatively related to class size. Yet growth in district size, growth in teacher wages, and decreases in class size were related elements of a general reform movement in education over this period. One may wonder whether these observed effects are merely proxies for unobserved characteristics of states with greater propensity for educational reform.

School size stands apart from the other quality variables. If communities with greater concern for educational quality were more likely to consolidate or to create larger schools, and if it were this (unobserved) community concern with education rather than school size itself that affected outcomes, then we would expect observed school size to be *positively* related to returns to education. But we find just the opposite. Although increasing school size was part of the same overall movement, its effects run counter to the other reform variables. That is, schools got bigger just as districts got bigger and classes got smaller. But whereas larger districts and smaller classes appear to increase returns to education, larger schools reduce returns to education. Thus, if district size and class size effects can be attributed to omitted community background variables, these same omitted variables cannot explain the school size effects.

Is it possible, on the other hand, that communities desiring better education would have been less, not more, likely to consolidate schools? This seems unlikely, given that the educational experts of the time were promoting larger schools. Educational luminaries such as Ellwood Cubberley early in the century and later James Conant, were powerful advocates of larger schools. A 1948 publication by the National Commission

²³ Heckman, Layne-Farrar, and Todd (1996) suggest several refinements to the Card and Krueger analysis. First, the Heckman group suggest that returns to education are nonlinear, finding evidence of significant "sheepskin effects" in the first-stage model at 12 and 16 years of education. Second, they suggest that school characteristics may influence the first-stage state-of-birth intercepts, in addition to slopes, and that these effects can be modeled in a second-stage regression analogous to equation (2) above. Third, they show that school characteristics affect educational attainment (years of schooling completed). Fourth, Heckman, Layne-Farrar, and Todd argue that returns to education region-of-residence specific. I plan to pursue each of these issues, as they pertain to the consolidation effects, in future work.

on School District Reorganization, entitled *Your School District*, is typical of the time in its call for larger schools through consolidation. According to the report, “Educational leaders have recognized the difficulty of providing a satisfactory educational program in schools with small enrollments.” Like many at the time, the authors were concerned with developing *minimum* size standards.²⁴ Therefore, there is good reason to believe that education-focused parents would have had a preference for larger schools, not smaller ones. Thus, in order to conclude that the observed negative effects of school size merely proxy for early community influences, we would also be forced to conclude that “better” communities provided smaller schools, contrary to the recommendations of the education experts of the day.

In short, I see no simple, plausible omitted variables story that can explain away the observed negative effects of school size. That schools size displays a negative relationship with educational returns, contrary to the other educational reforms (larger districts and smaller classes) *and* contrary to the expectations of contemporary education experts, bolsters the interpretation of this as a causal relationship rather than an artifact of unobservable early community influences.

We can do better than speculate on the relationship between consolidation and early environmental characteristics, however. Although individual-level data on parental characteristics or other early environmental variables are not contained in the census, we can obtain information about the characteristics of the population in the state at the time when the men in our sample were in school. In this regard, I collected two relevant variables: the real per capita income in the state at the time each cohort entered school and the corresponding percentage of the population classified as *rural* by the census. Spearman rank correlations between these two variables and school size, district size, and class size are shown in Table 3. Income is negatively related to the pupil-teacher ratio, indicating that more affluent states provided smaller classes, and this correlation is increasing over time. Thus, parental income presents itself as a possible partial explanation for the observed effect of class size. However, income is unrelated to district size for all three cohorts, and *positively* related to school size. That more affluent states had larger schools reinforces the argument the school size effects are not attributable to early environmental characteristics. At least based on these simple correlations, it appears unlikely that school or district size is merely a proxy for parental income. On the other hand, school size is strongly negatively correlated with the proportion of the population classified as rural, and this correlation increases somewhat across the cohorts. So if it is not family income but rather some factor associated with being raised in a rural community that is associated with returns to education, then the estimated effects of school size in Table 2 may be confounded.

²⁴ The authors reached the “unequivocal conclusion” that both elementary and high schools should have an enrollment of 300, and an average of 25 pupils per teacher (National Commission on School District Reorganization, 1948, pp. 22-23). While these would not be large schools by today’s standards, it is important to recall that the average elementary school in 1947-1948 had only 125 students, the average high school had 220, and the average pupil-teacher ratio was 28 (NCES, 1993).

In order to get at these issues more directly, I ran a series of models incorporating the income and rural variables, reported in Table 4. As seen from model (1), the income of the parents' generation actually shows a statistically significant *negative* relationship with returns to education. This result suggests that children raised in less affluent states derived greater returns from a year of education. That the income of the parents' generation is significant here challenges Card and Krueger's (1992a) assumption that family background variables influence the level of earnings but not the rate of return to education. Nevertheless, the estimated effects of school and district size are largely unaffected by the inclusion of the income variable. The rural variable, on the other hand, shows no significant association with returns to education. In fact, even when all of the school characteristics are removed from the model (not shown), rural population does not approach statistical significance. Thus, it appears unlikely that school or district size merely proxy for unobservable factors associated with a rural upbringing.

Another concern is that the observed effects of school and district size are due to biases that arise as larger fractions of the population attend school and stay in school longer. That is, if the expansion of schooling opportunities draws less able students into the educational system, *and* if it is the addition of such students that causes district and school sizes to increase, then the estimated effects of school and district size may merely reflect the changing skill distribution of students. A similar problem could arise as less able students stay in school longer, *if* this accounts for increases school and district size. On the other hand, it is possible that the consolidation of schools and districts, which generally resulted in longer travel times for students, may have deterred students with lower expected returns to education from attending school, which would bias our results toward finding positive effects of school and district consolidation. Thus two possible sources of bias related to shifting enrollment, but pointing in opposite directions, are possible. Indeed, the two effects could offset one another. Ultimately these are empirical questions.

In order to gain insight into the relationship between consolidation and attendance, I collected data on the fraction of the state population in average daily attendance at the time each cohort was in school. The proportion in attendance was not correlated with district or school size for any of the cohorts.²⁵ When this attendance measure is added to the model it is nowhere near significant, nor does it influence the estimated effects of school or district size. Based on this crude measure, then, there is little evidence to suggest that the observed effects of school and district size merely proxy for changes in the skill composition of the student population.

The preceding analyses have considered each of the environmental variables individually. In model (4), all of the background variables (income of the parents' generation, percent of the population classified as rural, and the proportion of the population attending school) are included jointly. Only the income variable is statistically significant, and even this only at the 10 percent level. Importantly, the estimated effects for school and district size are basically unperturbed. In summary,

²⁵ The Spearman rank correlation between proportion in attendance and school size was significant at the 10 percent level for the 1930-39 cohort. The correlation coefficient was $-.26$.

although income and rural status are correlated with school size, neither background variable appears to account for effect of school size on returns to education. District size is unrelated to background income and only modestly correlated with rural status. Hence it is unsurprising that these variables do not absorb the effects of district size when included in the model. The attendance rate is essentially unrelated to school or district size, and exhibits little influence in either of the models in which it is present. Taken together, the results reported in Table 4 suggest that school and district size are unlikely proxies for early environmental characteristics.

A second criticism leveled against the Card and Krueger method by HLT is that the estimates could be biased by nonrandom migration. As explained above, the Card and Krueger model is identified by individuals born in one state but working in another, assuming that migration across states is random. If migration is not random, state-of-residence effects may be conflated with background effects. However, the direction of the potential bias is unclear. As HLT acknowledge, “Theoretically, not accounting for the bias caused by nonrandom migration could result in either over- or underestimating quality effects” (1996, 27f). In the present case, given the high correlation between school size and district size, and their opposing effects on returns to education, it is difficult to imagine how these results could arise simply from selective migration bias.

An final issue that has dogged the Card and Krueger (1992a) analysis deserves comment. Card and Krueger’s (1992a) finding of a positive effect of school quality on earnings – in particular, the effect of class size – runs counter to many studies using individual-level data, and has led others to question whether the Card and Krueger result is an artifact of “aggregation effects” (e.g., Hanusheck, Rivkin, and Taylor, 1996). Without delving into the aggregation debate, it is reassuring that the present findings related to school size are broadly consistent with other studies using individual-level data, as reviewed in Andrews, Duncombe, and Yinger (2002) and Cotton (1996). While the literature on district size is smaller and less consistent in its results, the finding of a small positive effect is not out of line with some other studies using individual-level data (e.g., Ferguson, 1991; Walberg and Fowler, 1987).

Directions for Future Research

As indicated above, HLT suggest several important extensions to Card and Krueger’s (1992a) empirical strategy. Pursuing these lines of analysis in relation to the school and district size variables introduced here is an essential next step. In particular, gaining a better understanding of non-linearity in the return to education, the relationship between size and educational attainment, and regionally-specific returns to education would go far toward solidifying the basic conclusion reached here that size has an important effect on returns to schooling. Analyzing whether district or school size influences educational attainment, in particular, is an important complement to the analysis presented here related to the returns to a year of schooling.

Analyzing the effects of school and district size using individual-level data would go far toward allaying concerns about aggregation bias raised in relation to the Card and Krueger (1992a) study. Although many individual-level studies have found positive effects of school size on various student outcomes (Cotton, 1996), none has focused on earnings. In addition, future work should distinguish the effects of size for elementary and secondary schools. Because the *Biennial Survey* did not begin reporting enrollment for secondary school separately until the early 1940s, I was not able to differentiate size by level of education. However, such analysis could easily be conducted with the individual-level data used in more recent studies.

Most of the work exploring the relationship between school characteristics and earnings has focused on average earnings across students. Perhaps an equally important question is whether school characteristics influence the variance of earnings across students. In the present context, one effect of district consolidation presumably was to reduce the variance in educational quality across students (Kenny and Schmidt, 1994). Even if, as indicated by the results shown here, district consolidation had nominal effects on mean earnings, it may have had substantial effects on the variance. In other words, if equality of educational opportunity is a policy goal, consolidation may have benefits (or costs) that are not observed in analyses of mean earnings, such as presented here.

Finally, the results presented here do little to explain what it is about small schools or large districts that affects student outcomes. Potential explanations for the positive effects of school size range from participation in extracurricular activities and attachment to community to parental involvement and self-esteem (Cotton, 1996). Benefits of district size are typically seen to include cost reductions due to economies of scale (Duncombe and Yinger, 2000). Narrowing the analysis from general considerations of size to identify the specific mechanisms by which size matters will be essential for effective policy.

5. Discussion

The analysis presented above represents the first attempt to assess the effects of the consolidation movement across states and over time, during the period of greatest consolidation, 1930-1970. The results would not please Elwood P. Cubberley and other early twentieth century “administrative progressives” (Tyack, 1974). Contrary to the disparaging picture of small schools painted by the reformers, we find evidence that students from small schools obtained higher returns to education. While there do appear to be modest gains associated with larger *districts*, these gains are outweighed by the harmful effects of larger *schools*. From the estimates reported above, it appears that a one standard deviation change in school size has about one-and-a-half times as large an effect as a one standard deviation change in district size.

If the results indicate that the combination of larger districts with smaller schools provides the greatest returns for students, policymakers have historically not followed such a recipe. The Spearman rank correlation between school and district size across

states remained nearly constant at about 0.70 from 1930 to 1970. That is, larger districts tend to operate larger schools. Indeed, the mix of school and district size is central to issues of authority and governance in education. The number and size of schools within a district directly influence the extent to which central authorities, such as superintendents and school boards, can be directly involved in the operations of their schools. In other words, school board members may face a conflict of interest between good education policy and the maintenance of their own authority over schools. For instance, given a district of 10,000 students, it will be less costly for central authorities to monitor the operation of 10 schools with 1,000 students each than to monitor 25 schools of 200 students apiece. A shift toward smaller schools would require central authorities either to spend more time and money on oversight or to become less directly involved in the operation of individual schools. Thus, any consideration of moving toward smaller schools is intertwined with the decentralization of authority within school districts.

After professional educators and centralized education bureaucracies have struggled long and hard to centralize power, would it be in their interest to give some control back in the form of smaller schools? Although the evidence points to the benefits of smaller schools, education policy is seldom made on the basis of good research alone, especially policy that would involve transfers of power. Other contributors to this volume offer clues about the likely politics of decentralization in large school districts. The chapters by Hess and Moe agree that unions are most active in large school districts.²⁶ Campbell (this volume) finds that ordinary citizens, in contrast, are least likely to attend school board meetings in large districts. Thus, as in so many areas of education policy, proposals for decreasing school size must confront the power of the teacher unions. The high observed correlation between school size and teacher salaries (Table 1) suggests that unions may be at least skeptical of, if not downright hostile toward, decreases in school size. On the other hand, teachers, like students, may appreciate the intangible benefits of community attachment and individual participation associated with small schools (Cotton, 1996). In either case, there can be little doubt that decentralization proposals that attain union support will go farther faster than plans that raise union ire.

Despite the long history of centralization and professionalization in education, and the undoubted influence of teacher unions, ultimate authority still lies with voters, although their channels of influence may be less direct than in the past. Limited survey research suggests that voters generally support smaller schools (Cotton, 1996), and as additional evidence accumulates in their favor, public sympathy could grow. Even if voters in large districts are less likely to attend school board meetings (Campbell, this volume), they still rule at the ballot box, where board members are regularly voted in and out of office. And, encouragingly, Berry and Howell suggest, in another chapter of this volume, that voters do, in fact, reward and punish school board members at the polls. Thus, whether smaller schools, and the decentralization of power they represent, have the potential to become anything more than policy prescriptions will depend on the voters'

²⁶ Moe (this volume) provides a more detail analysis of how union activity translates into political influence, and suggests that unions may be equally influential in small districts, even though they are less active politically.

appetite for challenging what William Fowler (1992) has described as “the natural predilection in American education toward enormity.” If history is any guide, it will be an uphill battle.

Implications for contemporary education policy, however, must be drawn only with caution from the analysis presented here, for several reasons. First, I have not examined any school or district size data more recent than 1966. Much can change over thirty-some years. Second, the findings pertain to state *average* school and district size. One must therefore be cautious in trying to ascertain the “right” size for any individual school or district. Third, the analysis focused on only one of the channels by which school and district characteristics influence earnings: the rate of return to a year of education. As Heckman, Layne-Farrar, and Todd (1996) point out, there are at least two additional channels to be considered: effects on the level of earnings, and effects on educational attainment, which in turn affects earnings. Finding that school or district size influences earnings differentially through these various channels could lead to a more complex picture than that presented here.

Data Appendix

1980 Census Data

Samples are taken from the 5% Public Use A file, which is a self-weighting sample of the U.S. population. To maintain comparability with Card and Krueger (1992a), I follow their case selection criteria. Specifically, cases are restricted to white men born between 1920 and 1949 in the 48 mainland states and District of Columbia.²⁷ Cases with imputed data for age, race, sex, education, weeks worked, or earnings are dropped. Individuals who reported no weeks of work, annual wage and salary income of less than \$101, or average weekly wage and salary income of less than \$36 or more than \$2,500 were excluded. Using these selection criteria reported by Card and Krueger, I was not able to reproduce their exact sample. They report a total sample size of 1,019,746.²⁸ I obtain a total sample size of 994,883. Nevertheless, the discrepancy does not appear consequential. The correlation between my first-stage estimates of the rates of return to education and those reported by Card and Krueger is 0.99. The correlation between my estimates of first-stage errors and theirs is 0.99. When I use my estimated rates of return and their school quality variables, I am able to reproduce their main results very closely.

School Characteristics

Data on the pupil-teacher ratio, term length, and relative teacher wages were obtained from Table 1 of Card and Krueger (1992a). Data on average daily attendance, the number of public schools, the number of school districts, and the state share of funding for public education were obtained from various issues of the *Biennial Survey of Education* and, after 1960, the *Digest of Education Statistics*. Because data from the *Biennial Survey* are available only every two years, I coded each estimate to the odd year of the issue and linearly interpreted values for the even year. For instance, the values reported in the 1931-32 and 1933-34 editions were assigned to 1931 and 1933, respectively. The value for 1932 was then computed as the average of the 1931 and 1933 values. Each cohort was assigned the average of the school characteristics during the years people born in that cohort would have attended school. Following Card and Krueger (1992a) and Heckman, Layne-Farrar, and Todd (1996), I assume that all individuals completed 12 years of schooling.²⁹ For instance, a high school graduate born in 1920 would have entered school in 1926 and graduated in 1937. So school characteristics were averaged over 1926-1937 for individuals born in 1920. For the 1920-29 cohort, averages were taken for years of birth from 1920 through 1929, weighted by the number of births in each year.

²⁷ Year of birth was estimated from information on quarter of birth and age.

²⁸ This is the sample size reported in the original article (Card and Krueger, 1992a). However, in a later paper summarizing that article, they report its sample size as having been 1,018,477 (Card and Krueger, 1996).

²⁹ Both Card and Krueger (1992a) and Heckman, Layne-Farrar, and Todd (1996) report estimating models using individual-specific averages of the quality variables, and finding that the results did not change.

Other Variables

The income of the patents' generation for each cohort is per capita income from the *State Personal Income Estimates* of the Bureau of Economic Analysis. The state-level per capita income from 1930 was assigned to the 1920-29 cohort, 1940 to the 1930-39 cohort, and 1950 to the 1940-49 cohort. I used the consumer price index to convert all of the estimates into 1950 dollars.

The percent of the population classified as rural was taken from the 1930, 1940, and 1950 U.S. censuses. As with the income estimates, the 1930 value was assigned to the 1920-29 cohort, 1940 to the 1930-39 cohort, and 1950 to the 1940-49 cohort.

The percent of the population in average daily attendance was computed based on estimates of average daily attendance reported in the *Biennial Survey of Education* and *Digest of Education Statistics*, and the annual state population estimates from the census bureau.

REFERENCES

- Alesina, Alberto, Reza Baqir, and Caroline Hoxby. 2000. "Political Jurisdictions in Heterogeneous Communities." NBER Working Paper 7859.
- Andrews, Mathew, William Duncombe, and John Yinger. 2002. "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?" *Economics of Education Review* 21, 245-262.
- Berry, Christopher, and William Howell. 2003. "Retrospective Voting in Single Function Elections: Schools Boards, Test Scores and Incumbents' Electoral Fortunes." Working Paper, Harvard University, Department of Government.
- Card, David, and Alan Krueger. 1992a. "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100, 1-40.
- Card, David, and Alan Krueger. 1992b. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics* 107, 151-200.
- Conant, James. 1967. *The Comprehensive High School*. New York, NY: McGraw Hill Book Company.
- Cotton, Kathleen. 1996. "School Size, School Climate, and Student Performance." School Improvement Research Series, Office of Educational Research and Improvement.
- Ferguson, R. F. 1991. Paying for Public Education: New Evidence on How and Why Money Matters. *Harvard Journal of Legislation*, 28, 466-498.
- Ferguson, R. F., and Ladd, H. F. 1996. Additional Evidence on How and Why Money Matters: A Production Function Analysis of Alabama Schools. In H. F. Ladd, *Holding Schools Accountable: Performance-Based Reform in Education*.
- Fowler, William. 1992. "What Do We Know about School Size? What Should We Know?" Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 347 675). Cited in Cotton (1996).
- Hanusheck, Eric, Steven Rivkin, and Lori Taylor. 1996. "Aggregation and the Estimated Effects of School Resources." *Review of Economics and Statistics* 78, 611-627.
- Heckman, James, Anne Layne-Farrar, and Petra Todd. 1996. "Human Capital Pricing Equations with an Application to Estimating the Effect of School Quality on Earnings," *Review of Economics and Statistics* 78, 562-610.
- Hooker, Clifford, and Van Mueller. 1970. *The Relationship of School District Reorganization to State Aid Distribution Systems*. National Education Finance Project Special Study No. 11.
- Fox, W. F. 1981. "Reviewing Economies of Size in Education." *Journal of Education Finance*, 6.
- Hoxby, Caroline. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90, 1209-1238.
- Kenny, Lawrence. 1982. Economies of Scale in Schooling. *Economics of Education Review* 2, 1-24.
- Kenny, Lawrence, and Amy Schmidt. 1994. "The Decline in the Number of School Districts in the U.S.: 1950-1980." *Public Choice* 79, 1-18.

- Lee, V. E., and Smith, J. B. 1997. High School Size: Which Works Best and for Whom? *Educational Evaluation and Policy Analysis* 19, 205–227.
- Reynolds, David. 1999. *There Goes the Neighborhood: Rural School Consolidation at the Grass Roots in Early Twentieth-Century Iowa*. Iowa City: University of Iowa Press.
- Sebold, F. D., and Dato, W. 1981. School Funding and Student Achievement: An Empirical Analysis. *Public Finance Quarterly* 9, 91–105.
- Strang, David. 1987. “The Administrative Transformation of American Education: School District Consolidation, 1938-1980.” *Administrative Science Quarterly* 32, 352-366.
- Summers, A. A., and Wolfe, B. L. 1977. Do Schools Make a Difference? *American Economic Review* 67, 639–652.
- Tyack, David. 1974. *The One Best System: A History of American Urban Education*. Cambridge, MA: Harvard University Press.
- Walberg, Herbert. 1993. “Losing Local Control of Education.” Heartland Institute Policy Study No. 59.
- Walberg, H. J., and Fowler, W. J. 1987. Expenditure and Size Efficiencies of public School Districts. *Educational Researcher* 16, 5–13.

Figure 1

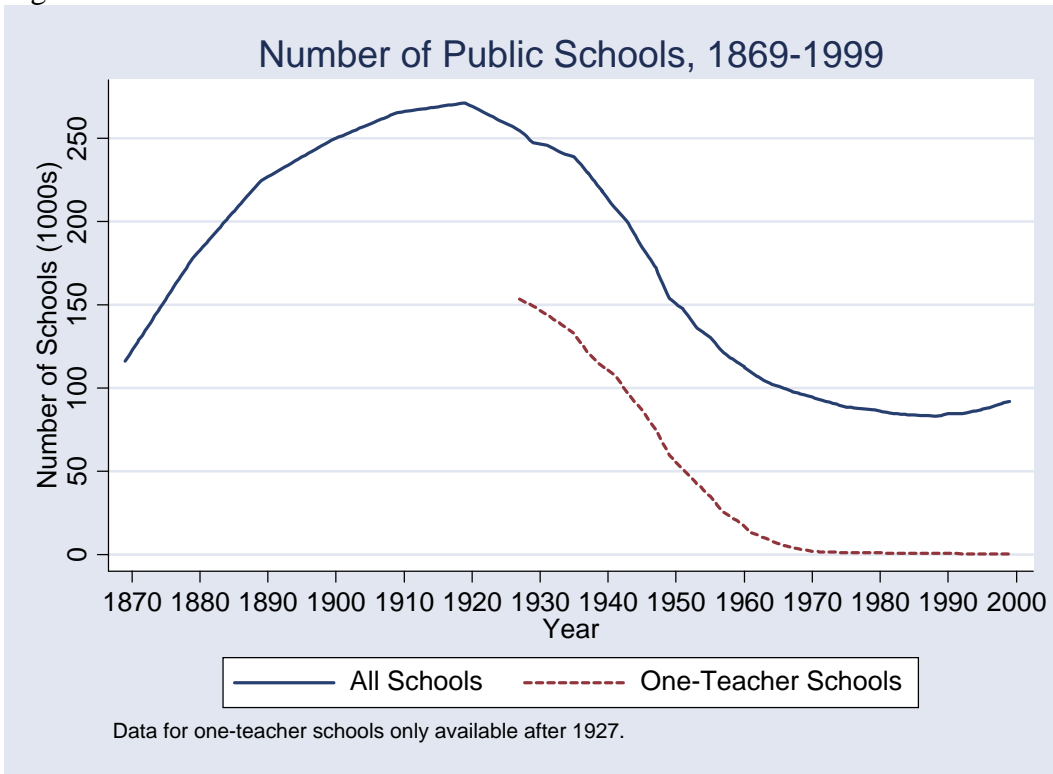


Figure 2

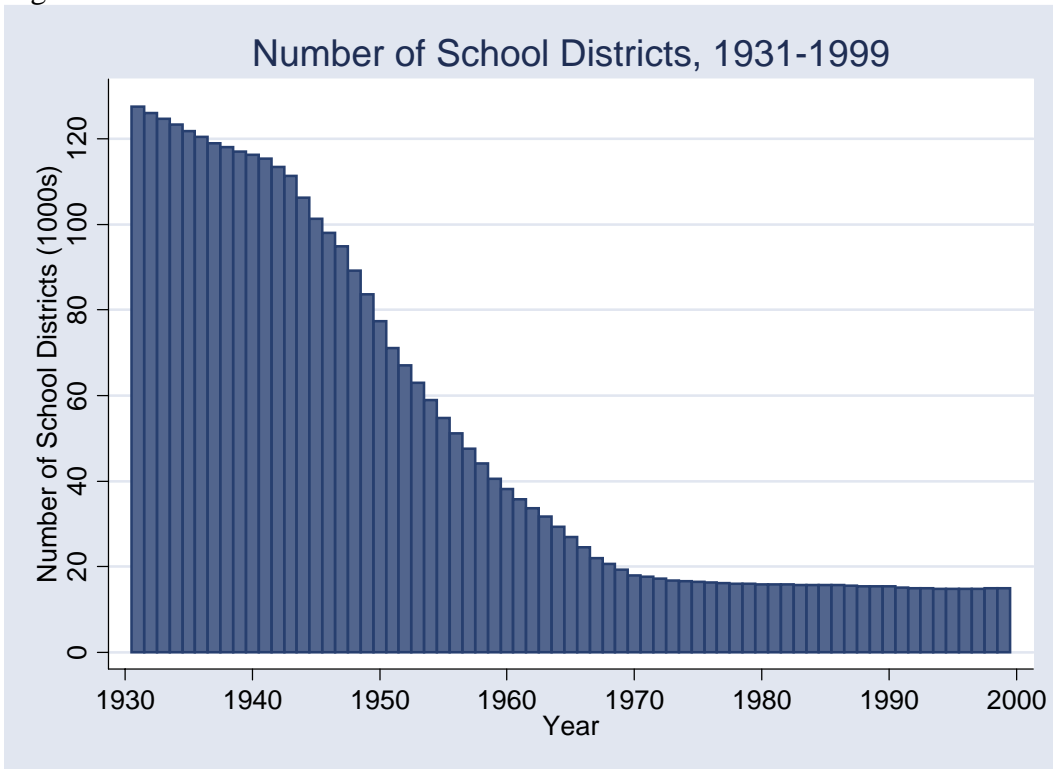


Figure 3

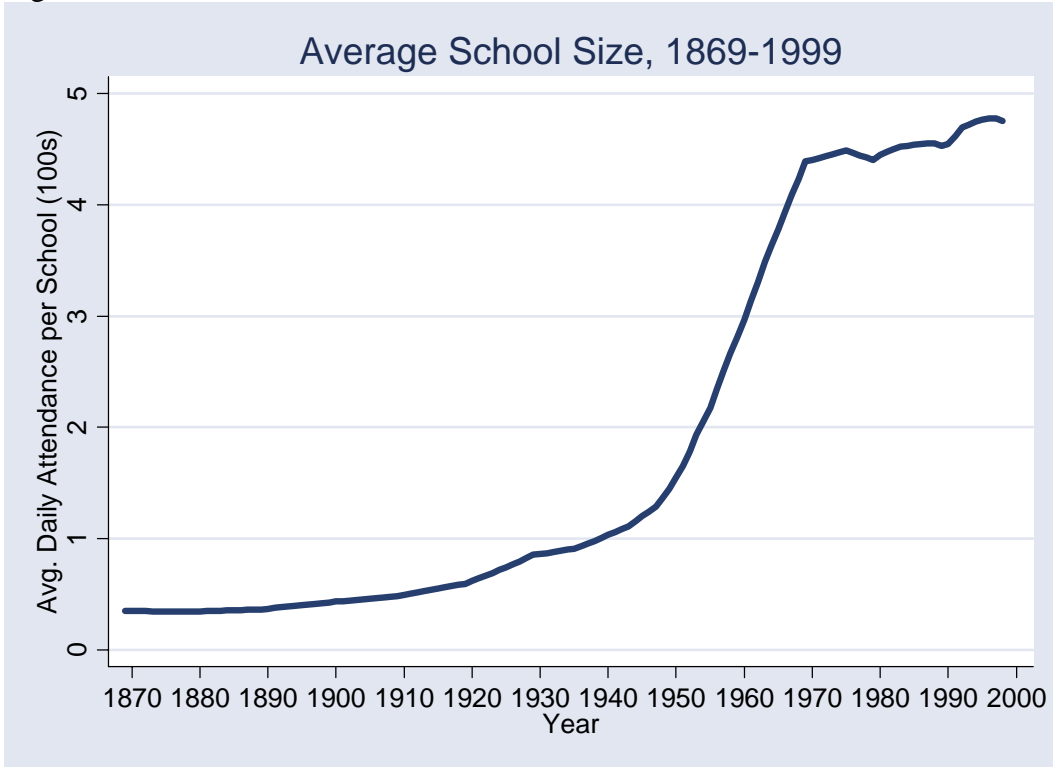


Figure 4

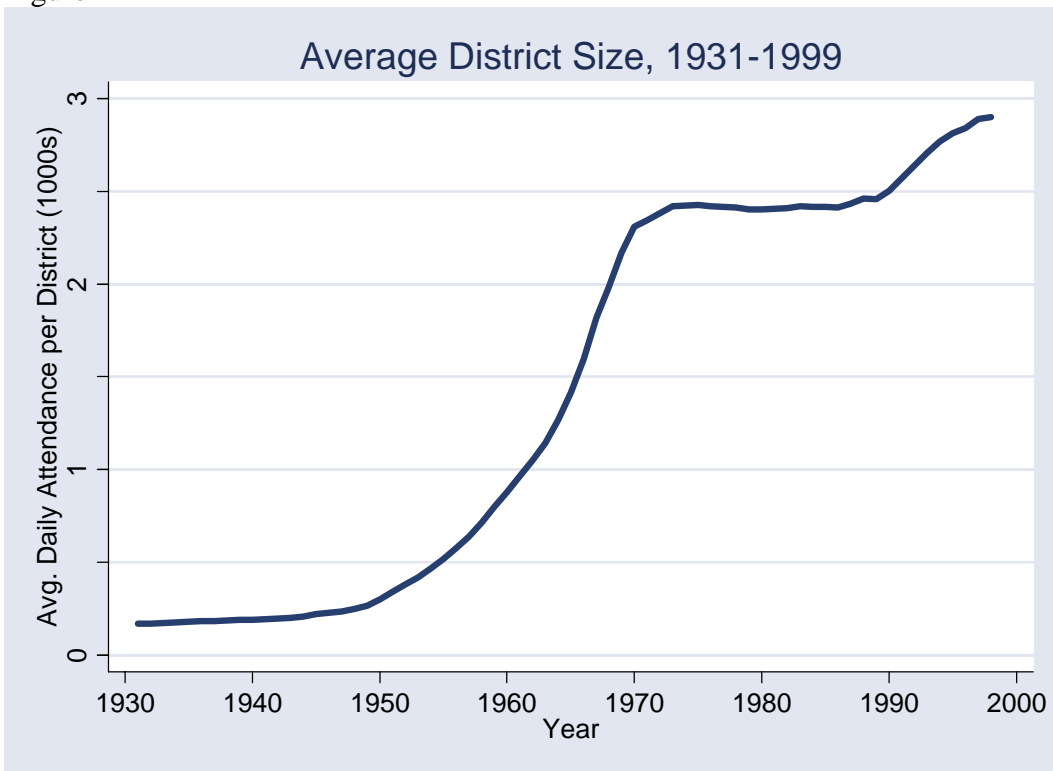


Figure 5

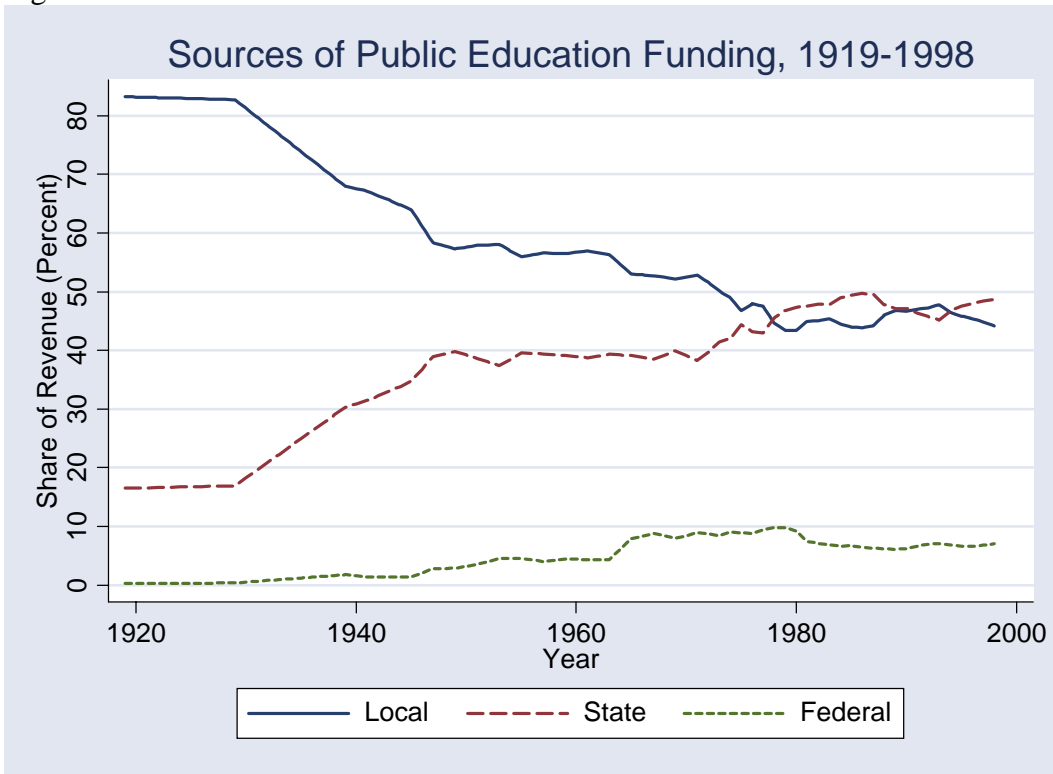


Table 1: Spearman Rank Correlations between School Characteristics, by Birth Cohort*

	Pupil- teacher ratio	Term length	Teacher salary	District size (ADA)	School size (ADA)
1920-29					
Term length	-0.381 <i>0.007</i>				
Teacher salary	0.034 <i>0.817</i>	0.632 <i>0.000</i>			
District size (ADA)	0.626 <i>0.000</i>	-0.026 <i>0.858</i>	0.381 <i>0.007</i>		
School size (ADA)	0.593 <i>0.000</i>	0.246 <i>0.089</i>	0.696 <i>0.000</i>	0.743 <i>0.000</i>	
State share of funding	0.714 <i>0.000</i>	-0.353 <i>0.013</i>	-0.044 <i>0.762</i>	0.486 <i>0.000</i>	0.368 <i>0.008</i>
1930-39					
Term length	-0.274 <i>0.056</i>				
Teacher salary	0.047 <i>0.748</i>	0.520 <i>0.000</i>			
District size (ADA)	0.664 <i>0.000</i>	0.051 <i>0.726</i>	0.363 <i>0.010</i>		
School size (ADA)	0.491 <i>0.000</i>	0.356 <i>0.012</i>	0.764 <i>0.000</i>	0.715 <i>0.000</i>	
State share of funding	0.771 <i>0.000</i>	-0.247 <i>0.087</i>	-0.054 <i>0.712</i>	0.554 <i>0.000</i>	0.316 <i>0.024</i>
1940-49					
Term length	-0.165 <i>0.256</i>				
Teacher salary	0.059 <i>0.688</i>	0.396 <i>0.005</i>			
District size (ADA)	0.713 <i>0.000</i>	0.183 <i>0.208</i>	0.418 <i>0.003</i>		
School size (ADA)	0.380 <i>0.007</i>	0.396 <i>0.005</i>	0.780 <i>0.000</i>	0.677 <i>0.000</i>	
State share of funding	0.658 <i>0.000</i>	0.037 <i>0.802</i>	0.099 <i>0.498</i>	0.617 <i>0.000</i>	0.320 <i>0.022</i>

*Numbers in italics are *p*-values. ADA = average daily attendance.

Table 2: Determinants of the Return to Education (GLS), Dependent Variable: Percentage Return to Education

Independent Variable	Base Models					With Quadratic Terms		Without DC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pupil-teacher ratio/100	-7.368 (2.507)***	-7.449 (2.507)***	-8.110 (2.436)***	-7.269 (2.502)***	-8.02 (2.43)***	-8.028 (2.355)***	-7.319 (2.354)***	-7.487 (2.484)***	-7.054 (2.371)***
Term length (hundreds of days)	0.320 (0.779)	0.399 (0.787)	0.221 (0.754)	0.227 (0.783)	0.14 (0.77)	-0.060 (0.736)	0.106 (0.740)	0.003 (0.905)	0.301 (0.744)
Relative teacher wage	0.788 (0.360)**	0.753 (0.364)**	0.567 (0.355)	0.779 (0.359)**	0.56 (0.36)	-0.005 (0.378)	0.261 (0.355)	0.261 (0.377)	0.196 (0.360)
District size (1000s)		-0.015 (0.024)			-0.00 (0.02)	0.109 (0.037)***	0.098 (0.037)***	0.116 (0.080)	0.074 (0.032)**
School size (100s)			-0.208 (0.065)***		-0.21 (0.07)***	-0.017 (0.171)	-0.313 (0.072)***	-0.321 (0.075)***	-0.306 (0.071)***
State share of funding				0.006 (0.006)	0.01 (0.01)	0.001 (0.006)	0.002 (0.006)	0.001 (0.006)	0.002 (0.006)
District size squared						-0.001 (0.000)***	-0.001 (0.000)***	-0.002 (0.003)	
School size squared						-0.047 (0.025)*			
Dummy for born 1930-39	0.905 (0.073)***	0.905 (0.073)***	0.984 (0.075)***	0.852 (0.092)***	0.93 (0.09)***	0.951 (0.096)***	1.007 (0.092)***	1.018 (0.093)***	1.017 (0.093)***
Dummy for born 1940-49	1.854 (0.098)***	1.865 (0.099)***	2.150 (0.132)***	1.802 (0.112)***	2.10 (0.14)***	2.096 (0.160)***	2.236 (0.144)***	2.243 (0.145)***	2.247 (0.145)***
Constant	5.619 (1.700)***	6.994 (2.618)***	7.287 (1.667)***	5.512 (1.696)***	7.46 (2.54)***	7.916 (2.420)***	7.773 (2.449)***	6.797 (1.871)***	6.244 (1.609)***
Observations	147	147	147	147	147	147	147	144	144
R-squared	0.95	0.95	0.96	0.95	0.96	0.96	0.96	0.96	0.96

Notes: Standard errors in parentheses. All models include state-of-birth fixed effects. Equations are weighted by the inverse sampling variance of the dependent variable. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Spearman Rank Correlations between School Characteristics and Population Characteristics*

	Birth Cohort					
	1920-29		1930-39		1940-49	
	Real per capita income of parents' generation	Percent rural	Real per capita income of parents' generation	Percent rural	Real per capita income of parents' generation	Percent rural
Percent Rural	-0.863		-0.849		-0.631	
	<i>0.000</i>		<i>0.000</i>		<i>0.000</i>	
Pupil-teacher ratio	-0.348	0.068	-0.339	0.177	-0.500	0.111
	<i>0.014</i>	<i>0.643</i>	<i>0.017</i>	<i>0.224</i>	<i>0.000</i>	<i>0.449</i>
District Size (ADA)	0.070	-0.294	0.059	-0.207	-0.145	-0.235
	<i>0.631</i>	<i>0.037</i>	<i>0.689</i>	<i>0.145</i>	<i>0.311</i>	<i>0.097</i>
School Size (ADA)	0.380	-0.610	0.500	-0.624	0.349	-0.720
	<i>0.007</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.012</i>	<i>0.000</i>

*Numbers in italics are *p*-values. ADA = average daily attendance. Correlations between pupil-teacher ratios, school size, and district size are shown in Table 1.

Table 4: Population Characteristics and the Return to Education (GLS)

	(1)	(2)	(3)	(4)
Pupil-teacher ratio/100	-6.358 (2.366)***	-7.070 (2.370)***	-6.413 (2.458)***	-5.894 (2.442)**
Term length (hundreds of days)	-0.061 (0.754)	0.361 (0.742)	0.426 (0.743)	0.064 (0.768)
Relative teacher wage	0.670 (0.427)	0.295 (0.430)	0.254 (0.364)	0.769 (0.480)
District size (1000s)	0.062 (0.032)**	0.076 (0.031)**	0.077 (0.031)**	0.065 (0.032)**
School size (100s)	-0.287 (0.070)***	-0.301 (0.072)***	-0.280 (0.076)***	-0.261 (0.077)***
Real per capita income of parents' generation (\$1000s)	-0.790 (0.394)**			-0.743 (0.397)*
Rural population (percent)		0.396 (0.914)		0.315 (0.902)
Percent of population attending school			-4.392 (4.508)	-3.383 (4.491)
Dummy for born 1930-39	1.159 (0.096)***	1.031 (0.074)***	0.990 (0.086)***	1.116 (0.110)***
Dummy for born 1940-49	2.779 (0.288)***	2.284 (0.139)***	2.258 (0.132)***	2.759 (0.291)***
Constant	7.107 (1.711)***	5.854 (1.833)***	6.354 (1.677)***	7.258 (1.923)***
Observations	144	144	144	144
R-squared	0.96	0.96	0.96	0.96

Table A1: Summary Statistics

Cohort	Estimated Returns to Education	Pupil-Teacher Ratio	Term Length (Days)	Relative Teacher Wage	District Size (ADA)	School Size (ADA)	State Share of Funding	Real Income of Parents' Generation	Percent Rural
1920-29									
Mean	4.98	28.30	172	0.94	2,871	116	26.5	\$804	52.9
Median	4.96	28.20	175	0.89	258	92	23.1	\$750	59.7
SD	0.64	4.85	10	0.24	11,257	84	18.2	\$327	21.2
Min	3.05	17.70	139	0.65	34	26	1.1	\$291	0.0
Max	6.89	37.60	187	1.85	78,767	495	90.8	\$1,809	83.4
1930-39									
Mean	5.97	26.29	175	0.97	3,292	156	36.7	\$953	51.6
Median	5.95	27.00	176	0.96	506	142	36.5	\$898	57.3
SD	0.55	3.87	6	0.17	12,114	97	21.2	\$362	19.6
Min	4.71	16.60	154	0.66	33	29	2.1	\$370	0.0
Max	7.29	34.40	187	1.54	84,620	528	89.8	\$2,026	80.2
1940-49									
Mean	7.00	25.00	178	0.98	4,353	250	37.6	\$1,438	43.5
Median	7.10	25.60	178	0.96	1,248	239	34.4	\$1,437	45.2
SD	0.55	2.99	3	0.12	14,142	121	18.5	\$333	17.1
Min	5.73	17.60	168	0.76	40	42	3.5	\$770	0.0
Max	7.89	31.40	184	1.28	98,376	537	84.9	\$2,228	73.4

Notes: Income of parents' generation is expressed in 1950 dollars. ADA = average daily attendance. When Washington, D.C. is excluded, mean district sizes for the three cohorts are 1,290, 1,598, and 2,394, respectively.