

Social Stratification: Economic Class and Education

Author: Claire Alley

University of Essex

Abstract

Income and wealth have an impact on the academic performance of children, and this effect extends to elementary education. This study used the CASchools data frame to analyse how the stratification of a region's income affects the reading scores of elementary school children. Results showed that children within low-income regions perform worse on standardized tests. Additionally, school expenditure was found to be a confounding variable to income and did not have an impactful relationship with reading scores.

Keywords: Stratification, Child Education, Income, Societal Class

Date of Submission: 15.05.2026

Date of Acceptance: 17.08.2026

Introduction

This analysis will examine the relationship between economic class and a student's reading performance using the CASchool data frame. The data was collected from 420 school districts within California from 1998 to 1999 and details the average household income of school districts within California and 5th grade students' results on a standardised reading test (Stock and Watson, 2007). The data also includes additional variables such as expenditure per student, number of students who qualify for CalWORKs, and percentage of students who are learning English as an additional language. CalWORKs is a socioeconomic support program that provides financial assistance to children and families with children who experience extreme financial hardships (Danielson et al., 2020). The variables contained in this data frame will be used to further express the connection between income and education. This dataset was chosen because it provided a variable that represented a child's academic performance and a variable that represented the average income of a community, in the form of school districts, rather than individual families. Additionally, this dataset contained variables that represented additional factors that influence a child's academic performance, such as expenditure and English as an additional language learners.

For this analysis, economic class will be defined by a district's average income to indicate the financial resources available to individual communities and will be separated between high and

low classes, where the high class has a higher average gross income than the lower class. The analysis of this data hopes to illustrate the relationship between income and education, and advocate for lower income regions to receive additional support so there is less inequality in academic performance between high-income and low-income regions.

Previous research into wealth's impact on children's education have utilized the theory of socioeconomic status as a model to describe their relationship (Lareau and Weininger, 2003; Sirin, 2005; Tramonte and Willms, 2009). This research has also primarily defined economic class by a child's household income and have categorized children as individually high-class or low-class by this distinction (Cooper and Stewart, 2020; Ermisch et al., 2001). Research with this definition of class has found that low-class students perform worse than high-class students in terms of academic success (Cooper and Stewart, 2020; Reardon, 2014). Additionally, Reardon (2014) states that the gap in performance on standardized reading tests between high and low-income students has grown significantly between the 1950s to the 2000s. The 2000s to the early 2010s also reached historic highs for economic inequality (Saez, 2012). Therefore, researching the impact income inequality has on education is important to learn to mediate its effects on children's education and their educational opportunities.

Moreover, there is research that has supporting evidence that one's socioeconomic status and the socioeconomic status of their community affect one's academic success (Farkas, 2003; Hung et al., 2019). To provide supporting research to that theory, this analysis has defined children as members of school districts which are either high-class or low-class depending on the district's average income of its inhabitants. This provides a differing viewpoint than previous research because instead of individual families determining the class of their child, this analysis makes community as the determining factor of a child's class.

This distinction is important because it represents the overall economic class of the school district that provides their education and the community that surrounds the child. As economic inequality becomes more prevalent in society, it becomes more important to understand its effects on communities and on its members.

Theoretical Model

For the duration of this analysis, the theory of social stratification will be employed to describe the analytical findings. McLeod (2012, p. 229) defines social stratification as "the unequal distribution of valued resources across social groups" where 'social groups' refer to groups of people in society that are distinguished based on their race, income, education and any other factors, and "dominant groups assert and maintain control over valued social resources" where 'valued resources' are defined as being both tangible and intangible.

Platt et al. (2026) defines social stratification similarly as an unequal distribution of resources, opportunities, and rewards by means of social processes that do not reflect differences in people's innate abilities. In the definition from Platt et al. (2026), not only is the unequal

distribution of resources a defining feature of social stratification, but the social process and the reproduction of those processes create stratification too.

Grusky (2019) details these resources further by sorting them into seven categories which are economic, political, cultural, social, honorific, civil, and human. However, we will primarily use the first and seventh categories because they represent the explanatory variable, average income and the outcome variable, academic performance. The first category, 'Economic', describes the ownership of production and income (Grusky, 2019). Grusky (2019) titles the seventh category 'Human', which includes resources related to human skills and knowledge which includes education. Using the theory of stratification, we will examine the CASchool's data on reading scores and observe how inequality in income affects reading performance. Additionally, this framework will be utilized to inspect how the other variables may influence academic performance, and if they can provide any additional information on stratification's effect on education.

These concepts of resources also provide a basis for the definition of class for this analysis. Vallet (2025) describes class as economic or social groups that are similar regarding certain criteria. Additionally, Vallet (2025) and Sørensen (2000) describe Max Weber's theory of class as defined by one's economic standing within society, and how one's standing affects one's opportunities. This definition of class and the description of resources from Grusky (2019) provide a basis for this analysis' understanding of class. Class, for this analysis' purpose, is defined by one's possession of economic resources, in this case income, and how differing groups of people can be distinguished based on the number of economic resources they possess, the high and low class. This explanation of class will be the basis for the theoretical model of social stratification.

The confirmatory research of this analysis pertains to the positive correlation between the outcome variable, reading scores, and explanatory variable, income. Whilst the income of a district increases, the reading scores of students will increase. The control variables are also expected to influence reading scores. Expenditure, for instance, is expected to have a similar correlation to reading scores as income, but expenditure's impact should not be as great as income's impact because this variable is less affected by stratification. The control variables of students with CalWORKs and students of English, are expected to have a negative correlation with reading scores. CalWORKs is anticipated to have a similar correlation as income, because CalWORKs represents those within the lower class, while income represents the overall class of the district. The value of this distinction is described in the data section of this analysis.

Data

Table 1 - Description of Variables in the Analysis

Explanatory Variable	Outcome Variable
Average Income	Average Reading Score

<i>Numerical Variable</i>	<i>Numerical Variable</i>
Variable Name	Variable Name
income	Read
Variable Measurement	Variable Measurement
The average income of the surrounding district in thousands of USD (\$).	The average reading score of a student.

Control Variables
1. Average Expenditure per Student (referred to as expenditure).
A numerical variable that details the average expenditure per student.
2. Percentage of Qualifiers for CalWORKs (referred to as CalWORKs).
A numerical variable that shows the percentage of students who qualify for CalWORKs, an income assistance program.
3. Percentage of English Learners (referred to as English learners).
A numerical variable that expresses the percentage of students who learn English as a second language.

CASchools, a dataset containing information about different districts within California, is the source for this analysis. The outcome variable is reading scores which are numerical values that represent the average reading score of the district's students. This variable is the outcome variable because this analysis aims to measure the impact stratification has on education and reading scores are a dependent variable that accurately represent a student's education.

The main explanatory variable is income which encompasses the average income of the district. This variable was chosen as the main explanatory variable because it represents the average economic class of a student, which is the main social group distinction this analysis aims to observe. This variable is also representative of the first category, 'Economic', described by Grusky (2019) which allows this analysis to view how stratification of income influences education.

The three control variables included within the investigation are the average expenditure per student, the percent of students qualifying for CalWORKs, and the percent of students learning English as a second language. Expenditure has been chosen to illustrate the economic class of the academic institution and to further observe how this expenditure may affect education. However, this variable is expected to be less impactful than income because expenditure is allocated evenly

between students which does not allow for individual students to control more monetary resources than another student and prevents stratification. CalWORKs has been incorporated because it depicts students with financial difficulties that make them a member of the low class. Children qualify for CalWORKs when their annual family income is below 78% of the poverty line (around \$16,900) in California (Danielson et al., 2020). Therefore, children who qualify for CalWORKs are in financial distress, making them members of the lower-class. This variable is important because it purely represents the low class, because members of the high class do not qualify for CalWORKs. This variable is not identical to income because income portrays the average class of the district and does not provide further knowledge into the low or high class. Moreover, income views the population as a homogenous high or low-class society and does not provide the nuance to view individuals as either class. Therefore, CalWORKs is necessary to depict the low class because the main explanatory variable is unable to do so.

The final control variable, English learners, accounts for students who are not at the expected level of English proficiency as their peers. This variable is important because it provides insight into which social groups have limited English knowledge, and how this knowledge interacts with stratification.

Results

Table 2 - Mean and Median of Variables

	Mean	Median
Reading Scores	654.9705	655.75
Income	15.31659	13.7278
Expenditure	5312.408	5214.517
CalWORKs Qualifiers	13.24604	10.52045
English Learners	15.76816	8.777634

Table 2 shows the means and medians of the variables which provides insight into which variables have outlying data that could skew results.

To begin the analysis of these variables, comparing the mean and median allows the identification of skewed averages amongst the variables. Reading scores, for example, have little to no variation between the mean and the median which conveys a lack of outliers and a lack of a skewed variable. The mean and median income, however, have a noticeable difference. The mean is greater than the median which implies the existence of high value outliers which skew the mean

income to be \$2,000 more than the median income. The inaccurately high income mean represents those with high income more which results in lower earning districts being underrepresented within further analysis. Fortunately, this is the reason for including CalWORKs, because this variable solely represents the low class, they will be represented despite the wavering mean.

As for the control variable, expenditure, there is little variation which implies that there are no outliers to skew the mean. The other control variables, CalWORKs and English learners, are skewed in favour of the mean due to the variables possessing high outliers. This implies that there is a higher concentration of CalWORKs qualifiers and English learners within either the high or low class. CalWORKs undoubtedly has a higher concentration within low class, lower income areas because only students from those backgrounds can qualify. Which class the high concentration of English learners reside will be discussed when further information is examined.

Table 3 - Standard Deviation of Variables

	Standard Deviation
Reading Scores	20.10798
Income	7.22589
Expenditure	633.9371
CalWORKs Qualifiers	11.45482
English Learners	18.28593

Table 3 shows the standard deviation of the variables which expresses how spread-out the data is from the mean. This provides insight into the variety of values within a variable.

Secondly, the observation of the standard deviation illustrates the range of values represented by a variable. The standard deviation of reading scores is small compared to the range of reading scores

acquired by the study (600-700). Income has quite a large deviation of an average of \$7,000 which means that the variability of income between districts ranges significantly. This large range, again, supports the inclusion of CalWORKs, as using an average for a wide range of values causes values on either end of the range to be ignored. The control variables also all have a large standard deviation, meaning that the variables range widely depending on the specific district.

Table 4 – Correlation of Variables

	Reading Scores	Income	Expenditure	CalWORKs Qualifiers	English Learners
Reading Scores	1.0000000	0.6978189	0.21792682	-0.61184714	-0.69028587
Income	0.6978189	1.0000000	0.31448448	-0.51265102	-0.30741949
Expenditure	0.21792682	0.31448448	1.0000000	0.06788857	-0.07139604
CalWORKs Qualifiers	-0.61184714	-0.51265102	0.06788857	1.0000000	0.31957593
English Learners	-0.69028587	-0.30741949	-0.07139604	0.31957593	1.0000000

Table 4 shows the correlations between variables which describes the relationships as either positive or negative, and expressing the strength of their correlation. The closer the value is to 1 or -1, the stronger the correlation.

Thirdly, the utilization of a correlation matrix will allow for the inspection of correlation between the variables. Starting with reading scores, the variable possesses a positive correlation with income and expenditure. More specifically, income retains a higher correlation to reading scores than expenditure does, which signifies that income has a greater influence on reading scores than expenditure. However, expenditure and reading scores correlation score is not negligible and still suggests a link between the two. Conversely, CalWORKs and English learners' correlation to reading scores are considerably negative. This suggests that as the percentage of CalWORKs and English learning students increases, the reading scores decrease. The correlation scores are also similar in strength to income which implies that, although income is an explanatory variable, it is not the only variable that influences reading scores. Additionally, with this new revelation, the question of which class has the higher concentration of English learners can be answered. As a

result of CalWORKs and English learners sharing a negative correlation with reading scores, the two variables represent two levels of stratification. A student's reading score is affected by their economic class and their linguistic proficiency in English. This aligns with this analysis' theory because social stratification cannot solely be described by a single variable but a series of multiple factors that impact one's access to resources. In this case, district income and English proficiency both have influence on reading scores and suggests that economic class is one factor among many that may influence reading scores.

When observing income and CalWORKs, the correlation is negative. As income increases the percentage of CalWORKs qualifiers decreases. This is rational because those living in economically wealthy areas will also be wealthy, on average, and will not qualify for CalWORKs. The percentage of English learners has a positive correlation with CalWORKs which implies that as the percentage of low-income families increases, the number of English learners also increases. This statistic shows a relationship between lower income areas and people with non-English backgrounds. This significance is evidenced by the negative correlation between English learners and income. As the surrounding districts' average income decreases, the number of non-English speakers increases.

Table 5 – P-Values of Variables

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept (Reading Scores)	625.23 (1.65) ***	640.77 (2.25) ***	641.22 (1.43) ***	625.49 (5.96) ***	639.52 (3.94) ***
Income	1.94 (0.10) ***	1.45 (0.10) ***	1.49 (0.07) ***	1.94 (0.10) ***	1.09 (0.08) ***
Expenditure				0.00 (0.00)	0.00 (0.00) **
CalWORKs Qualifiers		-0.61 (0.07) ***			-0.46 (0.05) ***
English Learners			-0.58 (0.03) ***		-0.53 (0.03) ***

*Table 5 shows the p-values of the variables which represents how significant and impactful the variables are to the analysis. Variables denoted with (***) have a high significance. Variables denoted with (**) have a medium significance. Variables with no denotation have no significance.*

Finally, investigating the regression and p-values of the variables will provide insight into the variables significance to reading score. The p-value of income is significant when alone with reading score and continues to be significant when paired with any of the other variables. Likewise, CalWORKs and English Learners also remain significant regardless of the variables included. Conversely, in model 5, expenditure receives a lower significance when compared to the other variables. Furthermore, in model 4 which pairs income and expenditure, expenditure is not significant. This implies that expenditure is a confounder and does not impactfully influence reading scores, instead income affects expenditure which generates an illusion of expenditure's impact on reading scores. This is a result of income influencing expenditure which causes the relationship between expenditure and reading scores to mirror the relationship between income and reading scores. The p-values when expenditure and reading scores are compared, expenditure has a lower, less impactful p-value. However, when expenditure is compared alongside income and reading scores, expenditure receives a p-value of 0 which indicates no impact on reading scores. Income, however, receives a high p-value which indicates a strong impact on reading scores. This interaction supports the claim that expenditure is a confounding variable between income and reading scores.

Discussion

The original hypothesis of this analysis is that, according to stratification, income would have a strong positive correlation and association with reading scores. This analysis has offered supportive evidence to this theory: the positive correlation between income and reading scores, and the significant p-value score of income when compared to reading scores strongly suggest an association between district income and reading scores which is consistent with the broader literature. Additionally, this paper analysed other potential variables that represent other factors of stratification. The percentage of English learners increased alongside average income which suggests that low-income areas contain more English learners and proposes an additional factor through which stratification impacts education. However, there is more research to be done to identify who 'English learners' describes and how non-economic stratification may influence this data. Furthermore, the variable expenditure was discovered to be a confounder of the explanatory variable income. If further analysis is to be carried out, this variable should be removed for a clearer analysis.

This analysis acts as a resource to further support the evidence of stratification within education, and the impact of district income on educational performance. However, this analysis had several restrictions, and further research is needed. This analysis describes the relationship between income and educational performance as an association and does not indicate that income directly causes educational performance. This analysis suggests that district income impacts

educational performance, but this analysis cannot provide evidence that income is the cause of educational performance. There may be other factors associated with high or low-income districts that impact educational performance more directly and future research should aim to discover these causal factors. Moreover, for future research, it would be more beneficial to separate income into high and low groups and observe their individual data relationships to have a better understanding on how economic class influences education. Additionally, future research should examine if the financial support a child receives from CalWORKs affects their academic performance. Previous studies into this topic have found some support and it would be beneficial to study this topic more (Clark-Kauffman et al., 2003; Duncan et al., 2011).

References

- Clark-Kauffman, E., Duncan, G. and Morris, P. (2003) 'How welfare policies affect child and adolescent achievement', *American Economic Review*, 93(2), pp. 299-303.
- Cooper, K. and Stewart, K. (2021) 'Does household income affect children's outcomes? A systematic review of the evidence', *Child Indicators Research*, 14, pp. 981-1005.
- Danielson, C., Thorman, T. and Bohn, S. (2020) *The importance of CalFresh and CalWORKs in children's early years*. San Francisco: Public Policy Institute of California. Available at: [The Importance of CalFresh and CalWORKs in Children's Early Years - Public Policy Institute of California](#)
- Duncan, G., Morris, P. and Rodrigues, C. (2011) 'Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments', *Developmental Psychology*, 47(5), pp. 1263-1279.
- Ermisch, J., Francesconi, M. and Pevalin, D. J. (2001) *Outcomes for children of poverty*. Department for Work and Pensions research report no. 158. Leeds: Corporate Document Services.
- Farkas, G. (2003) 'Cognitive skills and noncognitive traits and behaviors in stratification processes', *Annual Review of Sociology*, 29, pp. 541-562.
- Grusky, D. (2019) *Social stratification, class, race, and gender in Sociological Perspective*. 2nd edn. Routledge.
- Hung, M., et al. (2019) 'Exploring student achievement gaps in school districts across the United States', *Education and Urban Society*, 52(2), pp. 175-193.
- Lareau, A. and Weininger, E. (2003) 'Cultural capital in educational research: a critical assessment', *Theory and Society*, 32, pp. 567-606.

McLeod, J.D. (2012) 'Social stratification and inequality', In: aneshensel, C.S., Phelan, J.C., Bierman, A. (ed.) *Handbook of the Sociology of Mental Health. Handbooks of Sociology and Social Research*. Springer, Dordrecht. Available at: https://doi.org/10.1007/978-94-007-4276-5_12

Platt, L., Polavieja, J. and Werfhorst, H. (2026) 'Social stratification research: achievements, challenges, and the road ahead', in L. Platt, J.G. Polavieja, H.G. Werfhorst and M. Gangl (ed.), *The Oxford Handbook of Social Stratification*. Online edn. Oxford University Press. Available at: <https://doi.org/10.1093/oxfordhb/9780197539484.013.1>

Reardon, S. (2014) 'The widening income achievement gap', in R. Arum, I.R. Beattie, and K. Ford (ed.), *The Structure of Schooling: Readings in the Sociology of Education*. 3rd edn. SAGE Publications. pp. 288-294.

Saez, E. (2015) 'Striking it richer: the evolution of the top incomes in the United States', *University of California, Berkeley*. Available at: <https://eml.berkeley.edu/~saez/saez-UStopincomes-2014.pdf>

Sirin, S. (2005) 'Socioeconomic status and academic achievement: a meta-analytic review of research', *Review of Educational Research*, 75(3), pp. 417-453.

Sørensen, A. (2000) 'Toward a sounder basis for class analysis', *American Journal of Sociology*, 105(6), pp. 1523-1558.

Stock, J. H. and Watson, M. W. (2007) *Introduction to Econometrics*. 2nd edn. Prentiss Hall.

Tramonte, L. and Willms, J.D. (2009) 'Cultural capital and its effects on education outcomes', *Economics of Education Review*, 29(2), pp. 200-213.

Vallet, L. (2025) 'Class', in L. Platt, J.G. Polavieja, H.G. Werfhorst and M. Gangl (ed.), *The Oxford Handbook of Social Stratification*. Online edn. Oxford University Press. Available at: <https://doi.org/10.1093/oxfordhb/9780197539484.013.5>

Copyright Statement

©Claire Alley. This article is licensed under a [Creative Commons Attribution 4.0 International Licence \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).