Understanding the Influential Factors to Achievement Test Scores

for Students in New Jersey

Asuncion Los Banos

Saint Peter’s University

Abstract

Socio-economic factors and its influence on standardized test scores has been argued and multiple studies have been completed. This capstone examines the relationship between the different types of average income of household types: household, family, married couple, and non family and the previous achievement test the State of New Jersey administered to its high school students the High School Proficiency Assessment (HSPA). The data for average income by household and race was retrieved from the American FactFinder - U.S. Census Bureau. A single linear regression model of average test scores and to the average household income will be used to establish a positively correlated relationship. The average test scores will then be further analyzed by completing a regression by adding race: African American, Asian, Caucasian, and Hispanic or Latino. To rule out any consistencies the household type and race may have, the data will be from 2014, 2010, and 2007. The single linear regression for all three years for average household income showed to be statistically significant, whereas for the race only Asian was not statistically significant.

Understanding the Influential Factors to Achievement Test Scores

for Students in New Jersey

Do students who score high on standardized tests naturally good test takers or are there other factors that take place? Popham (1999) states that “standardized test is any examination that's administered and scored in a predetermined, standard manner”. Students in United States will likely take two major kinds: aptitude and achievements tests. For college bound high school students they will partake in an aptitude test that will predict how well they are likely to perform in educational settings. Those aptitude tests are the SAT and ACT, which forecast how well a high school student will perform in college (Popham, 1999). Students while in elementary, middle, and high school will be exposed to the other type of standardized test, the achievement test. These score results are what citizens and school board members rely on to base on how well the school is performing (Popham, 1999). Understanding what factors that can influence a student’s test score in an achievement test will help school districts become more efficient.

**Work by Competitors**

In Dahl and Lochner’s (2012) paper titled “The impact of family income on child achievement: Evidence from the earned income tax credit” a strategy of instrumental variables was used. They estimated the causal effect of income on children’s math and reading achievement. Their identification derived from large, nonlinear changes in the Earned Income Tax Credit. From their paper, they found “the largest of these changes increased family income by as much as 20 percent, or approximately $2,100, between 1993 and 1997. Our baseline estimates imply that a $1,000 increase in income raises combined math and reading test scores by 6 percent of a standard deviation in the short run. Test gains are larger for children from disadvantaged families and robust to a variety of alternative specifications.” Their estimation strategy was based on the fact that low to middle income families benefitted substantially from expansions of the EITC in the late-1980s and mid-1990s while higher-income families did.

Anger and Heineck’s (2010) paper examined the intergenerational transmission of cognitive abilities. Their finding was not compatible to a pure genetic model, but rather point to the importance of parental investments for children’s cognitive outcomes. They found that “individuals’ cognitive skills are positively related to their parents’ abilities, despite controlling for educational attainment and family background”.

In the paper titled “The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood” by Chetty, Friedman, and Rockoff (2011) measured if teacher’s impacts on students test scores a good measure of their quality. They analyzed school district data from grades 3 to 8 and linked the student’s test scores to tax records on parent characteristics and adult outcomes. Their paper found that “students assigned to high-value added teachers are more likely to attend college, attend higher- ranked colleges, earn higher salaries, live in higher SES neighborhoods, and save more for retirement.”

The paper by Matthew M. Chingos (2012) titled “Strength in Numbers: State Spending on K-12 Assessment Systems” provides the most current, comprehensive evidence on state-level costs of assessment systems. I felt this paper was important because achievement tests are used to evaluate how well a school is performing by the public and law makers.

The paper (2015) titled “What high-achieving low-income students know about college” is interesting because there are studies on the education gap and that a student from a low-income family does not perform well in school. This paper “provides individualized information about colleges' net prices, resources, curricula, students, and outcomes. Our prior study shows that the intervention raises students' applications to, admissions at, enrollment, and progress at selective colleges.” A survey data was used to show changes in students’ knowledge and decision-making.

Mayer (2010) focuses on the effect of parent’s income and isolated the effect of parental income on children’s outcomes, in particular the effect of low parental income on poverty. The author concluded that “evidence suggested that a 10 percent increase in parental income was associated with .024 to .104 additional years of schooling.”

In the paper (1999) “Why Standardized Tests Don't Measure Educational Quality” there is an argument of why standardized test should not be used as a measure for a student’s ability. The paper states, “The task for those developing standardized achievement tests is to create an assessment instrument that, with a handful of items, yields valid norm-referenced interpretations of a student's status regarding a substantial chunk of content. Items that do the best job of discriminating among students are those answered correctly.”

The paper (1990) “Income Level, Gender, Ethnicity, and Household Composition as Predictors of Children's School‐based Competence” they state that “In the United States, being black, male, or growing up in a low-income and/or single-parent household have all been identified as risk factors for maladjustment during childhood.” The authors compared predictions of 3 different forms of children's competence from each of these 4 variables. They used a sample of 868 black and white elementary school children from 2-parent and mother-headed 1-parent homes, studied 3 aspects of school-based competence: conduct, peer relations, and academic achievement. The result was “overall, income level and gender were thus the strongest predictors of children's competence.”

The paper (2005) “The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment” examined the process of how socioeconomic status, specifically parents' education and income, indirectly relates to children's academic achievement through parents' beliefs and behaviors. The data came from a national, cross-sectional study of children were used for this study. The subjects were 868 8-12-year-olds, divided approximately equally across gender (436 females, 433 males). This sample was 49% non-Hispanic European American and 47% African American. A structural equation modeling techniques, the author found that the socioeconomic factors were related indirectly to children's academic achievement through parents' beliefs and behaviors but that the process of these relations was different by racial group.

The authors of “Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal” (1996) examined differences in intelligence test scores of black and white 5-year-olds. The data from Infant Health and Development Program included 483 low birthweight premature children who were assessed with the Wechsler Preschool and Primary Scale of Intelligence. The children had been followed from birth, with data on neighborhood and family poverty, family structure, family resources, maternal characteristics, and home environment collected over the first 5 years of life. The paper found that “Black children's IQ scores were 1 SD lower than those of white children. Adjustments for ethnic differences in poverty reduced the ethnic differential by 52%”.

**Data**

The years for this capstone will be 2014, 2010, and 2007. Since I am searching for an influential factor on average test scores I decided to pick 2007 and 2010 because it was the start and end of the housing bubble and 2014 to view if there were any changes from the end of the bubble and four years later. The test scores will be from the 2014, 2010, and 2007 HSPA reading and math results. The data for test scores were retrieved from the State of New Jersey’s Department of Education website and these results will be my dependent variables. The dependent variables will be the test scores from the 21 counties in New Jersey. The HSPA scores for reading and math had a range of 100 to 300. Where a score of 199 and below indicated partially proficient, 200 to 249 indicated proficient and 250 to 300 indicated advanced proficient. For the purpose of this capstone the average score from the two subject areas will be used. An example of average scores from the HSPA 2014 is shown in shown in Table 1 below.

Table 1

|  |  |
| --- | --- |
| Average Combined Test Scores for 2014 by County | |
| County | Average Combined Test Scores |
| Atlantic | 225 |
| Bergen | 236 |
| Burlington | 228 |
| Camden | 221 |
| Cape may | 227 |
| Cumberland | 213 |
| Essex | 222 |
| Gloucester | 228 |
| Hudson | 224 |
| Hunterdon | 244 |
| Mercer | 229 |
| Middlesex | 231 |
| Monmouth | 236 |
| Morris | 242 |
| Ocean | 232 |
| Passaic | 216 |
| Salem | 221 |
| Somerset | 237 |
| Sussex | 232 |
| Union | 229 |
| Warren | 231 |

The independent variables of household income and ethnicity will be gathered from the U.S. Census Bureau’s American FactFinder. American FactFinder provides access to data about the United States and come from several censuses and surveys. I decided to utilize the Guided Search in the website to help search for the data I required. The constant filter for my search was the geographic type: County and New Jersey as the main state and all counties selected. The data for each year can be exported after the correct table has been selected in the website. Each independent variable can be easily located on the website using the ID column, shown in Table 2 below.

Table 2

|  |  |  |
| --- | --- | --- |
| Retrieving Data from U.S. Census Bureau’s American FactFinder | | |
| Label | Table, File or Document Title | ID |
| average\_income | Income In the Past 12 Months (In 2015 Inflation - Adjusted Dollars) | S1901 |
| race | RACE | DP05 |

The household incomes are divided into four parts: households, families, married couple, and non family households. A household includes all the people who occupy a housing unit as their usual place of residence. A family household includes a householder and one or more people living in the same household who are related to the householder by birth, married or adoption. All people in a household who are related to the householder are regarded as members of his or her family. A family household may contain people not related to the householder, but those people are not included as part of the householder’s family. Therefore, the number of family households is equal to the number of families, but family households may include more members than do families. Not all households contain families since a household may compromise a group of unrelated people or people living alone. A married couple, as defined by the census, is a husband and wife enumerated as members of the same household. The married couple may or may not have children living with them. The number of married couples equals the count of married couple families related and unrelated married couple subfamilies. A non family household may contain only one person, the householder, or additional persons who are not relatives of the householder. Non family households may be classified as either female non family or male non family households. An example of average income is shown in shown in Table 3 below.

Table 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Income by Household in 2014 | | | | |
| County | Household | Families | Married Couple | Non Family |
| Atlantic | $71,779 | $83,589 | $99,755 | $44,223 |
| Bergen | $116,079 | $135,531 | $152,585 | $63,593 |
| Burlington | $100,022 | $116,454 | $130,870 | $57,811 |
| Camden | $80,892 | $95,083 | $114,754 | $47,579 |
| Cape may | $77,949 | $92,068 | $102,358 | $49,053 |
| Cumberland | $65,299 | $73,627 | $90,031 | $40,608 |
| Essex | $87,496 | $106,813 | $144,627 | $47,760 |
| Gloucester | $91,614 | $105,523 | $119,050 | $51,125 |
| Hudson | $83,368 | $88,420 | $111,223 | $70,934 |
| Hunterdon | $136,413 | $156,246 | $176,079 | $78,898 |
| Mercer | $104,437 | $125,683 | $149,103 | $56,838 |
| Middlesex | $98,447 | $112,023 | $125,639 | $58,050 |
| Monmouth | $115,245 | $136,641 | $154,590 | $63,128 |
| Morris | $132,542 | $154,984 | $169,204 | $71,763 |
| Ocean | $79,508 | $93,658 | $101,888 | $47,026 |
| Passaic | $81,138 | $90,427 | $110,440 | $50,704 |
| Salem | $74,346 | $86,520 | $100,766 | $43,949 |
| Somerset | $134,595 | $155,890 | $172,571 | $73,580 |
| Sussex | $103,174 | $116,310 | $129,446 | $61,737 |
| Union | $99,818 | $115,257 | $140,195 | $54,896 |
| Warren | $87,399 | $100,443 | $112,533 | $53,733 |

To gather the data for race I had to use the Subject Hispanic or Latino and Race because individuals can either be Hispanic or Latino or the either types of race. As an example, in 2014 Atlantic County had a total population of 275,325. However, of this population only 48,865 was Hispanic or Latino and 226,460 was not Hispanic or Latino. The breakdown for the remaining race was 158,304 (Caucasian alone), 40,358 (Black or African American alone), and 22,037 (Asian alone). The remaining races that were not included in this capstone were: American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, some other race, and two or more races. These races were not added in the regression because they reflect very little of the population in New Jersey. An example of race is shown in shown in Table 4 below.

Table 4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Race by County in 2014 | | | | | |
| County | Total Population | African American | Asian | Caucasian | Hispanic or Latino |
| Atlantic | 275,325 | 40,358 | 22,037 | 158,304 | 48,865 |
| Bergen | 920,456 | 48,700 | 139,297 | 554,930 | 160,193 |
| Burlington | 450,155 | 70,793 | 20,124 | 312,829 | 31,743 |
| Camden | 512,632 | 93,339 | 27,569 | 302,836 | 76,956 |
| Cape may | 96,286 | 4,226 | 1,051 | 83,088 | 6,504 |
| Cumberland | 157,429 | 29,839 | 1,981 | 77,180 | 44,369 |
| Essex | 789,616 | 307,806 | 37,798 | 255,623 | 168,373 |
| Gloucester | 289,705 | 28,680 | 8,247 | 231,868 | 15,120 |
| Hudson | 654,878 | 73,368 | 92,362 | 194,006 | 279,712 |
| Hunterdon | 126,746 | 2,999 | 4,416 | 110,636 | 7,213 |
| Mercer | 369,526 | 72,365 | 35,364 | 195,590 | 58,733 |
| Middlesex | 824,046 | 74,558 | 187,496 | 387,573 | 158,164 |
| Monmouth | 629,702 | 42,498 | 32,688 | 479,949 | 63,488 |
| Morris | 497,103 | 15,095 | 46,834 | 365,232 | 60,488 |
| Ocean | 581,413 | 16,851 | 10,366 | 496,484 | 50,206 |
| Passaic | 505,403 | 56,112 | 25,922 | 221,995 | 194,313 |
| Salem | 65,501 | 9,063 | 585 | 49,796 | 4,882 |
| Somerset | 328,704 | 28,478 | 50,015 | 198,187 | 44,972 |
| Sussex | 146,888 | 2,747 | 2,587 | 129,351 | 10,105 |
| Union | 545,236 | 111,077 | 25,723 | 235,384 | 157,092 |
| Warren | 107,624 | 4,307 | 2,790 | 90,727 | 8,293 |

**Method**

To model my data, I will be using linear regression as my statistical method. Linear regression can “predict” the value of the dependent variable based upon the values of one or more independent variables. This statistical data analysis will be used to determine where there is a linear relationship between a dependent variable and one or more independent variables. I will be utilizing two types of linear regression: simple and multiple linear regressions. Regression analysis has three major uses and they are: causal analysis, forecasting an effect, and trend forecasting. For my causal analysis I am trying to answer the following question: “What is the strength of relationship between average test score and average income?” Using regression to forecast an effect I am trying to answer: “How much more income does a household need to earn for a student to score higher in their test score?”

In a simple linear regression a single independent variable is used to predict the value of a dependent variable. In a multiple linear regression two or more independent variables are used to predict the value of a dependent variable. The only difference between a single linear regression and a multiple linear regression is the number of independent variables. However, for both there is only one dependent variable. My dependent variable of average test scores is measured in a continuous measurement scale of 0 to 300. The independent variables are also measured in a continuous measurement scale.

A single linear regression for each year will be completed for the different types of households and race. After completing a single regression on each type a multiple linear regression will be done with all types together. I have kept my independent variables to four variables to keep it from overfitting which can make the model inefficient. I am going to keep my model as simple as possible because statistically if the model includes a large number of variables the probability increases that the variables will be statistically significant from random effects.

Regression Models:

The formula for single linear regression for New Jersey counties average HSPA scores with the average income by household type:

avg\_combined\_scores = β0 + β1(Average income by household type)

The formula for multiple linear regression for New Jersey counties average HSPA scores with all average income by household type:

avg\_combined\_scores = β0 + β1(Household)+ β2(Families)

+ β3(Married Couple)+ β4(Non Family)

The formula for single linear regression for New Jersey counties average HSPA scores with race by type:

avg\_combined\_scores = β0 + β1(Race by type)

The formula for multiple linear regression for New Jersey counties average HSPA scores with all race:

avg\_combined\_scores = β0 + β1(African American)+ β2(Asian)

+ β3(Caucasian)+ β4(Hispanic or Latino)

I will be utilizing the Gauss–Markov theorem which states that in a linear regression model in which the errors have expectation zero and are uncorrelated and have equal variances, the best linear unbiased estimator (BLUE) of the coefficients is given by the ordinary least squares (OLS) estimator, provided it exists. Assumption 1 - Linear in Parameters: The model is in linear parameters, but does not have to be linear in the x’s: Y = β 0 + β 1 X1 + β i Xi Assumption 2 - Random Sample of n Observations: An appropriately sized, random sample is used in the regression model. Assumption 3 – Zero Conditional Mean: The mean of the error terms has an expected value of zero given values for the independent variables.Assumption 4 – No Perfect Collinearity: The error term, *u*, is independently distributed and not correlated with any of the variables. The variables are not correlated. The assumption of no perfect collinearity states that *there is no exact linear relationship among the independent variables*. This assumption implies two aspects of the data on the independent variables. Assumption 5 – Homoskedasticity: The error terms all have the same variance and are not correlated with each other. In statistical jargon, the error terms are independent and identically distributed (iid). This assumption means the error terms associated with different observations are not related to each other.

**Results**

A single linear regression was completed for all the different types of household average incomes for 2014, 2010, and 2007. Then a multiple linear regression was completed on all household average incomes and the results were very interesting.

Household:

* The regression for average income by household in 2014 had a coefficient of 3.277e-04, an Adjusted R-squared of 0.7362 and p-value of 4.017e-07.
* The regression for average income by household in 2010 had a coefficient of 3.405e-04, an Adjusted R-squared of 0.6904 and a p-value of 1.855e-06.
* The regression for average income by household in 2007 had a coefficient of 2.953e-04, an Adjusted R-squared of 0.6374 and a p-value of 8.727e-06.
* This means that the average income of household has an influential factor where if there was a $10,000 increase in average household income of the county there would be a 3.28, 3.41, and 2.95 increase in the average HSPA scores for 2014, 2010, and 2007 respectively.

Table 5

|  |  |
| --- | --- |
| Linear regression of average income by household | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Family:

* The regression for average income by for a family household in 2014 had a coefficient of 2.672e-04, an Adjusted R-squared of 0.7341 and p-value of 4.336e-07.
* The regression for average income by for a family household in 2010 had a coefficient of 2.875e-04, an Adjusted R-squared of 0.6941 and p-value of 1.678e-06.
* The regression for average income by for a family household in 2007 had a coefficient of 2.494e-04, an Adjusted R-squared of 0.6318 and p-value of 1.014e-05.
* Again, a family household has an influential factor where if there was a $10,000 increase in average family income of the county there would be a 2.67, 2.87, and 2.49 increase in the average HSPA scores for 2014, 2010, and 2007 respectively.

Table 6

|  |  |
| --- | --- |
| Linear regression of average income by family | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Married Couple:

* The regression for average income by for married couple in 2014 had a coefficient of 2.332e-04, an Adjusted R-squared of 0.5665 and p-value of 4.987e-05.
* The regression for average income by for married couple in 2010 had a coefficient of 2.382e-04, an Adjusted R-squared of 0.5164 and p-value of 0.0001464.
* The regression for average income by for married couple in 2007 had a coefficient of 2.088e-04, an Adjusted R-squared of 0.4445 and p-value of 0.0005775.
* The average income for married couples was all statistically significant for 2014, 2010, and 2007.

Table 7

|  |  |
| --- | --- |
| Linear regression of average income by married couple | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Non Family:

* The regression for average income by for a non family household in 2014 had a coefficient of 5.742e-04, an Adjusted R-squared of 0.5821 and p-value of 3.484e-05.
* The regression for average income by for a non family household in 2010 had a coefficient of 5.435e-04, an Adjusted R-squared of 0.423 and p-value of 0.0008449.
* The regression for average income by for a non family household in 2007 had a coefficient of 4.607e-04, an Adjusted R-squared of 0.3963 and p-value of 0.00133.
* The coefficient results for a non family household were very surprising for me because non family households contain one person, the householder and additional persons who are not relatives of the householder. Therefore, an income increase has higher influential factor for student who might be in this type of household because their test scores increases higher than the other types of household. Just think about what type of student is living in a non family household, a foster child.

Table 8

|  |  |
| --- | --- |
| Linear regression of average income by non family | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

The results for multiple linear regression on all average income by household type was statistically significant at all levels for 2014, but the results for 2010 and 2007 varied. In 2010, only married couple was statistically significant, but the overall p-value was 5.928e-07. In 2007, family and married couple were statistically significant with an overall p-value of 4.908e-07. The results showed that average income of married couple had the most influential factor for all three years.

Table 9

|  |  |
| --- | --- |
| Multiple linear regression of average income by all household types | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

After completing the regression for average income by households, a regression for race was completed. The same method was done for race where a single a single linear regression was completed for the different types of race for 2014, 2010, and 2007 and then a multiple linear regression of all race was completed.

African American:

Before completing the regression for African American for 2014, 2010, and 2007 I hypothesized that test scores would be influential because there have been many articles regarding the “achievement gap”. After the single linear regression was completed I was not shocked with the results.

* The regression for African American in 2014 had a coefficient of -52.692, an Adjusted R-squared of 0.3059 and p-value of 0.005477.
* The regression for African American in 2010 had a coefficient of -58.868, an Adjusted R-squared of 0.3869 and p-value of 0.001552.
* The regression for African American in 2010 had a coefficient of -58.146, an Adjusted R-squared of 0.4775 and p-value of 0.0003146.
* The results for African American was statistically significant at all levels.

Table 10

|  |  |
| --- | --- |
| Single linear regression for African American | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Asian:

* The regression for Asian in 2014 had a coefficient of 36.155, an Adjusted R-squared of 0.0217 and p-value of 0.2443.
* The regression for Asian in 2010 had a coefficient of 33.990, an Adjusted R-squared of -0.002513 and p-value of 0.342.
* The regression for Asian in 2007 had a coefficient of 28.095, an Adjusted R-squared of -0.01947 and p-value of 0.4414.
* The results for Asian showed we cannot conclude that a significant difference exists because the p-value for all three years were over 0.05.

Table 11

|  |  |
| --- | --- |
| Single linear regression for Asian | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Caucasian:

* The regression for Caucasian in 2014 had a coefficient of 21.339, an Adjusted R-squared of 0.2111 and p-value of 0.02082.
* The regression for Caucasian in 2010 had a coefficient of 26.157, an Adjusted R-squared of 0.3005 and p-value of 0.005933.
* The regression for Caucasian in 2007 had a coefficient of 25.890, an Adjusted R-squared of 0.3448 and p-value of 0.003038.
* The regression results for Caucasian were statistically significant.

Table 12

|  |  |
| --- | --- |
| Single linear regression for Caucasian | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

Hispanic or Latino:

Besides African American students falling in the “achievement gap”, Hispanic or Latino student’s average scores are significantly lower than average scores for White and Asian students.

* The regression for Hispanic or Latino in 2014 had a coefficient of -38.266, an Adjusted R-squared of 0.2295 and p-value of 0.01623.
* The regression for Hispanic or Latino in 2010 had a coefficient of -42.092, an Adjusted R-squared of 0.2461 and p-value of 0.0129.
* The regression for Hispanic or Latino in 2007 had a coefficient of -35.775, an Adjusted R-squared of 0.1971 and p-value of 0.02514.

Table 13

|  |  |
| --- | --- |
| Single linear regression for Hispanic or Latino | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

The results for multiple linear regression on all races were statistically significant at all levels for all years. However, 2014 had the highest p-value of 0.004226, whereas 2007 had a p-value of 0.0001667.

Table 14

|  |  |
| --- | --- |
| Multiple Linear Regression for all race | |
| 2014 |  |
| 2010 |  |
| 2007 |  |

**Discussion**

My competitor’s research correlated family income with test scores. They supported their hypothesis, linking test scores with household income, by using data based on five repeated measures of cognitive test scores per child. Using data derived from the U.S. government’s Earned Income Tax Credit (EITC) records, they identified nonlinear changes that indicated upward shifts, or expansions, in family income over twelve years (1987-1999), a twenty-percent increase of around $2,100, between 1993 and 1997. They applied an “instrumental variable strategy” that established a causal relationship between measurable expansions in family income and corresponding increases in children’s math and reading scores. Testing to a variety of independent variables, they found short term score improvements of six percent, with one standard deviation for every $1,000 increase in income, for low-income families. The conclusions reached by recent studies suggest that unobserved heterogeneity and endogenous income shocks are important concerns. Furthermore, they suggest that income effects may be greatest among economically disadvantaged families. My competitors, Dahl and Lochner, outlined an instrumental variables strategy that eliminates omitted variable biases due to both permanent and temporary shocks correlated with family income and alleviates bias due to measurement error in income.

Dahl and Lochner completed a correlation on their study to find a cause and effect in family income and test scores. My approach was a regression where I used independent variables to find influential factors to a student’s test score. Dahl and Lochner’s results implied that a $1,000 increase in income raises combined math and reading test scores by 6 percent of a standard deviation in the short run. The regression I completed for the different types of average income by household type showed a $10,000 increase in in average income translate into an increase in the average test score of the county for 2014, 2010, and 2007.

The competitor completed a correlation versus I ran a regression on my data variables. They simply computed a correlation coefficient that tells how much one variable tends to change when the other one does. Whereas linear regression finds the best line that predicts Y from X, correlation does not fit a line. With linear regression, the X values can be measured or can be a variable that I could control. Finding the cause and effect is quite hard to find in exactly how correlated income can be to test scores because other events might be going on in the different types of household students live in. Also, while the competitor’s study reveal the correlations between income and child outcomes, they do not necessarily estimate a causal relationship. Children living in poor families may have a worse home environment or other characteristics that the researcher does not observe. These omitted variables may be part of the reason for substandard achievement and may continue to affect children’s development even if family income were to rise. Therefore, it is really hard to say if income really correlates to lower or higher test scores because there is inner working inside a household that cannot be seen or weighted.

**Conclusion**

The influential factor for a student’s test score is important to understand. Although a student cannot decide which race they are born into, understanding the powerful factor of household income is an important question. Regardless of household type, each one had a prominent factor to test scores. The results for average income by household and race shows positive influential factor between the variables and the average scores. From 2007 to 2010 the coefficients for all household types rose and then fell by a couple of points in 2014. The results from the simple linear regression for household, family, and married couple were significant at 0.001. The non family household were significant at 0.001 in 2014 and 2010, but was significant at 0.01 in 2007. The results from all years showed significant where a $10,000 increase in average income for any household type translated into an increased in the average test score in the county.

The coefficient for African American as a race for all three years were always statistically significant than the other race. African American was significant at 0.01 in 2014 and 2010 and significant at 0.001 in 2007. Hispanic or Latino was significant at 0.05 for all three years. African American and Hispanic students fall in the “achievement gap” where socio-economic factors including income levels, educational attainment, employment rates, housing options, neighborhood crime rates, and resources available to schools are worse for African Americans and Hispanics, on average, than for Whites (Patterson, C. J., et. al., 1990). The “achievement gap” often lead to fewer opportunities for African American and Hispanic children to access a wide range of activities and experience an enriched educational environment. Caucasian was significant at 0.05 in 2014 and 0.01 for 2010 and 2007. As for Asians, we cannot conclude that a significant difference exists because the p-value for all three years were over 0.05. However, because a regression is looking for the most influential factor it can mean that because Asian was not statistically significant it can mean that Asian American students excel in school because of culture. A study published in the journal PNAS, stated “Asian and Asian American youth are harder working because of cultural beliefs that emphasize the strong connection between effort and achievement. Studies show that Asian and Asian American students tend to view cognitive abilities as qualities that can be developed through effort, whereas white Americans tend to view cognitive abilities as qualities that are inborn” (Hsin, A., & Xie, Y., 2014).

In all, average income and race has an effective factor on standardized test scores, yet it may not be the main influencing factor. Students living in poor households may have a worse home environment or other characteristics that cannot be observed. Although, the regression shows income has an influential factor to test scores the omitted variables in the student’s home environment may be part of the reason for substandard and may continue to affect student’s development even if family income were to rise.

References

Anger, S., & Heineck, G. (2010). Do smart parents raise smart children? The intergenerational transmission of cognitive abilities. *Journal of Population Economics*, 23(3), 1255-1282.

Brooks‐Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child development*, *67*(2), 396-408.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood (No. w17699). *National Bureau of Economic Research*.

Chingos, M. M. (2012). Strength in Numbers: State Spending on K-12 Assessment Systems. *Brookings Institution*.

Dahl, G. B., & Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *The American Economic Review*, *102*(5), 1927-1956.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of family psychology*, *19*(2), 294.

Hoxby, C. M., & Turner, S. (2015). What high-achieving low-income students know about college. *The American Economic Review*, *105*(5), 514-517.

Hsin, A., & Xie, Y. (2014). Explaining Asian Americans’ academic advantage over whites. *Proceedings of the National Academy of Sciences*, *111*(23), 8416-8421.

Mayer, S. E. (2010). Revisiting an old question: How much does parental income affect child outcomes. *Focus*, *27*(2), 21-26.

Patterson, C. J., Kupersmidt, J. B., & Vaden, N. A. (1990). Income Level, Gender, Ethnicity, and Household Composition as Predictors of Children's School‐based Competence. *Child development*, *61*(2), 485-494.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational leadership*, *56*, 8-16.