

SPECIAL ISSUE

Value-Added: What America's Policymakers Need to Know and Understand

education policy analysis

archives

A peer-reviewed, independent,
open access, multilingual journal

epaa | aape

Arizona State University

Volume 21 Number 10

February 1st, 2013

ISSN 1068-2341

Findings vs. Interpretation in “The Long-Term Impacts of Teachers” by Chetty et al.

Moshe Adler

Columbia University & the Harry Van Arsdale, Jr. Center for Labor Studies
USA

Citation: Adler, M. (2013). Findings vs. interpretation in “The Long-Term Impacts of Teachers” by Chetty et al. *Education Policy Analysis Archives*, 21(10). This article is part of EPAA/AAPE’s Special Issue on *Value-Added: What America’s Policymakers Need to Know and Understand*, Guest Edited by Dr. Audrey Amrein-Beardsley and Assistant Editors Dr. Clarin Collins, Dr. Sarah Polasky, and Ed Sloat. Retrieved [date], from <http://epaa.asu.edu/ojs/article/view/1264>.

Abstract: The authors of the study “The Long-Term Impact of Teachers” claim that their study shows that increases in teacher value-added lead to significant and lasting increases in test scores and significant increases in income that will last throughout adulthood. Instead, I show that these claims are false because they are contradicted by the findings of the study itself. In fact, the results of the Chetty et al. study raise serious questions about the benefits of using the value-added method for evaluating teachers.

Keywords: Value added; teacher evaluation.

Resultados vs interpretación en "Los efectos a largo plazo de los docentes"

Resumen: Los autores del estudio "El impacto a largo plazo de maestros" (original en inglés “The Long-Term Impact of Teachers”) afirman que su estudio demuestra que aumentos en el valor agregado de los docentes conducen a aumentos significativos y duraderos en pruebas de resultados académicos y aumentos significativos en los ingresos económicos a lo largo de la edad adulta. Por el contrario, mostramos que estas afirmaciones son falsas, ya que se contradicen con las conclusiones

Journal website: <http://epaa.asu.edu/ojs/>

Facebook: /EPAAA

Twitter: @epaa_aape

Manuscript received: 11/12/2012

Revisions received: 12/03/2012

Accepted: 12/31/2012

del propio estudio. Así, los resultados del estudio de Chetty et al. plantean serias dudas sobre los beneficios de usar el método de valor agregado para la evaluación de los docentes.

Palabras clave: modelos de valor añadido (MVA); evaluación docente

Resultados *versus* desempenho “impacto de longo prazo de professores”

Resumo: Os autores do estudo "O impacto de longo prazo de professores" dizem que seu estudo mostra que os aumentos no valor agregado dos professores leva a aumentos significativos e duradouros em provas de desempenho acadêmico e aumentos significativos de renda durante a vida adulta. Contrariamente nós mostramos que essas alegações são falsas, e que contradizem as conclusões do próprio estudo. Assim, os resultados do estudo Chetty et al. levantam sérias dúvidas sobre os benefícios de usar modelos de valor agregado para avaliar os professores.

Palavras-chave: modelos de valor agregado (MVA); avaliação dos professores.

Introduction*

The study, “The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood,” by Raj Chetty, John N. Friedman and Jonah E. Rockoff (Chetty et al., 2011) has not yet been published in an academic journal, but it has nevertheless received wide attention. One of the authors, Raj Chetty, an economist at Harvard, received a 2012 MacArthur Foundation award for his “rigorous theoretical and empirical studies [that] are informing the design of effective government policy” and the foundation singled out “The Long-Term Impact of Teachers.” The foundation's statement read: “Chetty and colleagues found that, adjusting for other factors, students who by chance were assigned to talented teachers in elementary school had significantly higher incomes as adults and better future life outcomes more generally. By asking simple, penetrating questions and developing rigorous theoretical and empirical tests, Chetty’s timely, often surprising, findings in applied economics are illuminating key policy issues of our time” (The MacArthur Foundation, 2012).

The New York Times covered the study on its front page (Lowery, 2012) and President Obama relied on it in his 2012 State of the Union Address when he asserted “a great teacher can offer an escape from poverty to the child who dreams beyond his circumstance” (Massari, 2012). The value-added idea became a major bone of contention during the 2012 Chicago teachers' strike, when many opinion pieces about the strike made reference to the Chetty et al. study (Strauss, 2012). In New Zealand the Treasury Department posted the article on its website.¹

There is just one problem: as we explain below, the study does not show what the authors claim it shows. It does not show a long-term impact on earnings by high "value-added" teachers, nor does it show a lasting impact on test scores. But there have also been other concerns raised about the study; these will be reviewed first.

Michael Winerip, a *New York Times* reporter, highlighted an issue with the study that the authors themselves recognized: The adults in the study were children in the 1990s, a time when the students' test results had no effect on the pay of their teachers (Winerip, 2012). If teachers are judged by their students' test results, they may change the way they teach and resort to “teaching to the test.” As Winerip points out, the problem is that it is possible that the teachers who would

* The author wishes to thank Tom Russell, Ellen Adler, Alec Meiklejohn, and Henry Braun for helpful suggestions. Special thanks to Sarah Polasky and anonymous referees for their many helpful comments.

¹ <http://www.treasury.govt.nz/downloads/pdfs/gen-conf-14dec11-chetty.pdf>

become the most successful at “teaching to the test” would not be the same teachers who were the most successful in raising test scores when incentives for raising test scores did not exist. It is therefore unclear whether the association between higher test scores and better outcomes that Chetty et al. found would hold once teachers' pay depends on their value-added scores.

Richard Rothstein, a research associate at the Economic Policy Institute and a former national education columnist for *The New York Times*, has pointed out that teachers who do not excel at raising their students grades on standardized tests may excel at teaching other skills, such as cooperative behavior and social skills, skills that although perhaps antithetical to competitive behavior, are nevertheless crucial for success in adulthood (Rothstein, 2012). Thus discouraging or perhaps even firing teachers with low value-added scores may hurt, not help, students.

Professor Dale Ballou of Vanderbilt University reviewed the Chetty et al. study on behalf of the National Education Policy Center (Ballou, 2012), but his review is flawed. Ballou raises the possibility that unaccounted for factors have been the explanation for both the more successful adults and the higher value-added of their teachers in the Chetty et al. study. Of course, such a possibility always exists; this is the reason for the adage “correlation does not imply causation.” But the mere possibility is not sufficient to rule out causation. To rule out causation, an actual factor must be identified such that, once accounted for, the correlation would disappear. The factor that Ballou chose to illustrate his criticism, good parenting, would not have such an effect. According to Ballou good parents know both how to secure high value-added teachers for their children and how to prepare them for greater success in adult life. While this may be true, if the value-added of teachers does not affect adult results, why would good parents seek high value-added teachers for their children? Ballou's review is not illuminating regarding the validity of the Chetty et al. study.

Winerip, Rothstein, and Ballou focus on issues that are missing from the Chetty et al. study and argue that because of these issues incorporating value-added into teacher evaluations may not be useful and may even be harmful. We now turn to an examination of the Chetty et al. study itself.

Teacher Value-Added and Lifetime Income

Regarding the impact of teacher value-added on income, the Chetty et al. data yielded two results: 1) An increase of one standard deviation in teacher value added for one school year in childhood leads to an increase in annual income of \$182 at age 28; and 2) Teacher value-added in childhood has no effect on annual income at age 30. The only meaningful conclusion that can be drawn from these results is that a substantial increase in the value-added of a teacher for one year in elementary school leads to a small increase in income in early adulthood, but that this increase disappears by age 30. But this is not the conclusion the authors reached.

Chetty and his co-authors report that in their data the probability that teacher value-added does not have an effect on adult income at age 28 is less than 1%. In other words, the effect of teacher value-added on income at age 28 is “statistically significant.” However, although the result they found for 30 year olds is not statistically significant, the words “not statistically significant” are nowhere to be found in their study. Instead the authors write, “The 95% confidence interval for the estimate is very wide. We therefore focus on earnings impacts up to age 28 for the remainder of our analysis.”² The terms “statistically significant” and “confidence interval” are explained in the statistical terminology section below, but in plain English their statement means that they did not find conclusive evidence that teacher value-added affects the income of 30 year olds. Furthermore, they didn't just “focus” on earning impacts up to age 28; instead they proceeded as if the result for 28 year olds, an increase in income of .09%, was also the result that they found for 30 year olds, and

² Chetty et al., page 39.

made the assumption that this would have also been the result for any subsequent age group. Based on this assumption, which is contradicted by their own evidence, they calculated a life-time benefit of \$25,000 from an increase of one standard deviation in teacher value-added.

The authors claim that the reason for the “very wide” interval (translation: not statistically significant result) is that the sample of 30 year olds was smaller than the sample of 28 year olds. A smaller sample size *always* produces a wider interval, but this is not relevant. Samples almost always differ in size, but this is not license to throw away the results from smaller samples at will. Before they conduct a study, researchers typically calculate whether the sample is sufficiently large and decide whether to proceed. After conducting a study its authors cannot ignore the results; science does not permit cherry picking. The result that teacher value-added does not have a statistically significant impact on earnings at age 30 must be part of any conclusion drawn from the Chetty et al. study.

Nevertheless, did Chetty et al. miscalculate the sample size they needed for their study? Was their sample of 61,639 thirty years olds too small? The calculation of sample size requires some choices and assumptions³; using reasonable choices and assumptions the required sample size would have been 8,124.⁴ Chetty et al. therefore did not miscalculate when they decided to perform the test for 30 year olds.

Long-Lasting Test Score Gains

³ The basics of sample size calculations are explained in the statistical terminology appendix.

⁴ Sample size was calculated using the program of Professor Russell V. Lenth of the University of Iowa: <http://homepage.stat.uiowa.edu/~rlenth/Power/index.html>. This program is recommended by the NIH, which is interested in minimizing the number of animals used in medical research. See: Ralph B. Dell, Steve Holleran, and Rajasekhar Ramakrishnan, “Sample Size Determination,” *Institute for Animal Laboratory Research Journal*, 2002; 43(4): 207–213, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3275906/>. The list of parameters used is as follows:

- i. Detectable Beta (smallest meaningful value that can be detected): 2,000. I chose this value because the coefficient that is not statistically significant in the Chetty et al. equation for 30 year olds is \$2,058 (the standard deviation is \$1,953). (Chetty et al. Table 6, page 76.)
- ii. Standard Deviation of Teacher Value-Added: 0.1. This is the value that Chetty et al. found.
- iii. Number of Predictors: 26. Chetty et al. do not provide the actual equation they estimate in Table 6. In a note to Figure 1 they list all 25 explanatory variables that they use in their study, and this (plus an intercept) is the number we use for the calculation. Because sample size increases with the number of predictors, this is a conservative specification.
- iv. Variance Inflation Factor: 1. This value assumes that there is no multicollinearity among the explanatory variables. Chetty et al. do not indicate that there is multicollinearity in their data. Nevertheless, we calculated the sample size also with a Variance Inflation Factor of 5, and in this case the sample size is 41,020, still some 20,000 observations smaller than the actual sample.
- v. Alpha, the desired significance level of the test: .05.
- vi. Two tailed: Yes. Because of the possibility that students of high value-added teachers would earn less than students of low value-added teachers (see the discussion of Richard Rothstein's (Rothstein, 2012) critique above), a two tailed test is indicated. For a one tail test the required sample is smaller, 6,765.
- vii. Error Standard Deviation: \$5,000. Chetty et al. do not provide summary statistics for age 30. The standard deviation of annual income at age 28 is \$23,782. This permits the standard deviation of the error to be as large as 1/5 of the standard deviation of annual income.
- viii. Power: .95

An obvious question is how an increase in teacher value-added for just one year in elementary school can affect earnings in adulthood. Given that the increase in income at age 28 is only \$182 a year and that at age 30 there is no increase, this question may not really matter. Nevertheless, Chetty and his co-authors dealt with this issue in a way that deserves further investigation.

A teacher's value-added is measured by the effect that he or she has on his or her students' test scores. Therefore, if teacher value-added affects income in adulthood, the connection must be either directly through the student's higher test scores themselves, or through some other change—perhaps in work habits—that the teacher caused and that brought about both the increase in the student's scores in childhood and the increase in earnings later in life. But studies that preceded “The Long-Term Impacts of Teachers” discovered that increases in test scores that are due to teachers fade out after just a few years (Rothstein, 2010; Kane and Staiger, 2008; Jacob, Lefgren, and Sims, 2010).⁵ If the effect of a high-value-added teacher does not last even through a child's years in elementary school, how can it last for the rest of the child's life? “The Long-Term Impacts of Teachers” does not suffer from this problem because in this study Chetty and his co-authors discovered that a high teacher value-added leads to “long lasting test score gains.” Only that, once again, this is not what their study shows.

What the study *does* show is that an increase of one standard deviation in teacher value-added results in an increase of about 0.03 standard deviations in student test scores three years later. If the standard deviation of test scores is 27% (the standard deviation is 27% in SAT (Stanford Achievement Test) tests and 26% in the BSF (Tennessee Basic Skills First) tests; (Krueger, 1999, p. 531)), 0.03 of that value is 0.8% - less than 1%. In other words, after three years the effect of the high value-added teacher on test scores vanishes. But this is not how Chetty and his co-authors present the “fade out.” Instead, they write: “In our data, the impact of a one SD increase in teacher quality stabilizes at approximately 0.3 SD after three years, showing that students assigned to teachers with higher VA achieve long-lasting test score gains.” While this may have been an honest mistake, when asked about this, one of the authors, Jonah Rockoff, responded: “This is definitely a language error on our part. Instead of 'a one SD increase in teacher quality' we should have said 'an increase in teacher value-added of one (student-level) standard deviation.' Of course an increase of one in value added is roughly 10 teacher-level standard deviations, so your assessment of 0.03 in year three for a one teacher-level SD increase in year 0 would be correct” (Jonah Rockoff to Moshe Adler, personal communication, October 12, 2012). This answer is problematic, however.

If the authors were to change the language to reflect Rockoff's correction, it would read: “In our data, the impact of [an increase in teacher value-added of one (student-level) standard deviation] stabilizes at approximately 0.3 SD after three years, showing that students assigned to teachers with higher VA achieve long-lasting test score gains.” But this statement would not be true because the probability of finding a teacher whose value-added is 10 teacher-level standard deviations above the value-added of other teachers is practically zero. In fact, the range of teacher value-added in the Chetty et al. data is from about -0.18 to +0.18, or a maximum difference of about 0.36 teacher-level standard deviations.⁶ Students cannot be assigned to the high value-added teachers that the Chetty et al. statement mentions because such high value-added teachers simply do not exist.

⁵ Jacob et al. (2010) found that only 25% of the first year teacher effect remains in the second year. Kane and Staiger (2008) found that 50% remain in the second year and 25% remains in the third. J. Rothstein (2010) found that 30% remains in the second year. Furthermore, Rothstein also found that the correlation between the 1st year effect and the 3rd year effect is only 0.4, leading him to conclude: “A teacher's first-year effect is a poor proxy for his or her longer-run impact.”

⁶ Chetty et al., Figure 6.

The imaginary high value-added teacher is not a minor character in the Chetty et al. study. In fact, the authors draw the reader's attention to the non-existent high value-added teacher by producing a dramatic chart (Figure 1) that shows what an increase in teacher value-added of one unit would do. But, of course, this figure is false, because the maximum value-added increase possible in their own sample is only .36. The impact of increasing teacher value-added by one teacher-level standard deviation is shown in Figure 2.

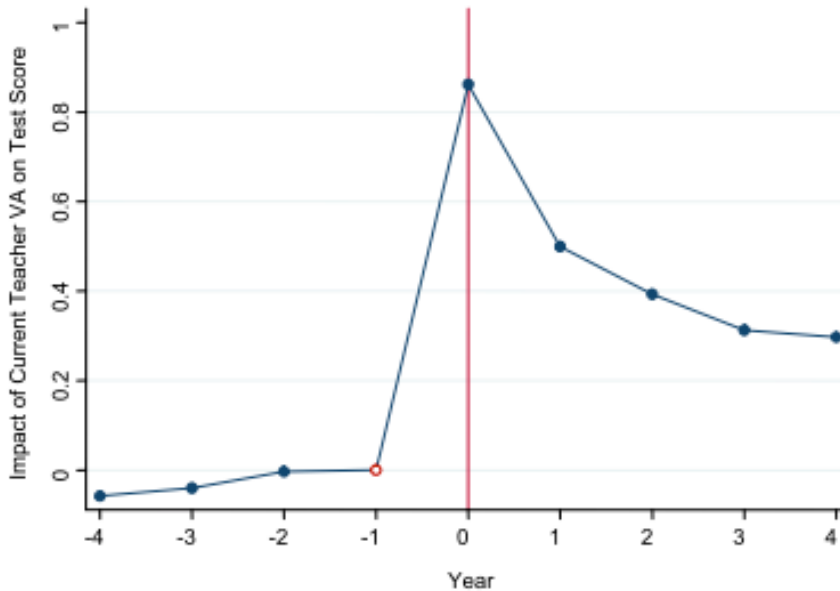


Figure 1. Impacts of teacher value-added on lagged, current, and future test scores

Source: Figure 2, Chetty et al.

Note (m.a.): These are the impacts of a one-unit increase in teacher value-added at year 0. This increase is equivalent to a 10 teacher-level standard deviations increase in teacher value-added.

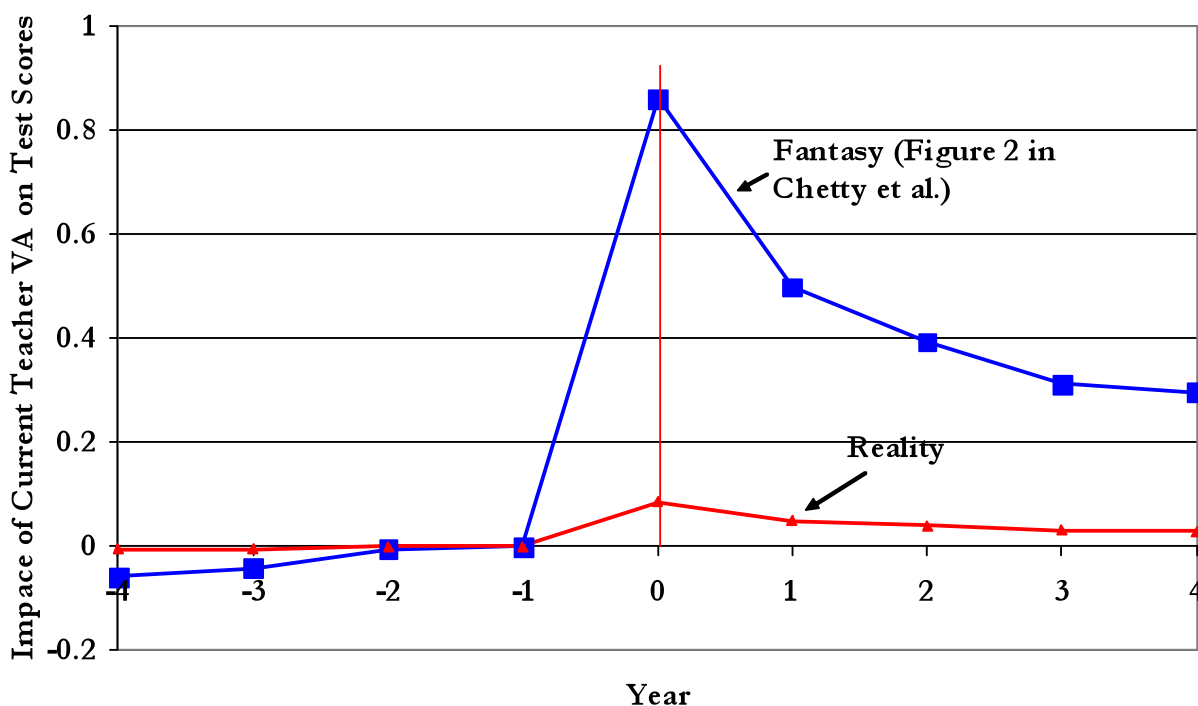


Figure 2: Impacts of teacher value-added on lagged, current, and future test scores

Source: Fantasy line: Figure 2, Chetty et al. Reality line: Author’s calculations

Note (m.a.): The Reality line shows the impacts of a 0.1 unit increase in teacher value added. This increase is equivalent to a one teacher-level standard deviation increase in teacher value-added. The Fantasy line is identical to the line in Figure 1.

Conclusion

The authors of the study “The Long-Term Impact of Teachers” claim that their study shows that increases in teacher value-added lead to significant and lasting increases in test scores and significant increases in income that will last throughout adulthood. As this article has shown, these claims are false because they are contradicted by the findings of the study itself. A substantial (one standard deviation) increase in teacher value-added increases test scores by 2.7% in the first year and by 0.8% after three years. The same increase in value-added adds \$182 of income at age 28, and no statistically significant increase in income at age 30, only two years later. Thus, the results of the Chetty et al. study raise serious doubts about the benefits of using the value-added method for evaluating teachers. Although the study does not successfully make the case for measuring the value-added of teachers, it does very eloquently make the case for the importance of peer review, since the claims made in the study have received wide attention and broad endorsement both in the United States and internationally. The gap between the results and their interpretation in the Chetty et al. study is so significant that there is little doubt that peer reviewers would have identified it.

References

- Ballou, D. (2012, February 16). Review of The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. National Education Policy Center. Retrieved from <http://nepc.colorado.edu/thinktank/review-long-term-impacts>
- Chetty, R., Friedman, J. & Rockoff, J. (2011, December, revised 2012, January). The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. Working Paper, National Bureau of Economic Research. *Working Paper No. 17699*. National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w17699>
- Jacob, B., Lefgren, L. & Sims, D. (2010). The Persistence of Teacher-Induced Learning Gains. *Journal of Human Resources*, 45(4): 915-943.
- Kane, T. J. & Staiger, D. (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. Working Paper 14607, National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14607>
- Krueger, A. Experimental Estimates of Education Production Functions. (1999, May). *The Quarterly Journal of Economics*, 497-532.
- Lowery, A. (2012, January 6). Big Study Links Good Teachers to Lasting Gain. *The New York Times*, A1.
- The MacArthur Foundation. (2012, October 2). Meet the Class of 2012: Raj Chetty <http://www.macfound.org/fellows/861/>
- Massari, P. Superstar Teachers. (2012, February 21). *The Harvard Gazette*.
- Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics*, 125(1): 175-214.
- Rothstein, R. (2012, September 16). What Research Really Says on Teacher Evaluation. *The Washington Post*. Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/what-research-really-says-on-teacher-evaluation/2012/09/16/2e9de9fa-ff44-11e1-8adc-499661afe377_blog.html
- Strauss, V. (2012, September 16). Introduction to Richard Rothstein. What Research Really Says on Teacher Evaluation. *The Washington Post*. Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/what-research-really-says-on-teacher-evaluation/2012/09/16/2e9de9fa-ff44-11e1-8adc-499661afe377_blog.html
- Winerip, M. (2012, January 15). Study on Teacher Value Uses Data From Before Teach-to-Test Era. *The New York Time*, A13.

Appendix: Statistical Terminology

Standard Deviation

The measurement unit that statistical studies -- including the Chetty et al. study -- use is the standard deviation. I will explain it through an example. Suppose two students, A and B, have test scores of 60% and 80% respectively. Their mean grade is 70%, and their individual score's deviations from the mean are -10% and +10% respectively. In absolute value the deviations are both 10%, and therefore the "standard deviation" in this case is 10%. When measured in standard deviations, student A's score is -1 standard deviation and student B's score is +1 standard deviation. The difference between their scores is two standard deviations.⁷

Measuring scores in standard deviations instead of percentages has two advantages. First, when measured in standard deviations, student scores are the same regardless of what units the test actually used. The SAT test, for example, uses a scale of 0-800 points. With this scale student A's score would have been 480, student B's score would have been 640 and the average score would have been 560. But in standard deviations A's and B's scores would have still been -1.0 and +1.0, respectively.

Measuring scores in standard deviations has an additional advantage. Suppose that there are many students, not just two, and suppose that their test scores are distributed along a bell ("normal") curve. Then a student whose score is one standard deviation above the mean is in the 84th percentile of scores, and a student who is 1.64 standard deviations above the mean is in the 95th percentile of scores. In other words, when scores are expressed in standard deviations, we can figure out the percentile rankings of the students.

Teacher Value-Added

A teacher's value-added score is equal to the number of standard deviations that he or she adds, on average, to the test scores of his or her students. What does this mean?

As was seen above, the standard deviation of student scores is about 27%. If the value-added score of a teacher is +1.0, that teacher adds, on average, one standard deviation, or 27%, to his or her students' test scores. Chetty et al. found that the differences in teacher value-added scores are small, their standard deviation is 0.1. What does this mean? Suppose that the value-added score of teacher B is one standard deviation above the value-added score of teacher A. Also suppose that the average score of the students in teacher A's class is 60%. Had teacher B been the teacher in that classroom instead of teacher A, then the average score of the students would have been 62.7%.

Three years later, the effect of teacher value-added fades out. If the students who three years earlier were in teacher A's class score 70%, these students would have scored 70.8% on average had teacher B been their instructor instead of teacher A.

Statistical Significance

Chetty et al. estimated the rate of increase in income in adulthood due to teacher value-added. Such a rate cannot be estimated with absolute precision, and what scientists report instead is a "confidence interval" for which, by convention, the level of confidence is of 95%.

Most people are familiar with confidence intervals from public opinion polls. A pollster always reports the result of his or her poll with a "margin of error." The margin is determined by

⁷ When there are many observations, statisticians use a slightly different method for calculating the standard deviation, but for our purposes -- understanding the basic concept -- the explanation above is sufficient.

the size of the sample and by the level of confidence that the pollster wants to have in his or her results. Suppose that the confidence level is 95% and that the pollster found that in his or her sample 52% of respondents said they would vote for candidate X. If the margin of error is $\pm 1\%$, he or she will predict, with 95% confidence, that X will be the winner. But if the margin of error is $\pm 3\%$ he or she will say that the poll is “too close to call” because the confidence interval is “too wide.” The span from 49% to 55% contains the possibility that X will lose. Academic researchers would that the result of 52% is “statistically insignificant.”

The result in Chetty et al. -- an increase of \$206 in annual income at age 30 due to an increase of one standard deviation in teacher value-added -- is not statistically significant at the 5% level because the margin of error is \$383.⁸ The confidence interval spans from -\$177 to +\$589, and it is therefore impossible to say with a 95% level of confidence whether an increase in teacher value-added would lead to a loss or a gain in income at age 30. This is why Chetty et al. state that the confidence interval is too wide.

Sample Size Basics

To understand the basic calculation involved in determining sample size I will continue with the election poll example. As already mentioned, the size of the margin of error depends on the size of the sample and the level of confidence, which by convention is set to 95%. With 1,067 respondents, the margin of error is 3%. With 4,268 respondents the margin of error is 1.5%, and with 9,604 the margin of error is only 1%. Of course, if all voters are in the sample the margin of error is 0% and the level of confidence is 100%. Polling all voters isn't practical, however, and this is why pollsters use smaller samples instead. What size sample should the pollster choose? A pollster who is not concerned that if one of the candidates is favored by 53% of respondents (or less) she would not be able to predict the elections would draw a sample of 1,067 respondents. A pollster who wants to be able to make a prediction even then would have to draw a larger sample. A similar choice of parameters is involved when calculating the sample size required for the Chetty et al. study.

⁸ The standard deviation of the effect of an increase of one standard deviation in teacher value-added is \$195 (Chetty et al., Table 6).

About the Author

Moshe Adler

Department of Urban Planning, Columbia University & the Harry Van Arsdale, Jr. Center for Labor Studies

Email: ma820@columbia.edu

Moshe Adler is the author of *Economics for the Rest of Us: Debunking the Science that Makes Life Dismal*, The New Press, 2010.

SPECIAL ISSUE

Value-Added: What America’s Policymakers Need to Know and Understand

education policy analysis archives

Volume 21 Number 10

February 1st, 2013

ISSN 1068-2341



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A2 (Brazil), SCImago Journal Rank; SCOPUS, Socolar (China).

Please contribute commentaries at <http://epaa.info/wordpress/> and send errata notes to Gustavo E. Fischman fischman@asu.edu

Join **EPAA’s Facebook community** at <https://www.facebook.com/EPAAAPE> and **Twitter feed** @epaa_aape.

education policy analysis archives
editorial board

Editor **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **David R. Garcia** (Arizona State University), **Stephen Lawton** (Arizona State University)

Rick Mintrop, (University of California, Berkeley) **Jeanne M. Powers** (Arizona State University)

Jessica Allen University of Colorado, Boulder

Gary Anderson New York University

Michael W. Apple University of Wisconsin, Madison

Angela Arzubiaga Arizona State University

David C. Berliner Arizona State University

Robert Bickel Marshall University

Henry Braun Boston College

Eric Camburn University of Wisconsin, Madison

Wendy C. Chi* University of Colorado, Boulder

Casey Cobb University of Connecticut

Arnold Danzig Arizona State University

Antonia Darder University of Illinois, Urbana-Champaign

Linda Darling-Hammond Stanford University

Chad d'Entremont Strategies for Children

John Diamond Harvard University

Tara Donahue Learning Point Associates

Sherman Dorn University of South Florida

Christopher Joseph Frey Bowling Green State University

Melissa Lynn Freeman* Adams State College

Amy Garrett Dikkers University of Minnesota

Gene V Glass Arizona State University

Ronald Glass University of California, Santa Cruz

Harvey Goldstein Bristol University

Jacob P. K. Gross Indiana University

Eric M. Haas WestEd

Kimberly Joy Howard* University of Southern California

Aimee Howley Ohio University

Craig Howley Ohio University

Steve Klees University of Maryland

Jackyung Lee SUNY Buffalo

Christopher Lubienski University of Illinois, Urbana-Champaign

Sarah Lubienski University of Illinois, Urbana-Champaign

Samuel R. Lucas University of California, Berkeley

Maria Martinez-Coslo University of Texas, Arlington

William Mathis University of Colorado, Boulder

Tristan McCowan Institute of Education, London

Heinrich Mintrop University of California, Berkeley

Michele S. Moses University of Colorado, Boulder

Julianne Moss University of Melbourne

Sharon Nichols University of Texas, San Antonio

Noga O'Connor University of Iowa

João Paraskveva University of Massachusetts, Dartmouth

Laurence Parker University of Illinois, Urbana-Champaign

Susan L. Robertson Bristol University

John Rogers University of California, Los Angeles

A. G. Rud Purdue University

Felicia C. Sanders The Pennsylvania State University

Janelle Scott University of California, Berkeley

Kimberly Scott Arizona State University

Dorothy Shipps Baruch College/CUNY

Maria Teresa Tatto Michigan State University

Larisa Warhol University of Connecticut

Cally Waite Social Science Research Council

John Weathers University of Colorado, Colorado Springs

Kevin Welner University of Colorado, Boulder

Ed Wiley University of Colorado, Boulder

Terrence G. Wiley Arizona State University

John Willinsky Stanford University

Kyo Yamashiro University of California, Los Angeles

* Members of the New Scholars Board

archivos analíticos de políticas educativas
consejo editorial

Editor: **Gustavo E. Fischman** (Arizona State University)

Editores. Asociados **Alejandro Canales** (UNAM) y **Jesús Romero Morante** (Universidad de Cantabria)

Armando Alcántara Santuario Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

Claudio Almonacid Universidad Metropolitana de Ciencias de la Educación, Chile

Pilar Arnaiz Sánchez Universidad de Murcia, España

Xavier Besalú Costa Universitat de Girona, España

Jose Joaquin Brunner Universidad Diego Portales, Chile

Damián Canales Sánchez Instituto Nacional para la Evaluación de la Educación, México

María Caridad García Universidad Católica del Norte, Chile

Raimundo Cuesta Fernández IES Fray Luis de León, España

Marco Antonio Delgado Fuentes Universidad Iberoamericana, México

Inés Dussel FLACSO, Argentina

Rafael Feito Alonso Universidad Complutense de Madrid, España

Pedro Flores Crespo Universidad Iberoamericana, México

Verónica García Martínez Universidad Juárez Autónoma de Tabasco, México

Francisco F. García Pérez Universidad de Sevilla, España

Edna Luna Serrano Universidad Autónoma de Baja California, México

Alma Maldonado Departamento de Investigaciones Educativas, Centro de Investigación y de Estudios Avanzados, México

Alejandro Márquez Jiménez Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Felipe Martínez Fernández University of California Los Angeles, USA

Fanni Muñoz Pontificia Universidad Católica de Perú

Imanol Ordorika Instituto de Investigaciones Economicas – UNAM, México

María Cristina Parra Sandoval Universidad de Zulia, Venezuela

Miguel A. Pereyra Universidad de Granada, España

Monica Pini Universidad Nacional de San Martín, Argentina

Paula Razquin UNESCO, Francia

Ignacio Rivas Flores Universidad de Málaga, España

Daniel Schugurensky Universidad de Toronto-Ontario Institute of Studies in Education, Canadá

Orlando Pulido Chaves Universidad Pedagógica Nacional, Colombia

José Gregorio Rodríguez Universidad Nacional de Colombia

Miriam Rodríguez Vargas Universidad Autónoma de Tamaulipas, México

Mario Rueda Beltrán Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Luis San Fabián Maroto Universidad de Oviedo, España

Yengny Marisol Silva Laya Universidad Iberoamericana, México

Aida Terrón Bañuelos Universidad de Oviedo, España

Jurjo Torres Santomé Universidad de la Coruña, España

Antoni Verger Planells University of Amsterdam, Holanda

Mario Yapu Universidad Para la Investigación Estratégica, Bolivia

arquivos analíticos de políticas educativas
conselho editorial

Editor: **Gustavo E. Fischman** (Arizona State University)
Editores Associados: **Rosa Maria Bueno Fisher** e **Luis A. Gandin**
(Universidade Federal do Rio Grande do Sul)

Dalila Andrade de Oliveira Universidade Federal de Minas Gerais, Brasil
Paulo Carrano Universidade Federal Fluminense, Brasil
Alicia Maria Catalano de Bonamino Pontifícia Universidade Católica-Rio, Brasil
Fabiana de Amorim Marcello Universidade Luterana do Brasil, Canoas, Brasil
Alexandre Fernandez Vaz Universidade Federal de Santa Catarina, Brasil
Gaudêncio Frigotto Universidade do Estado do Rio de Janeiro, Brasil
Alfredo M Gomes Universidade Federal de Pernambuco, Brasil
Petronilha Beatriz Gonçalves e Silva Universidade Federal de São Carlos, Brasil
Nadja Herman Pontifícia Universidade Católica –Rio Grande do Sul, Brasil
José Machado Pais Instituto de Ciências Sociais da Universidade de Lisboa, Portugal
Wenceslao Machado de Oliveira Jr. Universidade Estadual de Campinas, Brasil

Jefferson Mainardes Universidade Estadual de Ponta Grossa, Brasil
Luciano Mendes de Faria Filho Universidade Federal de Minas Gerais, Brasil
Lia Raquel Moreira Oliveira Universidade do Minho, Portugal
Belmira Oliveira Bueno Universidade de São Paulo, Brasil
Antônio Teodoro Universidade Lusófona, Portugal
Pia L. Wong California State University Sacramento, U.S.A
Sandra Regina Sales Universidade Federal Rural do Rio de Janeiro, Brasil
Elba Siqueira Sá Barreto Fundação Carlos Chagas, Brasil
Manuela Terrasêca Universidade do Porto, Portugal
Robert Verhine Universidade Federal da Bahia, Brasil
Antônio A. S. Zuin Universidade Federal de São Carlos, Brasil