In *The Class Size Debate,* two eminent economists debate the merits of smaller class sizes and the research methods used to measure the efficacy of this education reform measure. Alan Krueger (Princeton University) maintains that smaller class sizes can improve students' performance and future earnings prospects. He challenges Prof. Hanushek's widely cited analysis of the class size literature, arguing that it gives disproportionate weight to single studies that include a large number of estimates. An appropriate weighting, he says, would reveal that class size is indeed a determinant of student achievement.

Eric Hanushek (Stanford University) counters that Prof. Krueger's re-analysis achieves results different from his own by emphasizing low-quality estimates. He argues that other policies besides class size reduction, such as improving teacher quality, are more important.

Jennifer King Rice (University of Maryland) brings a third-party perspective to the debate. She addresses each author's arguments and focuses on the policy implications of the class size literature.

Lawrence Mishel is vice president of the Economic Policy Institute and co-author of *The State of Working America.* **Richard Rothstein** is an adjunct professor of public policy at Occidental College, the national education columnist for *The New York Times,* and the author of *The Way We Were? The Myths and Realities of America's Student Achievement.*

The **Economic Policy Institute** is a nonprofit, nonpartisan research organization that seeks to broaden the public debate about strategies to achieve a prosperous and fair economy. The Institute stresses real world analysis and a concern for the living standards of working people, and it makes its findings accessible to the general public, the media, and policy makers. EPI's books, studies, and popular education materials address important economic issues, analyze pressing problems facing the U.S. economy, and propose new policies.

Economic Policy Institute books are available in bookstores and at **www.epinet.org**.

ISBN 0-944826-92-X

THE CLASS SIZE DEBATE

Mishel & Rothstein, editors

EPI

THE CLASS SIZE DEBATE

Lawrence Mishel & Richard Rothstein, editors

Alan B. Krueger, Eric A. Hanushek, & Jennifer King Rice, contributors

ECONOMIC POLICY INSTITUTE

THE CLASS SIZE DEBATE

Other books from the Economic Policy Institute

The State of Working America

Market-Based Reforms in Urban Education

School Vouchers:
Examining the Evidence

Can Public Schools Learn From Private Schools?

Where's the Money Gone?
Changes in the Level and Composition of Education Spending

Risky Business:
Private Management of Public Schools

School Choice:
Examining the Evidence

THE CLASS SIZE DEBATE

Lawrence Mishel & Richard Rothstein, editors

Alan B. Krueger, Eric A. Hanushek,
& Jennifer King Rice, contributors

ECONOMIC POLICY INSTITUTE

Washington, D.C.

The Class Size Debate is a publication of
the Economic Policy Institute's Education Program.

Copyright © 2002

ECONOMIC POLICY INSTITUTE
1660 L Street, NW, Suite 1200
Washington, D.C. 20036

http://www.epinet.org

ISBN: 0-944826-92-X

Table of contents

About the editors & contributors vii

Acknowledgments viii

INTRODUCTION 1

CHAPTER 1: UNDERSTANDING THE MAGNITUDE AND EFFECT OF CLASS SIZE ON STUDENT ACHIEVEMENT 7
by Alan B. Krueger

I. Reanalysis of Hanushek's literature review 10
Expenditures per student 19
Summing up 21
II. Economic criterion 22
Lazear's theory of class size 22
Benefits and costs of educational resources 23
The 'critical effect size' 28
Caveats 29
III. Conclusion 31
Endnotes 33

CHAPTER 2: EVIDENCE, POLITICS, AND THE CLASS SIZE DEBATE 37
by Eric A. Hanushek

I. The history of class size reduction 39
II. Econometric evidence 41
The misleading Krueger presentation 42
The importance of quality 46
The source of difference in tabulations of results 50
III. The Tennessee class size experiment (Project STAR) 55
IV. Policy calculations 58
V. Conclusions 61
Appendix: Issues with the econometric data 62
Endnotes 63

CHAPTER 3: A RESPONSE TO ERIC HANUSHEK'S "EVIDENCE, POLITICS, AND THE CLASS SIZE DEBATE" 67
by Alan B. Krueger
A closer look at the nine studies that receive the most weight 73
Statistical significance of estimates 79
NAEP data 81
Minor points of disagreement 82
Conclusion 85
Endnotes 87

CHAPTER 4: MAKING THE EVIDENCE MATTER: IMPLICATIONS OF THE CLASS SIZE RESEARCH DEBATE FOR POLICY MAKERS 89
by Jennifer King Rice
The nature of the debate 90
Policy issues 91
Implications for research 94

References 95

About EPI 102

About the editors & contributors

Lawrence mishel is the vice president of the Economic Policy Institute and was its research director from 1987 to 1999. His areas of research are labor economics, wage and income distribution, industrial relations, productivity growth, and the economics of education. He is the co-author (with Jared Bernstein and Heather Boushey) of the forthcoming *State of Working America 2002-03* and the co-editor of *Beware the U.S. Model* (with John Schmitt) and *Unions and Economic Competitiveness* (with Paula Voos). He holds a Ph.D. in economics from the University of Wisconsin.

Richard Rothstein is a research associate of the Economic Policy Institute, the national education columnist for *The New York Times,* a contributing editor of *The American Prospect,* and an adjunct professor of public policy at Occidental College in Los Angeles. He is the author of *Can Public Schools Learn From Private Schools* (with Martin Carnoy and Richard Benveniste); *The Way We Were? The Myths and Realities of America's Student Achievement*; and *Where's the Money Gone? Changes in the Level and Composition of Education Spending 1967-91*; and co-editor (with Edith Rasell) of *School Choice: Examining the Evidence*.

Alan B. Krueger is the Bendheim Professor of Economics and Public Affairs at Princeton University. Since 1987 he has held a joint appointment in the Economics Department and Woodrow Wilson School at Princeton. He has published articles in academic journals on a wide range of subjects, including unemployment, social insurance, labor demand, the economics of education, income dispersion, technological change, health economics, and environmental economics. He is author of *Education Matters: Selected Essays on Education by Alan B. Krueger,* co-editor (with Robert Solow) of *The Roaring Nineties: Can Full Employment Be Sustained?* and, since 1996, editor of the *Journal of Economic Perspectives*, a journal of the American Economic Association. In 1994-95 he served as chief economist of the U.S. Department of Labor. He received a Ph.D. in economics from Harvard University in 1987.

Eric A. Hanushek is the Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University and a research associate of the National Bureau of Economic Research. He is a leading expert on educational policy with an emphasis on the economics and finance of schools. His books include *Improving America's Schools, Making Schools Work, Educational Performance of the Poor*, *Education and Race*, *Assessing Policies for Retirement Income*,

Modern Political Economy, *Improving Information for Social Policy Decisions*, and *Statistical Methods for Social Scientists*, along with numerous articles in professional journals. He has held a number of posts in government, including deputy director of the Congressional Budget Office, senior staff economist at the Council of Economic Advisers, and senior economist at the Cost of Living Council. He earned a Ph.D. in economics at the Massachusetts Institute of Technology.

Jennifer King Rice is an assistant professor in the Department of Education Policy and Leadership at the University of Maryland. Her research interests include education policy, education productivity, cost analysis applications to education, and educational reforms for at-risk students. Her publications and presentations have precipitated invitations to share her expertise with various organizations including the U.S. Department of Education, the Maryland State Department of Education, the New York State Board of Regents, the Maryland State Attorney General's Office, the National Center for Education Statistics, and the Society of Government Economists. She earned a Ph.D. from Cornell University.

Acknowledgments

Prof. Krueger's paper is a revised and extended version of a paper that was originally prepared for a conference sponsored by Temple University's Center for Research in Human Development and Education, titled, "What Do We Know About How to Make Small Classes Work?" held December 6-7, 1999 in Washington, D.C. The paper was written while Prof. Krueger was on leave at the Center for Advanced Study in the Behavioral Sciences at Stanford University. He is grateful to Diane Whitmore and Michael Watts for excellent research assistance, to Victor Fuchs for helpful comments, and to Eric Hanushek for providing the data used in Section I. Jesse Rothstein provided valuable editorial assistance. The Center for Advanced Study in the Behavioral Sciences, Temple's Center for Research in Human Development and Education, and Princeton's Industrial Relations Section provided financial support.

Helpful comments on Prof. Hanushek's paper were provided by John Kain, Steve Landsburg, Ed Lazear, Terry Moe, Paul Peterson, Macke Raymond, and Steve Rivkin.

Introduction

For three decades, a belief that public education is wasteful and inefficient has played an important role in debates about its reform. Those who have proposed new spending programs for schools to improve student achievement have been on the defensive. The presumption has been that changes in structure and governance of schools — like choice, vouchers, charter schools, standards, accountability, and assessment — are the only way to improve student outcomes. Traditional interventions, like smaller class size and higher teacher salaries, have been presumed ineffective.

Voters and state and local political leaders have never been as impressed with this statement of alternatives as have national policy makers and scholars. Throughout the last third of the 20th century, when the idea that "money makes no difference" held sway in academic circles, spending in public education increased at a steady rate, and class sizes declined. But, as we showed in a 1995 Economic Policy Institute report, *Where's the Money Gone?*, the spending has increased more slowly than most people believe. It can't be known whether the rate would have been more rapid in the absence of an academic consensus regarding public education's inefficiency.

The leading proponent of the prevailing view that money doesn't make a difference has been Eric A. Hanushek, now of the Hoover Institution. Dr. Hanushek has played two roles. As a scholar, he has conducted a series of influential literature reviews that support the conclusion that increased spending in general, and smaller class size in particular, do not "systematically" lead to improved student achievement. There have been hundreds of research studies that attempt to assess the relationship of spending and achievement. Dr. Hanushek has found that, in some cases, the relationship is positive, but in others no positive relationship can be discerned, either because the relationship is negative or because it is statistically insignificant.

These findings have led Dr. Hanushek to play another role — as a very visible public advocate for restraining the growth of spending in public schools. He chaired a task force of the Brookings Institution, leading to the publication of *Making Schools Work: Improving Performance and Con-*

trolling Costs, a very influential 1993 book that asserts, "Despite ever rising school budgets, student performance has stagnated.…[I]n recent years the costs of education have been growing far more quickly than the benefits." Dr. Hanushek has testified in many state court cases regarding the equity and adequacy of school spending, generally in support of the proposition that increased funds are not a likely source of improved student achievement. He is also frequently cited in newspapers and magazines in support of this proposition.

Dr. Hanushek's academic research, inventorying and summarizing existing studies of the relationship between spending and achievement, does not inexorably lead to conclusions about the desirability of restraining school spending. Even if his conclusion about the lack of a "systematic" relationship is unchallenged, it remains the case that some studies show a positive relationship, and therefore it might be possible to determine when, and under what conditions, higher spending produces student achievement. Dr. Hanushek states as much in almost all of his academic publications, but with the caveat that "simply knowing that some districts might use resources effectively does not provide any guide to effective policy, unless many more details can be supplied." However, Dr. Hanushek's research has not led a generation of scholars and policy makers to seek to supply these details. Rather, the impact has mostly been to encourage policy makers to look away from resource solutions and toward structural and governance changes.

In recent years, the most important challenge to this dominant trend has arisen because of an unusual experiment (STAR, or the Student Teacher Achievement Ratio study) conducted by the state of Tennessee. Attempting to determine whether achievement would increase with smaller class sizes, the state legislature authorized schools to volunteer to participate in an experiment whereby they would receive additional funds for lower class sizes for kindergarten to third-grade classes, provided that students and teachers were randomly assigned to regular (large) or small classes.

The result was significantly enhanced achievement for children, especially minority children, in smaller classes. This single study persuaded many scholars and policy makers that smaller classes do make a difference, because the study was believed to be of so much higher quality than the hundreds of non-experimental studies about which Dr. Hanushek had relied for his summaries. Most theoreticians have long believed that conducting true randomized field experiments is the only valid method for resolving disputes of this kind. The reason is that, in non-experimental studies, comparisons between groups must ultimately rely on researchers' assumptions about similarity of the groups' characteristics. This makes the studies subject to errors from mis-specification (for example, assuming that black

students who receive free or reduced-price lunch subsidies are similar in relevant respects to white students who receive these subsidies) or from omitted variables (for example, failing to recognize that parental education levels are important determinants of student achievement).

Randomized field trials, on the other hand, avoid these flaws because, if treatment and control groups are randomly selected from large enough populations, researchers can assume that their relevant characteristics (whatever those characteristics may be) will be equally distributed between the two groups. In a non-experimental study, retrospective comparison of student achievement in small and large classes may lead to the conclusion that small classes are superior only because of some unobserved characteristic that distinguishes the two groups, besides the size of their classes. In an experimental study, results are more reliable because the unobserved characteristics, whatever they may be, are evenly distributed.

It is hard to avoid the conclusion that however valid the Tennessee study will ultimately be judged to have been, enthusiasm for it has been somewhat excessive because another principle of scientific experimentation is that results should be confirmed over and over again before acceptance, in different laboratories where unobserved laboratory conditions may be different. In this case, even if the Tennessee results are entirely reliable, policy conclusions are being drawn that go beyond what the Tennessee results can support. For example, the Tennessee study showed that small classes are superior to large ones, but because both types of classes were mostly taught by teachers trained in Tennessee colleges, earning similar salaries on average, it is possible that the results would not be reproduced by teachers trained in different institutions, having different qualifications, or earning higher or lower salaries. As another example, the Tennessee study found that student achievement was higher in classes of about 16 than in classes of about 24. The Tennessee study itself cannot suggest whether other degrees of reductions in class size would also boost achievement.

Nonetheless, the Tennessee study has had great influence on policy makers. In California, the governor and legislature made the needed additional money available to all schools that reduced class sizes to 20 in grades K-3. California previously had nearly the largest class sizes in the nation, so the reductions were substantial. But implementation of this policy illustrates the dangers of rushing to make policy changes based on limited research. Because California increased its demand for elementary school teachers so suddenly, many teachers without training or credentials were hired. At the same time, many experienced teachers, working in lower-income and minority communities, transferred to districts with more affluent and easier-to-teach students, taking advantage of the vast numbers of sudden

openings in suburban districts. Class size reduction therefore had the result in California of reducing the average experience (and, presumably quality) of K-3 teachers in the inner city. Nonetheless, since the implementation of the class size reduction policy, test scores in California schools, including schools that are heavily minority and low income, rose. But because California simultaneously implemented other policy changes (abolition of bilingual education, a stronger accountability system), it is uncertain to what extent class size reduction has been responsible for the test score gains.

Thus, as we enter a new decade, these two controversial lines of research — Dr. Hanushek's conclusion that there is no systematic relationship between resources and achievement, and the STAR results that smaller class sizes do make a difference — while not entirely inconsistent, are contending for public influence.

In the following pages, the Economic Policy Institute presents a new critique of Dr. Hanushek's methodology by Alan Krueger, a professor of economics at Princeton, and a reply by Dr. Hanushek.

Dr. Krueger's paper has two parts. First, he criticizes Dr. Hanushek's "vote counting" method, or how Dr. Hanushek adds together previous studies that find a positive relationship and those that find none. In particular, Dr. Krueger notes that many of the published studies on which Dr. Hanushek's conclusions rely contain multiple estimates of the relationship between resources and achievement, and in particular between pupil-teacher ratio and achievement. In these cases, Dr. Hanushek counted each estimate separately to arrive at the overall total of studies that suggested either a positive, negative, or statistically insignificant effect for resources. But Dr. Krueger suggests that it would be more appropriate to count each publication as a single "study," rather than counting separately each estimate within a publication. By counting each publication as only one result, Dr. Krueger concludes that the effect of resources on achievement is much more positive than Dr. Hanushek found.

In the second part of his paper, Dr. Krueger applies the findings of the Tennessee STAR experiment to his own previous research on the effect of school spending on the subsequent earnings of adults, and to similar research conducted with British data. From assumptions about future interest rates, Dr. Krueger estimates the long-term economic benefits in greater income from class size reduction, and concludes that, with plausible assumptions, the benefits can be substantial, exceeding the costs.

In this respect, Dr. Krueger's paper is an important advance in debates about education productivity. By comparing the long-term economic benefits and costs of a specific intervention, he has shown that education policy making can go beyond an attempt to evaluate school input policies solely

by short-term test score effects. While, in this preliminary exploration, Dr. Krueger has had to make substantial assumptions about the organization and financial structures of schools (assumptions he notes in "caveats" in the paper), he has defined a framework for the cost-benefit analysis of school spending for other researchers to explore, elaborate, and correct.[1]

Dr. Hanushek responds to each of the Krueger analyses. With regard to the claim that "vote counting" should be based on only one "vote" per published study, Dr. Hanushek challenges the statistical assumptions behind Dr. Krueger's view and concludes, again, that his own method, of counting each estimate as a separate study, is more valid. Dr. Krueger's method, he suspects, was designed mainly for the purpose of getting a more positive result.

With respect to Dr. Krueger's estimates of the long-term economic effects of class size reduction, Dr. Hanushek notes that the estimates ultimately rely solely on evidence of labor market experiences of young Britons in the 1980s. "While it may be academically interesting to see if there is any plausibility to the kinds of class size policies being discussed, one would clearly not want to commit the billions of dollars implied by the policies on the basis of these back-of-the-envelope calculations."

It is unfortunate that the subject of public education has become so polarized that policy debates, allegedly based on scholarly research, have become more contentious than the research itself seems to require. A careful reading of the papers that follow cannot fail to lead readers to the conclusion that there is substantial agreement between these antagonists. It is perhaps best expressed by Dr. Hanushek when he states,

> Surely class size reductions are beneficial in specific circumstances — for specific groups of students, subject matters, and teachers.…Second, class size reductions necessarily involve hiring more teachers, and teacher quality is much more important than class size in affecting student outcomes. Third, class size reduction is very expensive, and little or no consideration is given to alternative and more productive uses of those resources.

Similarly, in his paper, Dr. Krueger states,

> The effect sizes found in the STAR experiment and much of the literature are greater for minority and disadvantaged students than for other students. Although the critical effect size differs across groups with different average earnings, economic considerations suggest that resources would be optimally allocated if they were targeted toward those who benefit the most from smaller classes.

It is difficult to imagine that Dr. Krueger would disagree with Dr. Hanushek's statement, or that Dr. Hanushek would disagree with Dr. Krueger's.

Too often, scholarship in education debates is converted into simplified and dangerous soundbites. Sometimes liberals, particularly in state-level controversies about the level, equity, or adequacy of per-pupil spending, seem to permit themselves to be interpreted as claiming that simply giving more money to public schools, without any consideration to how that money will be spent, is a proven effective strategy. In contrast, conservatives sometimes permit themselves to be interpreted as claiming that money makes no difference whatsoever, and that schools with relatively few resources can improve sufficiently simply by being held accountable for results.

But surely the debate should not be so polarized. All should be able to agree that some schools have spent their funds effectively, and others have not. All should be able to agree that targeting the expenditure of new funds in ways that have proven to be effective is far preferable to "throwing money at schools" without regard to how it will be spent. All should be able to agree that there is strong reason to suspect that minority and disadvantaged children can benefit more than others from a combination of smaller class sizes and more effective teachers. And all should be able to agree that much more research is needed to understand precisely what the most effective expenditures on schools and other social institutions might be if improving student achievement, and narrowing the gap in achievement between advantaged and disadvantaged children are the goals.

It is difficult to avoid the conclusion that continued debates about whether money in the abstract makes a difference in education, without specifying how it might be spent, are unproductive. Equally true, denying that specific resource enhancements, alongside policy changes, can be an essential part of any reform agenda is also unproductive. Hopefully, the Krueger-Hanushek dialogue that follows can help to focus future debates on where spending is more effective. And it can add a new dimension to these debates, by proposing a comparison of the longer-term economic benefits of school spending, compared to its costs, that has barely begun to be explored.

Endnote

1. Indeed, other researchers are starting to examine both the costs and the benefits of policy interventions such as lower class size. Doug Harris (2002) uses a simulation model to estimate the "optimal" use of resources, considering teacher salaries and class size. Other researchers have examined the return on class size relative to other interventions.