“Our goal with conferences such as this one is to open up . . . dialogue to achieve an effective delivery of education to all of our children—and, by doing so, to build innovation and prosperity for new generations.”

— Sandra Pianalto

Innovation in Education

A Federal Reserve Bank of Cleveland Research Conference
Proceedings of a Conference Held in Cleveland, Ohio, November 17–18, 2005

Conference Overview and Welcoming Address

Sandra Pianalto

On behalf of Mark Sniderman, senior vice president and our director of research, and all of our staff who have been involved in putting this conference together, I am delighted to welcome you to our second conference on education. I am pleased to see this great turnout and to see many familiar faces from our education conference last year.

We decided last year that a lot can be gained by bringing educators and civic leaders together with the economists. And this year we have done the same.

I think it is important that the lessons we learn, such as the ones we will hear over the next two days, be heard by all of the interested parties: school superintendents, politicians, teachers, union representatives, and parents. It is also important for the economists involved in this research to hear from the practitioners. I am also pleased to have an opportunity to launch two days of inquiry and discussion on a topic of such vital national and regional importance: education.

I spend a lot of time talking with civic and business leaders about our region's economy, and I find that those conversations always end up on education and its importance to our region's economic growth and development. If you attended last year's conference on education and economic development, you may recall that we discussed some of the inputs and outputs that apply to education.

We began a dialogue on how to think about educational inputs like teachers, books, and computers, and then consider their combinations in terms of things like school-board policies, curriculum, and instructional methods. The big challenge, of course, is to find ways to achieve success in important outputs such as graduation rates and achievement test scores. While that is a very basic framework, it turns out that it generates a lot of useful insights about how education is actually produced and how people who care so much about education might work together to improve the product.

This year, our conference will focus on the role of innovation in education.

I would like to provide a framework to get us started as we address this issue. First, I will provide a brief overview of the conference. Second, I want to point out that the longer-term benefits innovation can bring about are often obscured by short-term turmoil. Third, I will explain why it is important to understand the incentives that new education policies or programs may generate.

RECENT INNOVATIONS IN THE EDUCATION INDUSTRY

Let me begin, then, with a conference overview. We decided to focus the conference on innovation this year because we want to call attention to some of the interesting new ideas that are being implemented regionally, nationally, and internationally. Some of these ideas dovetail with what we learned last year. For example, we learned from last year's conference that many people in our region recognize the need for change, and they want to know more about the changes taking place elsewhere. We presented some research on how local school boards can be made more accountable for student outcomes. Choice and competition can go a long way in this regard, and we will likely hear more about these topics when my friend and colleague, Michael Moskow, the president of the Federal Reserve Bank of Chicago, speaks to us at lunch today. In fact, today's entire program focuses squarely on the issues of competition and choice.

Last year, we also presented some research suggesting that student achievement could improve

when teacher compensation is tied to student performance. This seems to be a very timely idea. Just a few weeks ago, on November 1, the voters in Denver, Colorado, voted to implement an innovative approach for taxing and funding an incentive-based compensation system for teachers. We are fortunate to welcome Brad Jupp, who was intimately involved in that process from the union side of the bargaining table. We will hear from Brad tomorrow at lunch.[1]

Last year, we also had a lively discussion about the returns to investment in early childhood education. It turns out that it is not very difficult to make a case for expanding access to early childhood programs. What is difficult is figuring out how to design and fund the programs. At what age should children begin? What kind of services should be offered? Where will the money come from? Art Rolnick, the research director at the Federal Reserve Bank of Minneapolis, has a specific proposal that he would like to "sell" to Minnesota, but I am sure that he thinks that we should adopt his proposal right here in Cleveland as well. Art will fill us in on the details of his proposal tomorrow.

Also in tomorrow's sessions, we will hear about new research on what drives students to attend school, to learn, and to achieve.

While there is a lot of continuity with last year's conference, this year's conference stands on its own by emphasizing innovation. We know that when innovations are adopted, they require change, and that change can come fast and hard.

Here in Ohio, we are in the early stages of learning about the pros and cons of school choice with vouchers and charter schools. These alternatives arose from a general unhappiness with the status quo. We are continually reminded of this through the media. Hardly a day goes by that we don't see some kind of report informing us of the sad state of affairs of education in the United States. The bottom line, typically, is that our test scores have been falling behind the rest of the world. Or we hear about our local schools performing worse than schools in other districts and that it is becoming increasingly difficult to pass school levies. However, despite the willingness of states and communities to experiment, lately we have been hearing that some of the innovations may not be working quite as well as we had hoped.

INNOVATION CAN BRING SHORT-TERM UPHEAVAL

This brings me to my second point: Innovation can bring upheaval. This upheaval makes it difficult, at least in the short run, to fully appreciate the benefits we receive from that innovation. Implementing a new technology or a new idea always disrupts the status quo. It takes time to work out the kinks. Old technologies and processes must give way to new ones, but normally this does not happen without a fight.

For innovation to actually make a difference, we should expect to see new technologies, new organizational practices, and sometimes even new occupations. We may not capture the full returns from the innovation for many years. For example, as a society, we have only recently started reaping the benefits that microprocessor, telecommunications, and biosciences companies have delivered by working with innovations that were developed a number of years ago. Perhaps we are seeing the same process with innovation in education.

For as long most of us can remember, choosing a new school for our children to attend meant moving to a new school district or paying for a private school. Local property taxes were the main source of school funds. But these models for organizing and funding education are no longer being taken for granted, and new ideas are coming forward. I think people are looking for, and are ready for, change.

Yesterday I happened to be meeting with some business people from Athens, Ohio, and the subject inevitably turned to education. I commented that so many of our current educational practices can still be traced back to the system we developed in response to our needs as an industrializing economy, 100 years ago. One of my guests politely said, "Sandy, our school calendar is still based on the planting, growing, and harvesting cycle of the agricultural economy of 200 years ago!"

We have seen tremendous change in almost every industry and institution in our economy. I think that we might finally be at the front end of a revolution in primary and secondary education.

The challenge we face today is to try to accelerate the reorganization process that accompanies change by

learning from the past and paying closer attention to the role that incentives can play.

THE INCENTIVES CREATED BY INNOVATION

This brings me to my last point, which concerns incentives. Any policy, law, or contract creates incentives. For example, I think we all understand the motivation behind the movement toward standards and accountability in our schools, but how many of us anticipated the lengths to which some administrators, teachers, and students might engage in unethical activities in an effort to evade these standards?

Today, there are more choices and more types of schools. In some areas, vouchers can be used so that kids can change schools without moving to a new location. The success of charter schools and voucher programs in competing with public schools depends on a number of incentive mechanisms operating effectively. Schools must be willing to offer something of value to the marketplace, and at a price their customers can pay. Parents must be able to differentiate among the suppliers in the market and must find a way to get their child to the school of their choice.

The emergence of charter schools leads us to ask many questions, such as: Who should be allowed to get a charter? Should we regulate these schools? How will we measure their success or failure?

Governments must establish taxes and subsidies that lead to socially desirable outcomes. The location of schools, and the cost of attending them, can affect housing choices and tax rates in various communities which, in turn, will have repercussions on other aspects of community life. We should work aggressively to resist different educational outcomes based on race, class, and geography. Incentives matter, and they often have unintended consequences.

You will hear about several studies today and tomorrow that illustrate these points very clearly. As we begin our two-day conference, I am excited about what we will be able to learn together. As I said earlier, it is important to bring together all of the interested parties. It is important that economists hear from the practitioners and to understand the problems they face. It is important that politicians hear from economists to learn the likely incentive effects of proposed policies.

Our goal with conferences such as this one is to open up that dialogue to achieve an effective delivery of education to all of our children—and, by doing so, to build innovation and prosperity for new generations.

ENDNOTE

[1] Brad Jupp's conference presentation, while not included in these proceedings, is available online at *www.clevelandfed.org/Research/EdConf2005/Nov/papers.cfm*.

Lessons from Private-School Vouchers in Colombia [1]

Eric Bettinger

Education vouchers are one of the most politically divisive issues in the United States. Debates over the constitutionality and potential effectiveness of vouchers have taken place at all levels and branches of government. In the United States, the underlying argument in favor of vouchers centers on the effectiveness of public education, particularly in poor, urban areas. Voucher proponents claim that poor, urban schools are failing, that private schools may provide a better education to their students, and that vouchers will increase competition between schools and subsequently raise performance in all schools. On the other hand, critics argue that vouchers can lead to additional funding problems in urban school districts and may siphon the best students away from public schools.

In contrast to the debates in the United States, discussions about vouchers in developing countries center on the supply of and access to primary and secondary education. In developing countries, private schools make up a much larger share of educational providers than in the United States. Additionally, the public infrastructure is often unable to meet the demand for public schools. Such was the case in Colombia in 1991.

In the early 1990s, secondary-school enrollment rates were extremely low for the poorest children in Colombia. Only 55 percent of eligible students actually attended secondary school. The gap in enrollment rates between the poorest and richest students in Colombia was almost 35 percent (Sanchez and Mendes 1995). At the same time, Colombia's public school system was at its capacity. The school day was only four hours long, and 98 percent of public schools hosted multiple sessions per day; 20 percent of schools hosted three sessions per day in the same building. Forecasts from the World Bank (1993) suggested that Colombia's public schools were not prepared to handle the increase in enrollment that would occur over the next decade.

The voucher debate in Colombia was not about competition; it was about how to increase the capacity of the country to promote and manage additional enrollment. The central idea was to exploit excess capacity in private schools by allowing education vouchers for students enrolled in public schools. The resulting shift of students from public to private schools could create additional spaces in the public sector and potentially lead to an expansion of overall enrollment.

The program, entitled the Plan de Ampliación de Cobertura de la Educación Secundaria (PACES), began in 1991 and offered vouchers to students entering sixth grade, the start of Colombian secondary school. The vouchers were only available to the poorest of Colombia's population; applicants had to present evidence that they lived in a poor neighborhood.

Students were only eligible if they had been attending a public primary school, and they had to be accepted at a private school prior to their application. Elite private schools did not participate in the program, but studies have shown that the private schools that accepted the vouchers had similar pupil–teacher ratios, test scores, and access to technology (King, Rawlings, Gutierrez, Pardo, and Torres 1997). If students were selected to receive a voucher, they could renew it each year through graduation as long as they did not repeat a grade.

By 1997, PACES was one of the world's largest private-school voucher programs; over 125,000 vouchers had been awarded. While PACES was large relative to other voucher programs, it was small relative to Columbia's overall secondary school system. In 1995, approximately 3.1 million students

attended secondary schools in Colombia, with roughly 37 percent of students in private schools.

From the beginning, the demand for PACES vouchers far exceeded the supply. As a fair way to allocate vouchers when there was excess demand, the use of lotteries—one of the distinctive elements of PACES—was implemented. These lotteries created a laboratory with which to evaluate the effectiveness of the vouchers. Similar to a randomized trial in medicine, the voucher lottery created "control" and "treatment" groups. Students who applied to the voucher lottery and did not win formed an unbiased comparison group for students who did win, and comparing the academic and non-academic outcomes of students involved in the voucher lottery shows the effects of the voucher program.

There have been two major studies utilizing these voucher lotteries to measure the effects of PACES. The first study was conducted by Josh Angrist, Erik Bloom, Elizabeth King, Michael Kremer, and me (Angrist et al. 2002). Using survey data, we examined the effects the use of vouchers had after three years on students who had applied for the vouchers in Bogotá in 1995. As a longer-run follow-up to the first study, Josh Angrist, Michael Kremer, and I focused on high school graduation and college-entrance-exam data for these same students (Angrist et al. forthcoming). In the remainder of this paper, I review the evidence from these studies and discuss some of the lessons that the Colombia voucher program sheds on education in both developed and developing countries.

EFFECTS AFTER THREE YEARS

For the first study, conducted in 1998 and 1999, we surveyed almost 3,000 students who had applied for PACES vouchers in selected cities throughout Colombia. Our survey included questions examining students' education histories, their siblings' subsequent education experiences, and students' non-academic outcomes. Because of the randomness of the voucher, we only need to compare the average outcomes of voucher winners and voucher losers to learn about the impact of the program.

One of the most obvious outcomes of the program was its effect on private-school attendance. Because students had to be accepted at a private school prior to the voucher lottery, most applicants had a preference for private schooling. Among students who won the voucher, 96 percent attended private school that year; among students who applied for and did not win the voucher, 90 percent attended private school the next year regardless. While lottery winners stayed in private schools, applicants who did not win began leaving private schools in grades seven and eight. By the time of our survey (eighth grade), only 54 percent of voucher lottery losers were in private schools, compared to 70 percent of lottery winners.

One of the program's interesting effects on attendance patterns occurred among students who had applied to vocational private schools prior to the voucher lottery. These students who did not win a voucher behaved quite differently than students who did. Voucher lottery winners attended the vocational schools to which they had applied; lottery losers, by contrast, transferred to academic schools. If one measures school quality by the educational outcomes of its students, vocational-voucher lottery losers attended *better* schools than the vocational-voucher lottery winners.

We generally found that lottery winners had better educational outcomes than lottery losers. Although lottery winners only completed about one-tenth of a year more of school than lottery losers and were just as likely to drop out, one big difference did stand out: repetition rates. About one in five lottery losers had to repeat grades six or seven. Among voucher winners, only one in seven students repeated. The difference translates to a 25 percent reduction in repetition rates. In terms of test scores, voucher winners scored about one grade-level higher on standardized exams than voucher losers.

Another striking finding involved the program's effects on non-academic outcomes. Colombia, like many developing countries, has a substantial number of youths working outside the home. Voucher lottery winners were less likely to be working outside the home and worked fewer hours on average than lottery losers. Given that there was virtually no difference in drop-out rates between these groups of students, this suggests that vouchers alleviated working commitments of voucher winners, potentially freeing up time that they could devote to their studies. Additionally, we tracked whether students were married or cohabitating with a significant other at the time of our survey (approximately age 15): Lottery winners were less likely than lottery losers to be involved in such a relationship.

And what of the voucher winners who attended vocational schools? They had less grade repetition, completed more years of schooling, and worked fewer hours outside the home than students who applied for the vocational voucher and did not win. Given that these voucher-winning students attended schools that appeared inferior academically, this finding may have some interesting implications for how the voucher program affected students.

LONG-RUN EDUCATIONAL EFFECTS

In Colombia, 90 percent of graduating seniors take the college entrance exam. With the cooperation of the national testing office (ICFES), we were able to match applicants' records to their subsequent exams. As before, we only needed to compare the average outcome for lottery winners and losers to identify the effects of the vouchers.

Of the students who initially applied for the voucher, almost one-third eventually took the college entrance exam. Lottery winners were about 5 to 7 percentage points more likely than lottery losers to take the exam. The voucher hence increased their likelihood of taking the college entrance exam by about 20 percent.

Angrist et al. (forthcoming) also describe the effects of the voucher program on college-entrance-exam scores. (Differences in exam scores are harder to interpret because the voucher program induced many students to take the exam who would not have otherwise done so.) When we compare the raw scores, we find that lottery winners have higher language scores than lottery losers. Moreover, using a variety of econometric methods, we attempt to identify the effects the voucher program has on test scores of both average and high-achieving students who applied to the voucher lottery. We find that the program improves test scores at the mean as well as for the highest achievers (i.e., students over the 90^{th} percentile).

Again, when we look at applicants who applied to vocational schools, we see that lottery winners had higher test scores and a greater likelihood of taking the college entrance exam than lottery losers. This is interesting, given that vocational-voucher winners were more likely than vocational-voucher losers to attend vocational schools, whose records for getting students to take the ICFES and whose students' subsequent performances on the exam are inferior to those of academic schools. Voucher winners at vocational schools did well despite their less-academic surroundings.

DISCUSSION AND CONCLUSION

The voucher program in Colombia proved to be very successful for students who were able to participate. Not only did their academic outcomes improve, but so did many of their non-academic outcomes. (Evidence in the United States has been less clear about the effect of voucher programs on students' success.) In the case of Colombia, the answer is unambiguous—voucher winners' outcomes improved relative to what they would have achieved in the absence of the voucher.

The voucher program also improved outcomes for students at vocational schools. Among students who applied to vocational schools prior to applying for the voucher, the program increased the likelihood of attending vocational schools. Yet these schools were inferior according to most academic standards. Why did these students do so well? Ongoing work by Bettinger, Kremer, and Saavedra (2005) investigates this. One potential explanation for the findings is the effect that the voucher program had on students' incentives; students lost the voucher if they did not pass a grade. The threat that the government would cancel a student's voucher may have been enough to persuade the student to work harder in school. A recent series of economic papers have focused on understanding the effects of incentives on students (e.g., Angrist and Lavy 2002; Kremer and Miguel 2004). Families and students respond to financial incentives, and these incentives may lead to better health and educational outcomes.

In 1998, Colombia dissolved the voucher program. A new administration did not see the program as central to its educational initiatives and, at the time, there was no measure of the effectiveness of the program; the first evidence of its effectiveness didn't come until 2002. As a result of the evidence on PACES, Colombia has renewed discussions with the World Bank about whether it should restart a voucher initiative.

The Colombia voucher program provides a valuable social lesson: A randomized experiment can influence education policy. In recent years, the United States government has emphasized the role of evaluation of randomized policy interventions. Randomization is the "gold standard" in evaluation. While other methodologies can provide hints of the effectiveness of a

policy, they generally cannot provide irrefutable evidence. Randomization, however, can provide definitive evidence when the randomization is conducted properly. In the case of the Colombian vouchers, randomization was used to assign the voucher to attain fairness: more people wanted the vouchers than there were vouchers available. Rather than assign vouchers based on previous performance, wealth, or other characteristics, randomization ensured that each applicant had an equal chance.

The definitiveness of the Colombian evidence has attracted the attention of policymakers and academics alike. It has shown that vouchers do improve outcomes for students. It has also demonstrated, and provided a model for, how policies can be implemented using randomization and how the resulting evaluations can provide definitive evidence of a program's efficacy.

ENDNOTE

[1] This paper is largely based on "Vouchers for Private Schooling in Colombia: Evidence From a Randomized Natural Experiment" by Joshua D. Angrist, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer (*American Economic Review* 2002) and "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia" by Joshua D. Angrist, Eric Bettinger, and Michael Kremer (*American Economic Review*, forthcoming).

REFERENCES

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review*, 92 (5): 1535–58.

Angrist, Joshua, Eric Bettinger, and Michael Kremer. Forthcoming. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review*.

Angrist, Joshua, and Victor Lavy. 2002. "The Effect of High School Matriculation Awards: Evidence from Randomized Trials," NBER Working Paper Number 9389. Cambridge, Mass.: National Bureau of Economic Research.

Bettinger, Eric, Michael Kremer, and Juan Saavedra. 2004. "How do Vouchers Work? Evidence from Colombia," mimeo, Case Western Reserve University.

King, Elizabeth, Laura Rawlings, Marybell Gutierrez, Carlos Pardo, and Carlos Torres. 1997. "Colombia's Targeted Education Voucher Program: Features, Coverage and Participation," Working Paper No. 3, Series on Impact Evaluation of Education Reforms, Development Economics Research Group, The World Bank, September.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2004. "Incentives to Learn," NBER Working Paper Number 10971. Cambridge, Mass.: National Bureau of Economic Research.

Sanchez, Fabio, and Jairo Mendez. 1995. "Por Que los Niños Pobres No Van A La Escuela? (Determinantes de la asistencia escolar en Colombia)," mimeo, Departmento Nacional de Planeación Republica de Colombia.

The World Bank. 1993. *Staff Appraisal Report: Colombia, Secondary Education Project*, Latin America and the Caribbean Region, Report 11834-CO.

Education Vouchers and the Cleveland Scholarship Program [1]

Clive R. Belfield

Vouchers are a method for financing government services by which clients are given coupons of prescribed value for use at any eligible service provider. Vouchers for K–12 education continue to attract interest, offering the promise of greater parental choice, enhanced school efficiency, and improved educational outcomes for students.

The first formal voucher program was established in Milwaukee in 1990; by 2003–04, 12,778 students were participating across 107 schools. Its practical success was followed by programs developed in Cleveland, Florida, Colorado, and Washington, D.C. (see Belfield and Levin 2005). At their most basic, these programs offer an "existence proof" for vouchers. With them has come sustained academic inquiry into education vouchers and their anticipated effects.

Here we summarize this evidence to establish stylized facts about vouchers. The evidence is then used to inform evaluation of the Cleveland Scholarship Program (CSP), including its effects on students' test scores. Established in 1995, the CSP now has particular prominence because, in 2002, the U.S. Supreme Court ruling *Zelman vs. Simmons-Harris* resulted in federal approval of inclusion of religious schools in voucher programs. Although operating for almost 10 years, the status of the Cleveland program had still been uncertain; with the legal challenge resolved, the CSP is now affirmed, and more vouchers are being offered across all school grades. In 2005, Ohio ratified a statewide version of the CSP to be introduced in 2006: This will provide as much as $5,000 each for up to 14,000 students enrolled in schools that receive the state's lowest performance ranking for three straight years. Evidence on the efficacy of the program is therefore critical, both for the direct development of policy in Ohio and for voucher initiatives in other states.

NEW RESEARCH ON EDUCATION VOUCHERS

Research on vouchers has not only examined their academic benefits; it has also considered how parents choose schools, how schools operate, and how vouchers influence public finances. These investigations are useful for informing program design and for setting vouchers within the broader context of school-choice reform.

In looking at school-choice decisions by parents, it is clear that many affluent families already have choices; attention therefore focuses on how voucher programs might open up choices for those families who are constrained. Thus far, all voucher programs have been targeted to low-income families or to districts with low-performing schools. Clearly, choice sets are being expanded for low-income families. However, several mediating factors make voucher programs less equitable in actuality than is implicit in a simple reading of the program eligibility rules. First, religion pervades family preferences of schools (Campbell et al. 2005). Certainly, parents value high test scores, but preferences are varied, with many families choosing their neighborhood school. Second, race has a strong influence on school-choice decisions. This relationship is complicated by different patterns in Hispanic and African American children and by the fact that public schools show strong patterns of racial segregation. However, vouchers consistently lead to greater student segregation (even in a highly minority public school system such as Washington, D.C.). Third, school choice is an action; that is, parents must actually use the voucher to change their child's school, conditional on eligibility and having new choices. Consistently, usage rates are much lower—perhaps by one-third—than offer rates (Howell and Peterson 2002, table 2). Moreover, even within low-income groups, the children most likely to succeed in school are the ones most likely to utilize the voucher (as are

white students). Also, a nontrivial proportion of those who receive vouchers are already in private school.

Research has also investigated the supply of private schooling. Without supply, family preferences become meaningless, and, if private schools do not operate in ways distinct from public schools, there is no advantage to students choosing them.

Several consistent findings emerge from the research on supply. Most participating schools are religious (across various faiths); secular schools have a small market share. The supply of new schools appears reasonably elastic: For example, almost one-half of the schools participating in the Milwaukee program were founded after it was introduced. But voucher student enrollees are increasingly a majority within their schools: by 2001, 40 percent of the schools participating in the Milwaukee program had more than 80 percent of their students claiming vouchers. However, research on the inputs and technologies that private schools use (beyond selecting their students) and on which are more efficient has yielded very little: Economists are still no further ahead in identifying the separate benefits of ownership, innovation, and technical efficiency, that is, which inputs work best (Hanushek 2004). After controlling for student characteristics, most research finds only very modest advantages to private school. Finally, greater competition is likely to improve schools' performance (Belfield and Levin 2002), but the competitive pressures exerted by small-scale voucher programs are also likely to be modest.

Further research has examined how voucher programs might impinge on the existing public school system and its local financing. Because schooling is financed locally, individual school choices will feed back into house prices and district spending and, therefore, school quality. Nechyba (2003) reports several novel results from modeling the introduction of large-scale voucher programs. First, competition for high-ability students would increase; these students would pay lower tuitions, both as a result of the voucher and as a consequence of schools' greater eagerness to enroll them. Second, public schools would engage in more ability tracking to prevent high-ability students from switching to private schools. Both factors suggest further educational inequalities, with greater rewards (and resources) for high-ability students. Third, public-school quality is most likely to decline in middle- or high-income school districts. Importantly, opinions about vouchers are likely to be driven more by perceived effects on property values than by educational outcomes; homeowners may be wary of education reforms that may raise uncertainty as to the value of their home.

The most high-profile research on vouchers has looked at whether they raise student achievement. The evidence here shows, at best, moderate advantages for voucher participants.

For the Milwaukee Program, Rouse (1998) found small but positive effect-size differences of 0.08–0.12sd per year for math but no effect for reading. However, the data were from the first five years of the program—religious schools were not participating, and the voucher students were concentrated in a few schools. For the Florida program, Figlio and Rouse (2005) found modest results from data on over 180,000 students. Voucher users in initially low-performing schools do post higher test scores, but much of this is attributable either to student characteristics or to teaching to the high-stakes test. Randomized field trials for vouchers in three cities found small test-score gains after three years (Howell and Peterson 2002). These treatment effects were primarily for African Americans in one setting, with no evidence of cumulative gains for those who used the voucher for the longest periods. Finally, it is worth noting that new evidence from expanded public-school choice points to the same conclusion, with few achievement gains from placement in a choice school (Cullen et al. 2005).

CLEVELAND SCHOLARSHIP PROGRAM

The Cleveland Scholarship Program operates in the Cleveland Municipal School District (CMSD), which has 75,000 students across 130 schools. Eligible schools are nonpublic, chartered schools located within the CMSD and approved by the state superintendent. Surrounding public-school districts are eligible to apply, and 5,734 students currently participate in the program. Initial enrollment in 1996 was 1,996, with total funding of $5 million. Eligible children had to be in grades K–8, reside within the CMSD, and not require segregated special education. Low-income families were given preference, with those below 200 percent of the poverty level given 90 percent of tuition or $2,250, whichever was lower;

families above 200 percent of the poverty level were given 75 percent of tuition or $1,875, whichever was lower. About one-fourth of the students came from the latter group. In comparison, per-pupil expenditure in CMSD in 1996 was $7,500 (including transport). In 2003–04, scholarships were made available for ninth grade and beyond, and funding was increased to $3,000 for grades K–8 and set at $2,700 for higher grades. Where voucher applications exceeded available placements, a lottery system was used.

Despite its usefulness for informing future voucher reforms, the Cleveland Scholarship Program has been the subject of little academic inquiry. The program is large enough to allow for samples of students according to voucher status and religious schooling, and with recent data, it offers an up-to-date evaluation of vouchers in light of recent school-choice reforms. Given the duration of the program, it is possible to examine the question of dose response (i.e., whether persistence in the program yields higher rewards). Also, CMSD has a high proportion of African American students, for whom vouchers are held to be most beneficial. Importantly, the CSP voucher is relatively ungenerous: if effects can be found for this program, it is likely that achievement gains would be even larger for more generous programs.

To identify effects, it is necessary to classify students according to voucher status: *users* (offered and used a voucher to attend private school); *non-users* (offered a voucher but did not use or stopped using it); *applicant rejects* (applied for but not offered a voucher); *eligibles* (could apply but did not); and *ineligibles* (not eligible according to program rules). Here, these last two groups are conflated into a general public-school comparison group.

The potential for bias in identifying effects from vouchers is high. Three kinds of bias are particularly important, but their effects are (probably) offsetting for this program. Applicant bias occurs when only those who apply for the voucher are likely to benefit from it. This will bias gains toward users because applicants are typically from motivated families. Eligibility bias occurs when those who are eligible differ both from those who apply and (separately) from those who do not apply. This will bias gains away from users, because CSP eligibility is conditional on low family income. (It is possible with these data to control for eligibility). Usage bias occurs when those who use the voucher differ from those who do not use it, conditional on application and eligibility. This will bias gains in favor of users relative to non-users because usage is positively correlated with ability and family resources. In addition, data collection inevitably generates some response bias as survey attrition rates are higher for non-users.

EVALUATING THE CSP

Since 1996, the CSP has been evaluated by the Center for Education Evaluation at Indiana University (see Metcalf et al. 2003). The evaluation began with collecting data for those entering kindergarten in 1997 and has continued up to sixth grade in 2003. (No substantive changes in program design occurred during this period, but the legal status of the program was only resolved in 2002). The data set includes over 4,000 students who attend more than 100 separate schools. In terms of voucher status, the data set is composed of users (23 percent), non-users (10 percent), applicant rejects (16 percent), and a public-school comparison group (51 percent).

The data set has three advantages: it is longitudinal (including achievement measures); it includes students from multiple comparison groups; and it is a large sample spread across many different schools. However, the public group does not precisely conform to the above categorization: It includes both eligibles and ineligibles as part of a general comparison group. Also, the sample sizes are considerably lower because of attrition and missing responses.

Evidence on the CSP comports with extant research. Specifically, most students chose religious schools, and high (and growing) proportions of voucher applicants and users had previously been enrolled in private schools. Table 1 shows information on selected student characteristics across voucher status. This, too, conforms with other studies. African American students and low-income students (free-lunch eligible) are much less likely to use the voucher, conditional on being offered it.

To identify the achievement gains from the voucher program, we estimate a series of education production functions. The impacts of voucher status on achievement are reported in table 2. Terra Nova test scores in second and fourth grade are regressed against a set of student and school characteristics, including prior-year test scores (see table 2, notes).

Table 1: Descriptive Statistics by Voucher Status (Second-Grade Students)

	Voucher user	*Non-user*	*Rejected applicant*	*Public-school comparison group*
Black	56%	75%	58%	48%
Hispanic	7%	4%	6%	3%
Free-lunch eligible	58%	83%	50%	49%
N	*624*	*326*	*438*	*971*

These impacts are effect-size gains relative to the public-school comparison group in reading, math, and language.

The top panel of table 2 shows that voucher status has mixed effects. For reading, there are no statistically significant differences across the four groups in second grade. For math, voucher users report the lowest scores: the statistically significant effect size is –0.09sd against the public-school group and of comparable size against the other two groups. In contrast, for language, the public-school group does considerably worse than the other three groups; however, the voucher-user group gains the least—both non-users and rejected applicants show larger advantages. Given the biases that might lead to statistically significant gains for voucher users, we cannot find evidence that voucher students outperform relevant comparison groups in second grade. Moreover, the bottom panel shows that by fourth grade, the math penalty for voucher users persists, but the gains in language have dissipated to insignificance.

Additional testing using second-grade data affirms the weakness of any voucher effects. When we do not control for prior achievement, the results do not favor voucher users: The math penalty is maintained, and the language advantage over the public-school group is eliminated. Moreover, the rejected applicants report statistically significant test-score gains in reading and language. When we compare users only against those who were not offered a voucher (the treatment effect), users report scores that are lower by 0.14sd in reading and 0.11sd in math, with no difference in language. When we examine whether those who have participated the longest in the program obtain the strongest effects, we find mixed effects across the three subjects. Finally, we test for whether the impact of vouchers differs by race. Restricting the sample to African American students, voucher users appear more disadvantaged: Their reading scores are now statistically significant and lower (–0.14sd); the math penalty is still evident (but not statistically significant); and the language advantage is not discernible. These sub-analyses give no indication that vouchers have a differential and beneficial impact for African American students.

Overall, there is no clear advantage for voucher students; if anything, there is a slight academic

Table 2: Effect-Size Test-Score Gains over Public-School Students

	Reading	*Math*	*Language*
Second Grade:			
Voucher user	–0.060	–0.092**	0.097**
Non-user	–0.019	–0.021	0.162*
Rejected applicant	0.083	0.026	0.136**
N	*1733*	*1786*	*1736*
Fourth Grade:			
Voucher user	0.043	–0.113***	0.038
Non-user	–0.065	0.044	0.076
Rejected applicant	–0.030	–0.055	–0.063
N	*2089*	*2102*	*2085*

Notes: Terra Nova test. OLS estimation. Effect sizes control for subject-specific first- and third-grade scores; African American; Hispanic; female; free-lunch eligible; unsubsidized lunch; class size; and years of teacher experience. * significant at 10%; ** significant at 5%; *** significant at 1%.

penalty. However, given that in 2001 the value of the voucher was less than $2,400 and the opportunity cost in the public school system was approximately $8,800, it might be concluded that the voucher program is cost-effective. Yet, back-of-an-envelope calculations show that this conclusion is premature. The CSP voucher does not include transportation, which must be paid by the district, nor does it include standardized assessments, which public schools must impose. It does not cater to special-education students or the most disadvantaged students (who do not take up the voucher). And, because the program is sub-additive, for every three vouchers distributed, approximately one student would have attended private school anyway. There are also additional administration costs.

Finally, the $8,800 figure for CMSD is the average cost per student, not the marginal cost. With approximate costs of transport, assessment, special education, sub-additivity, and administration taken into account, the anticipated fiscal saving falls from $6,400 ($8,800-$2,400) to perhaps $2,500–$3,000. Although this is still a sizeable cost saving, it does not account for marginal costs, student disadvantage, or any reorganization costs.

CONCLUSION

Recent research sheds light on the efficacy of vouchers. Broadly, it may be questioned whether vouchers—even with some modest achievement gains—will be a catalyst for educational improvements. On the demand side, there are many steps before parents actually exercise choice. On the supply side, most of the participating schools are religious, with little evidence of new secular schools either opening or accepting vouchers. Competition will probably be muted. General equilibrium models explain why homeowning voters are wary about expanding voucher programs.

The Cleveland Scholarship Program readily fits into this general pattern, evincing similar features to voucher programs in Milwaukee, Florida, and now Washington, D.C. Although targeted at students from low-income families, these programs in fact serve those somewhat closer to the middle of the income distribution, when usage rates and prior schooling are accounted for. Students are highly likely to be in religious schools, and African American students are less likely to use their vouchers. Importantly, the CSP results are not encouraging with respect to achievement: The program shows a slight academic penalty for voucher users relative to other comparison groups. Thus, it seems unlikely that an expanded statewide program will radically enhance educational outcomes in Ohio.

ENDNOTE

[1] For a full version of this paper, see C.R. Belfield's "The Evidence on Education Vouchers: An Application to the Cleveland Scholarship and Tutoring Program," Working Paper, NCSPE, *www.ncspe.org*.

REFERENCES

Belfield, C.R., and H.M. Levin. 2002. "The Effect of Competition on Educational Outcomes: A Review of the U.S. Evidence," *Review of Educational Research* 72: 279–341.

Belfield, C.R., and H.M. Levin. 2005. *Privatizing Educational Choice*. Paradigm Publishers: Denver, CO.

Campbell, D.E., M.R. West, and P.E. Peterson. 2005. "Participation in a National, Means-Tested School Voucher Program," *Journal of Policy Analysis and Management* 24: 523–41.

Cullen, J.B., B.A. Jacob, and S. Levitt. 2005. "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries," *Journal of Public Economics* 89: 729–60.

Figlio, D., and C.E. Rouse. 2005. "Do Accountability and Voucher Threats Improve Low-Performing Schools?" NBER, Working Paper, 11597.

Hanushek, E.A. 2004. "What If There are No 'Best Practices'?" *Scottish Journal of Political Economy* 51: 156–72.

Howell, W.G., and P.E. Peterson. 2002. *The Education Gap. Vouchers and Urban Public Schools.* Washington, D.C.: Brookings Institution.

Metcalf, K.K., S.D. West, N.A. Legan, K.M. Paul, and W.J. Boone. 2003. *Evaluation of the Cleveland Scholarship and Tutoring Program*, Summary Report 1998–2002. Bloomington, Indiana: Indiana University School of Education.

Nechyba, T.J. 2003. "What Can Be (and Has Been) Learnt from General Equilibrium Simulation Models of School Finance?" *National Tax Journal* LVI: 387–414.

Rouse, C.E. 1998. "Private-School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113: 553–602.

School Size and Student Outcomes: A Nontechnical Paper [1]

Christopher Berry
Martin West

In the middle of the twentieth century, a not-so-quiet revolution remade American public education. As late as 1930, schools in the United States were small, community-run institutions, most employing but a single teacher. Over the next four decades, the number of schools fell by more than 100,000, as nearly two-thirds of all schools were eliminated through a process of consolidation. Average school size increased fivefold over this short period. In the process, school districts evolved into professionally run educational bureaucracies, some educating hundreds of thousands of students.

Despite the scale and pace of these changes in the organization of public education, little is known about the consequences of consolidation. Did the quality of public education rise as schools became larger and more professional, as proponents of consolidation promised? Answering this question takes on particular importance in the context of the contemporary "small-schools" movement, which includes education heavyweights such as the Gates Foundation, the Annenberg Foundation, and school systems in New York, Chicago, and other big cities. The historical experience with consolidation provides a valuable context for contemporary reformers because seldom have we seen such dramatic changes in school size over such a short period of time.

This paper aims to begin filling the gap in our understanding of the consequences of the consolidation movement. We use data from the Public-Use Micro-Sample of the 1980 U.S. census to estimate the effects of changes in school and district size on students' labor market outcomes and educational attainment. Our results indicate that students born in states with smaller schools obtained higher returns to education and completed more years of schooling. While larger *districts* were associated with somewhat higher returns to education and increased educational attainment in most specifications, any gains from consolidation were outweighed by the harmful effects of larger *schools*. Reduced form estimates of the effects of consolidation on labor market outcomes confirm that students from states with larger schools earned substantially lower wages later in life.

BACKGROUND

The consolidation of schools was part of a larger effort to professionalize education that began in the late nineteenth century (Tyack 1974). To the "administrative progressives" of the day, the concentration of authority over schools in the hands of professional educators seemed a cure for both the corruption of city school systems and the parochialism of rural ones. In imagining a professionally run school, reformers drew their inspiration from the modern corporation, with its principles of "scientific" management by experts. To these reformers, consolidated schools seemingly offered economies of scale in administration, instruction, and facilities.

It is clear that the impetus for consolidation seldom came from local communities. Local resistance to consolidation was often fierce, especially in rural areas where the school was the central institution of the community. In the face of local resistance, state governments often resorted to using fiscal incentives to induce consolidation or simply mandated consolidation by unilaterally redrawing district boundaries (Hooker and Mueller 1970; Strang 1987). "Defensive consolidation," in which districts rushed to consolidate in anticipation of a more radical plan proposed by the state, was also common (Reynolds 1999).

Few communities withstood the financial and political pressures for long. Figure 1, which is based on data from the federal government's *Biennial Survey of Education*, shows that the number of American public

Figure 1

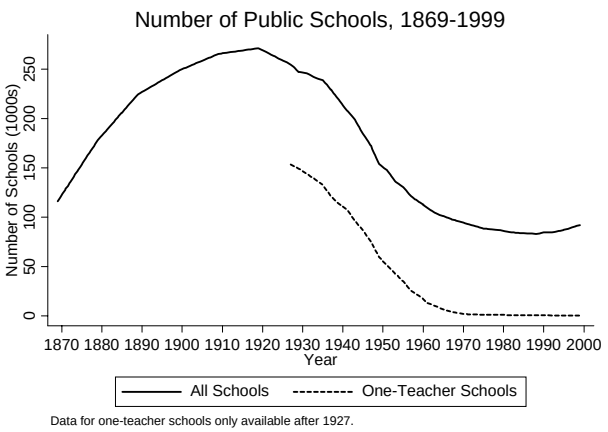
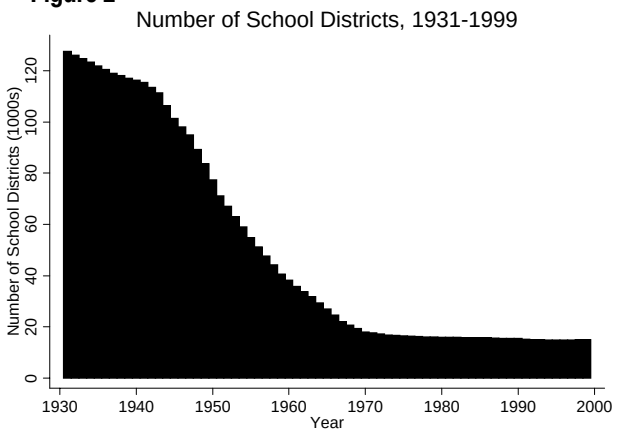


Figure 2



schools peaked at 217,000 in 1920 and declined rapidly over the succeeding 50 years.[2] The decline's pace slowed in the 1970s, and the number of schools reached a nadir in the late 1980s at around 83,000. Since then, approximately 10,000 schools have been added nationwide, in the first significant burst of (net) new school construction in over 60 years.[3]

The number of districts also declined dramatically from the 1930s to the 1970s. The earliest reliable data on the number of school districts in each state come from the 1931–32 edition of the *Biennial Survey* and show that the number of districts fell by half between 1931 and 1953, as over 60,000 districts were dissolved (Figure 2). It declined by half again between 1953 and 1963 and by yet another 50 percent over the following ten years. The number of districts stabilized in the early 1970s and has not changed appreciably since.

As schools and districts were consolidating, the number of pupils attending public schools was on the rise. From 1929 to 1969, average daily attendance (ADA) in public elementary and secondary schools doubled, rising from approximately 21 to 42 million students.[4] The combination of declining numbers of schools and districts and rising attendance produced substantially larger educational institutions over the course of the twentieth century. From 1930 to 1970—the period of most rapid consolidation—ADA per school increased from 87 to 440 students (see figure 3). At the same time, ADA per school district increased from approximately 170 to 2,300 students (see figure 4).[5] Both schools and districts witnessed their most rapid burst of growth in the years from 1950 to 1970, as increasing attendance rates, the baby boom, and institutional consolidation coincided.

As discussed above, school consolidation was part of a broader movement of school reform. Between 1930 and 1970, the school term grew longer, class sizes shrank, and teachers became better paid. The average state share of funding for public education more than doubled between 1930 and 1950, from less than 20 percent to roughly 40 percent, and made a smaller jump again in the late 1970s. The overall effect of these changes was to transform the small, informal, community-controlled schools of the nineteenth century into centralized, professionally run educational bureaucracies. The American public school system as we know it today was born during this brief, tumultuous period.

PREVIOUS RESEARCH

There have been two identifiable waves of literature on school size (Howley 1996). The first wave studies, appearing roughly from the 1920s through the 1970s, focused primarily on input measures of school quality.[6] Larger schools were consistently found to be superior in this regard, with better facilities, more qualified teachers and administrators, and a greater depth and variety of course offerings and extracurricular activities. The well-known Conant Report, in which former Harvard University president James Conant reported on a nationwide survey of 2,000 high schools, represents the high point of this first wave of literature (Conant 1959, 1967).

Beginning in the 1980s, the focus of the school-size literature shifted from school inputs to student

Figure 3

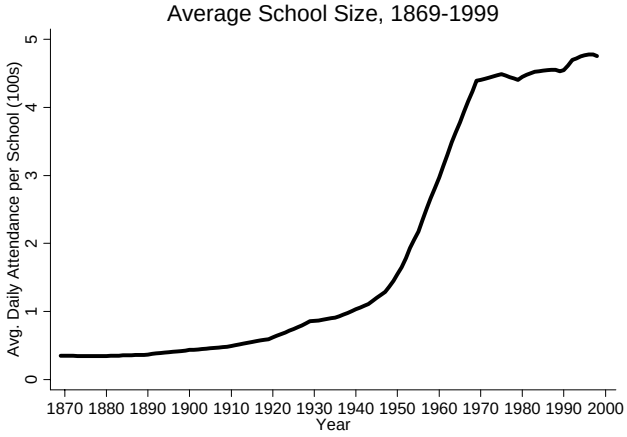
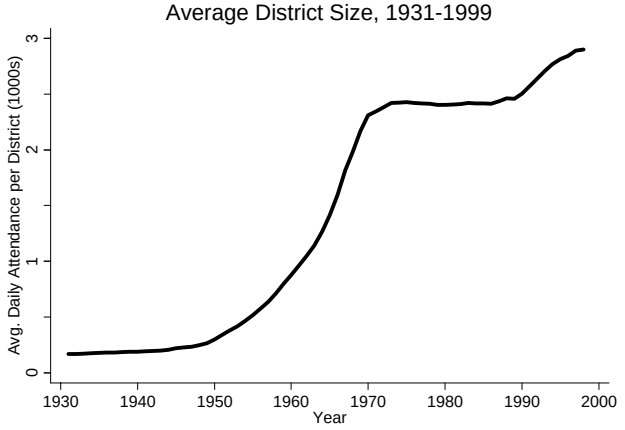


Figure 4



outcomes. This ongoing second wave of studies has been less favorable to large schools. In fact, six of the seven studies of school size and student performance reviewed by Andrews et al. (2002) found decreasing returns to scale.[7] Summers and Wolf (1977) find that African American students in particular are harmed by large school size, while Lee and Smith (1997) find that students of low socio-economic status do especially poorly in large schools. Although the reasons for the superior performance of students in small schools have not been identified, speculative explanations have focused on non-academic factors such as a greater sense of community belonging among students, closer interaction with adults, and more parental involvement (e.g., Cotton 1996).

ANALYTICAL FRAMEWORK AND DATA

A key shortcoming of the recent literature on school and district size and student outcomes is its general inattention to methodological challenges inherent in the estimation of economies of size. Schools and districts that are smaller than the norm likely share other unusual features that are not well measured by the variables included in standard education-production functions. Absent random assignment of students, they are also likely to draw a population of students that differs from the students in larger schools and districts in unmeasured ways. The expected direction of these biases due to unobserved differences is theoretically unclear. In addition, over time, highly effective schools and districts may attract more students, creating a general bias in observational studies toward finding increasing returns to size.

We chose the empirical approach taken in this paper with these issues in mind. Our empirical analysis uses the Public-Use Micro-Sample of the 1980 U.S. census to relate changes in school and district size during the consolidation movement to student outcomes in the labor market later in life. We focus in particular on the effects of consolidation on the *slope* of the relationship between earnings and education. That is, we examine how changes in school and district size affected the labor market value of an additional year of schooling. We implement this strategy in two stages: In the first stage, we identify the state-of-birth-specific component of the return to education, and in the second stage we relate these state-of-birth-specific returns to characteristics of each state's public schools.[8]

By relating differences in state-average district and school size to long-term outcomes of students raised in that state, we avoid the problem of families choosing where to live based on the size or quality of the school. (It seems implausible that families would choose to move to a new state because of the average size of its schools.) We also restrict our analysis to within-state, over-time variation (to eliminate the influence of time-invariant state characteristics) and control for a range of institutional and demographic variables likely to be associated with changes in size. Finally, although we present reduced form estimates of the relationship between school and district size, our main analysis centers on the return to an additional year of schooling, which appears less likely to be affected by unmeasured background characteristics (Card and Krueger 1992).

RESULTS

The Rate of Return to Education

Using this analytical approach, we found that smaller schools had a significant positive effect on students' wages as adults. Moreover, the effects are substantial. Our findings suggest that an increase of one standard deviation in average school size is associated with a decrease of 1.23 standard deviations in the rate of return to education. Put differently, increasing a state's average school size by 145 students, equivalent to the difference in average school size between the median state in the 1920–29 cohort and the median state in the 1940–49 cohort, is associated with about a 9 percent decline in earnings for high school graduates (those with exactly 12 years of education).

In the same analysis, we also found a positive effect of large district sizes on students' adult wages. In other words, the results suggest that larger schools were detrimental—whereas larger districts were beneficial—to the return to education. An increase of district size by 947 students, again the difference in average size between the 1920–29 and 1940–49 median states, is associated with a 2.1 percent increase in earnings for high school graduates. However, the findings on district size were not robust enough to further analytical checks, so we are cautious about putting much weight on the positive effect of larger districts.

A concern with these results is that consolidation did not occur randomly. As discussed above, consolidation was one of a series of progressive-movement reforms in American politics that centralized and professionalized public education. The timing of consolidation varied across states, and it is possible that states whose residents put a higher priority on education also embraced reform more readily. Thus, one might be concerned that the effects of consolidation reported above merely reflect the influence of being raised in a community that places a high value on education. Because we cannot directly measure the value placed on education in a state, and therefore cannot control for it statistically, it is possible that these unobservable background characteristics bias the estimated effects of consolidation.

With respect to the findings on school size, the influence of such unobservable background characteristics is less of a concern. According to prominent education experts of their day, such as Ellwood Cubberley and James Conant, larger schools were better. Moreover, we also find a positive correlation between the income of the parents' generation in a state and the average school size. Therefore, there is ample reason to believe that education-minded parents would have had a preference for larger schools, not smaller ones. If early community influence was really driving the results, increases in school size would be associated with higher wages. But we find just the opposite. In addition, when we control for state background characteristics, such as parental income and percent rural, our estimates of the effects of school size do not change significantly. That school size displays a negative relationship with educational returns, contrary to the relationship with district size *and* contrary to the expectations of contemporary education experts, bolsters the interpretation of this as a causal relationship rather than as an artifact of unobserved early community influences. Thus, the findings on school size appear to be relatively safe from concerns about unobserved community influences.

On the other hand, the findings on district size are open to question. Because consolidation was pushed by education reformers, the positive association between district size and wages might simply reflect the effects of being raised in a state where parents placed a higher value on education. In addition, the district-size results are not significant in alternative statistical models.

Educational Attainment

The results discussed thus far pertain to the returns to education, that is, the wage increase that a worker can expect to earn for each additional year of schooling completed. A related issue is what effect consolidation had on the average number of years of education completed. For instance, if consolidation discouraged students from staying in school, then the additional return to education would be offset by reduced attainment. To get at this issue, we estimated models of the school-size effect on educational attainment using an analytical approach similar to that described previously. We find students born in states with larger schools also completed significantly fewer total years of schooling. This relationship persists even after controlling for the percentage of the state's population that is rural and for average parental-generation income. For instance, increasing a state's average school size by 145

students, equivalent to the difference in average school size between the median state in the 1920–29 cohort and the median state in the 1940–49 cohort, is associated with a decline of about 0.12 years of completed education. In other words, smaller schools were associated with more years of education completed as well as with a greater labor market return to each year of education.

SUMMARY AND POLICY IMPLICATIONS

Our results indicate that students educated in states with smaller schools earned higher wages as adults by both completing more years of schooling and earning a higher return to education. Thus, our study provides general support for the growing small-schools reform movement in contemporary education circles.

Implications for contemporary education policy, however, must be drawn only with caution from the analysis presented here, for several reasons. First, we have not examined any school- or district-size data more recent than 1966. Second, the findings pertain to state *average* school size. One must therefore be cautious in trying to ascertain the "right" size for any individual school or district based on our results.

Finally, and perhaps most importantly, the results presented here do little to explain what it is about small schools that affects student outcomes. Potential explanations for the positive effects of school size range from participation in extracurricular activities and attachment to the community to parental involvement and self-esteem (Cotton 1996). Narrowing the analysis from general considerations of size to identify the specific mechanisms by which size matters is essential for effective policy design.

ENDNOTES

[1] This paper, on which Christopher Berry based his remarks at the conference, is a nontechnical version of the following research paper: "Growing Pains: The School Consolidation Movement and Student Outcomes," by Christopher Berry, Harris School of Public Policy, University of Chicago, and Martin West, Harvard University, available at *www.ksg.harvard.edu/pepg/PDF/Papers/PEPG05-04%20Berry%20West.pdf*.

[2] The *Biennial Survey of Education*, which began publication in 1869, was the federal government's first attempt to track statistics related to state and local education. In 1960, it changed its title to the *Digest of Education Statistics*.

[3] This same period—the 1920s to the 1970s— was also notable for a pronounced shift away from one-teacher schools. In 1927, the first year for which data on one-teacher schools are available, they composed 60 percent of all public schools. By 1970, the one-teacher school was all but extinct; only about 400 remained as of 1999.

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

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

[6] This literature is reviewed by Fox (1981) and Stemnock (1974).

[7] The exception, Kenny (1982), found increasing returns to scale. Four of the studies also identified constant returns to scale over at least some range of the data, suggesting that returns to scale in school size may be nonlinear.

[8] Although our model is inspired by Card and Krueger (1992), the identification strategy also differs in important respects.

REFERENCES

Andrews, Mathew, William Duncombe, and John Yinger. 2002. "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?" *Economics of Education Review* 21: 245–62.

Card, David, and Alan Krueger. 1992. "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100: 1–40.

Conant, James B. 1959. *The American High School Today: A First Report to Interested Citizens*. New York, NY: McGraw-Hill Book Company.

Conant, James B. 1967. *The Comprehensive High School*. New York, NY: McGraw-Hill Book Company.

Cotton, Kathleen. 1996. "School Size, School Climate, and Student Performance," School Improvement Research Series, U.S. Department of Education, Office of Educational Research and Improvement.

Fox, W.F. 1981. "Reviewing Economies of Size in Education," *Journal of Education Finance* 6.

Heckman, James, Anne Layne-Farrar, and Petra Todd. 1996. "Human Capital Pricing Equations with an Application to Estimating the Effect of School Quality on Earnings," *Review of Economics and Statistics* 78: 562–610.

Hooker, Clifford, and Van Mueller. 1970. "*The Relationship of School District Reorganization to State Aid Distribution Systems*," National Education Finance Project, Special Study No. 11.

Howley, Craig. 1996. "Review of the Relevant Research." Doctoral Dissertation, Chapter 2, accessed at *http://oak.cats.ohiou.edu/~howleyc/chapter2.htm*

Kenny, Lawrence. 1982. "Economies of Scale in Schooling," *Economics of Education Review* 2: 1–24.

Lee, V. E., and J. B. Smith 1997. "High School Size: Which Works Best and for Whom?" *Educational Evaluation and Policy Analysis* 19: 205–27.

Reynolds, David. 1999. *There Goes the Neighborhood: Rural School Consolidation at the Grass Roots in Early Twentieth-Century Iowa*. Iowa City: University of Iowa Press.

Strang, David. 1987. "The Administrative Transformation of American Education: School District Consolidation, 1938–1980," *Administrative Science Quarterly* 32: 352–66.

Summers, Anita A., and Barbara L. Wolfe. 1977. "Do Schools Make a Difference?" *American Economic Review* 67: 639–52.

Tyack, David. 1974. *The One Best System: A History of American Urban Education*. Cambridge, MA: Harvard University Press.

Workforce Quality and Public-School Reform

Michael H. Moskow[1]

I'd like to thank the Federal Reserve Bank of Cleveland for hosting this Conference on Innovation in Education. The benefits of education are easy to see. Education is essential for citizens to participate in a responsive democracy, and it has meant growth and progress for Americans. Historically, gains in educational achievement have gone hand-in-hand with the adoption of new technologies and improvements in our standard of living. And in our increasingly complex society, education is essential to making wise saving, investment, and occupational decisions that determine our lives' financial prospects and economic well-being.

I'm going to begin by discussing how education, workforce quality, and labor productivity are connected. I will then follow by addressing the need for school reform—especially choice and competition—in our large central cities. Here, I will draw on what Chicago has learned from its sometimes arduous—but hopefully progressive—path to improving its public schools.

EDUCATION AND WORKFORCE QUALITY

Why should a Federal Reserve Bank president be interested in education? One reason is that, as monetary policymakers, we are constantly tracking productivity growth because it is a key determinant of our standard of living. And an important factor driving productivity growth is worker quality, which includes the education and the experience of the workforce.

Estimates by Dan Aaronson and Dan Sullivan at the Chicago Fed find that, of the 2.7 percent average annual growth in labor productivity from 1965 to 2000, almost a quarter of a percentage point (0.22) is attributable to labor quality. Changes in age, education, and labor force participation cause this contribution to vary over time. In the late 1980s and early 1990s, improvements in worker skills were adding 0.40 percentage points per year to the growth of output. By the end of the 1990s, this had fallen to 0.18 percentage point. And we could see a decline to 0.05 percentage point by 2010, as the highly experienced workers of the baby boom generation retire in increasing numbers. Gains in education and other workforce skills by new entrants, and skill improvements by remaining workers, could offset much of this decline. But if we consider recent education trends, I'd say we have our work cut out for us on this.

The quarter century after World War II was a period of especially rapid gains in worker skills. High school graduation rates increased throughout this era, and college graduation rates tripled. The expansion of secondary and postsecondary education caused labor-quality growth to average nearly 1 percent per year, as younger, more highly educated workers replaced retirees with less educational attainment.

THE NEED FOR SYSTEMIC SCHOOL REFORM

Since that time, this process has played out as the educational attainment of retirees and new entrants has converged. One disturbing element of this convergence is that high school completion rates have stalled. In addition, relative to the population of 17-year-olds, the number of traditional high school diplomas granted in the late 1990s was 7 percentage points lower than in the early 1970s. It is only when GED holders are included that current overall high school graduation rates match the earlier ones.

College graduation rates are growing as families have noted the very high and climbing economic returns to education. The gaps in wages and unemployment rates between skilled and less-skilled workers in the U.S. economy have widened dramatically since the late 1970s. Currently, unemployment is only about 2.5 percent among those with a college degree but

over 8 percent for high school dropouts. Furthermore, research shows that each additional year of education tends to raise income by about 10 percent. And these returns are not merely private. Researchers are finding that education raises the productivity of other workers, lowers crime, and raises public involvement in the policy process. Some believe that these social benefits alone may exceed the 10 percent per year private return.

Our investment in education is already enormous. We currently spend almost one-half trillion dollars on elementary and secondary schools in the public sector, or about 4 percent of GDP. The United States has nearly the highest spending per pupil in the world. Yet there is a great deal of dissatisfaction with the results, particularly the dismal outcomes generated by many of our urban schools. We must spend the money more wisely and achieve a substantially greater return on our investment. This has driven the spate of interest in experimentation and systemic reform to improve our schools, especially for those disadvantaged students who have limited choices and opportunities in our current system.

Many school reforms have been tried. Reforms in states in my district, similar to those in Ohio, have been both earnest and varied—from Milwaukee's full-fledged voucher program, to the extensive charter program in Michigan, to Chicago's central authority with the mayor having substantial executive power. As one who believes that customers know what they want, I think that choice and competition should be given a full trial in these reform efforts. This is because market mechanisms relentlessly work in most instances to deliver the right services to their customers.

Some believe that widespread customer dissatisfaction with urban schools is not sufficient to justify greater choice and broader opportunity in the delivery of school services. They want to wait for definitive statistical evidence before proceeding with widespread reforms. As researchers, I'm sure you appreciate that statistical studies—even those based on randomized experiments—have not provided entirely informative or conclusive evidence on outcomes from reform. Still, it is important to keep pursuing evidence on all fronts, as you are doing here at this conference, from case studies to cross-sectional comparisons to carefully designed randomized trials, so that good science can drive good public policy. We must continue to acquire data—but we can't wait for definitive results before moving forward with programs aimed at significant improvement.

IMPROVING TEACHING EFFECTIVENESS

When discussing reform in education, a good place to start is where the "rubber meets the road," that is, where our teachers interact and deliver services to their students.

Now, teaching is a very difficult job. I know, having been a high school English and history teacher myself many years ago. Our schools are extremely fortunate to have many excellent and dedicated teachers. But we clearly need more. Teacher quality matters greatly for student achievement. Several carefully designed research studies from across the country—several of which have been or will be presented at this conference—show that teacher quality varies significantly, even within the same school. To give you some idea of the magnitude, Dan Aaronson and Lisa Barrow of our staff found that the test-score gains for an average student in the Chicago Public Schools would increase by at least 20 percent if that student were reassigned from a classroom with an average-quality math teacher to a classroom in the same school with a math teacher ranked at the 95th percentile of his or her quality distribution.

But what can we do to improve our teacher corps? Can we simply sort through our teachers, identify the ineffective ones, and then replace them? Well, that's not so easy to do in our current system.

Surprisingly, the typical credentials that determine compensation in our schools today—advanced degrees in education, certification, years of teaching experience—do not help much in identifying who the effective teachers are, as Aaronson and Barrow also found. Their research shows that these factors account for only about 3 percent of the variation in teacher quality.

Can we substantially improve the effectiveness of our teachers through enhanced training? While we know that teachers matter, we know less about programs that could improve the quality of our teachers. Research shows that having a math or science degree is helpful for teaching those subjects. Specific skills do matter. But since the general education credentials seem to matter little, traditional teacher training programs—including those college curriculums specializing in

education—apparently don't do the job. There are, however, heroic efforts and exemplary program models underway to improve teacher preparation throughout the nation. These include better recruiting strategies, mentoring, and enhanced training academies. In many cases, these programs have been generously funded by the business and philanthropic communities. Still, they do not reach the majority of our teaching corps.

Blunt instruments, such as raising teachers' salaries across the board, also do not look promising. Raising salaries would attract and keep better teachers, but it would also encourage many poorly performing teachers to remain on the job even longer—in other words, such a policy would be a slow and expensive way to raise teacher quality. And, even with higher salaries for new teachers, excellent teachers will find themselves hostage to rigid pay scales that do not compensate them for their excellence. Nor will they be able to advance their careers without leaving the classroom.

However, one important finding from studies of teacher quality and student outcomes is that high-quality teacher performance tends to persist from year to year. This means that we can predict good teachers from their past performances. So through observation and timely assessment of data on student achievement, a motivated and empowered school principal could accurately identify high-quality teachers.

Therefore, we know some of the ways in which we can build a higher-quality teacher workforce. It will require a system that starts with more selective hiring, includes a lengthy apprenticeship with comprehensive evaluation, and follows up with regular, rigorous personnel evaluations with pay-for-performance rewards. In addition, we should have higher pay for teachers in subject areas where it is more difficult to find qualified instructors, such as math and science.

So far, so good. But what kind of systemic changes can actually bring about such personnel policies—indeed, ones that are taken for granted in many other professions and businesses? Unfortunately, public education has many elements of a monopoly—a publicly owned and operated monopoly. And consumers are served far better in just about any market when there is competition by many as opposed to having just one provider.

So I believe that systemic change in public education would be encouraged and accelerated by the infusion of market incentives. Choice and competition have the power to lead us to better educational outcomes by spurring both new and existing schools to innovate, keep costs low, and better serve the students.

In competitive environments, successful firms and organizations discover their customers' needs. They make strong and innovative efforts to tailor their services to meet those needs at a reasonable cost. When doing so, they adopt pay-for-performance and other types of personnel and compensation practices that we should be aspiring to in the education profession.

CHICAGO'S REFORM EFFORTS

Let me now turn to Chicago as an example of a school system that has taken a series of steps that are beginning to bring about a modest amount of choice and competition. It has, however, been a winding path, preceded by several other attempted reforms.

In the late 1980s, then–Secretary of Education William Bennett called the Chicago public school system the worst in the nation. Correct or not, Chicago's school system was characteristic of older, large-city school systems across the Northeast and Midwest. Flight from the city in both its economic base and middle-class population during the 1960s and 1970s left behind a school system that was poorly suited for the daunting job of ameliorating poverty.

A series of reforms began in 1988, when the state legislature passed the Chicago School Reform Law, which diminished the authority of the Chicago Board of Education and pushed some local decisionmaking down to the individual school and community level. Reformers hoped that the schools would function more efficiently if they could simulate the decentraized decisionmaking autonomy of small suburban school districts. One major provision of the law was the creation of Local School Councils (LSCs) for each school. Accountability for school performance and pupil performance was vested with LSCs, whose members were elected by the residents of each school attendance area. The LSCs had the power to appoint principals, largely free of interference by the Board of Education; they also had some budget authority. In addition, principals ostensibly gained the power to appoint and fire teachers, although in practice, these

powers were not as effective as envisioned by some supporters of the legislation.

Initial hopes were high, and some neighborhoods effectively managed and improved their local schools. Some continue to do so today. At too many school sites, however, LSCs could not effectively grapple with the entrenched power of the system, the political influence of the teachers' union, and the vagaries of community politics. Some LSCs were never able to organize effectively to make the necessary improvements.

By 1995, Chicago school reform shifted toward a different approach: strong central control vested in the office of the mayor. The Illinois state legislature passed a new education reform bill that generally redistributed power from LSCs to the city government. Specifically, the Chicago Board of Education was abolished, and the mayor was given the power to appoint a new five-person, corporate-style School Reform Board of Trustees with a CEO.

The Reform Board addressed waste, fraud, and inefficiency. It privatized janitorial and maintenance services, tightened the purchasing process, and cleaned up the system's finances.

This strong central authority was also given the ability to close schools for poor performance. For the first time, in 1997, seven Chicago high schools were "reconstituted"—the principals were removed and most of the teachers were fired.

These reforms have been modestly effective in raising student performance. As measured by the state's school achievement test (ISAT), both reading and math scores have risen at the third-, fifth-, and eighth-grade levels since 1999. These gains have been realized in high-poverty schools as well as selective-enrollment schools.

However, holding on to these gains in the high school years has proven to be more difficult. In fact, we have not seen gains in test scores at the high school level. And even for the lower grades, achievement gains have tapered off over the last couple of years. It is perhaps for these reasons that the mayor, the public schools, the business leadership, and community groups have come to the realization that further reforms are necessary. New and superior schools, employing innovation and creativity, must be created to replace nonfunctional schools in high-poverty neighborhoods.

THE RENAISSANCE 2010 PROGRAM

Renaissance 2010 is a recent product of this effort. This bold program aims to create 100 excellent new schools—charter schools, contract schools, and performance schools—with more independent and entrepreneurial leadership. These schools have more freedom to innovate because they are less encumbered by the historical layers of rules and procedures. In some instances, schools will exercise greater latitude to recruit and choose their own staff and to evaluate and reward their performance. Such an environment surely will attract teachers with desire, drive, talent, and commitment.

As they are elsewhere, charter schools are an important part of reform in Illinois. The majority of the new Renaissance 2010 schools this year and those planned for next year are charters.

Illinois passed charter school legislation in 1996, but the tough political environment caused this legislation to be weak. It limited the total number of charters allowed to just 15 for Chicago (later expanded to 30), compared with over 600 Chicago public schools total. And it only permitted the public-school district and the State Board of Education to charter new schools.

Still, this early and tentative attempt at choice and competition has built a basis for expanding reform. The Chicago charter school support network and evaluation process were carefully crafted and nurtured. Importantly, the business community has supported charter schools with funds and programs, despite often being at odds with parts of the CPS bureaucracy. Furthermore, charters are fully accountable for results, and some have been closed down for poor performance. And we see that students at charter schools have made promising gains. A recent CPS evaluation concluded that "charter schools performed as well or better than comparable neighborhood schools on 79 percent of the performance indicators." This performance has generated a broader-based trust in charters by the public. So the process, relationships, and trust created by the early charter programs are now in place to expand these existing efforts within the Renaissance 2010 framework.

At the beginning of this academic year, 18 Renaissance 2010 schools were opened to replace schools that were closed due to declining enrollment and deteriorating facilities. Sixteen additional Renaissance 2010 schools are in the final stage of approval to start operating next year. The business community and family foundations provided start-up funds for these new schools. The objective, of course, is not only to improve the education of the students enrolled in the new schools but also to provide incentives for traditional public schools to improve.

Renaissance 2010 schools will stay in operation only as long as they "deliver the goods" to their deserving customers. They will be held accountable in two ways. First, Renaissance schools can expect close scrutiny and rigid assessment by the Chicago Public Schools administration—not of their pedagogy and procedures, but of how well their students perform. Second, parents will be able to review their options and, if they choose, move their children to another school—just as many suburban parents now do when selecting the community in which to live.

Looking forward, the problems inherent in raising educational attainment for our most disadvantaged students are not easy. Too often, the hurdles to student achievement involve family background, resources, neighborhood, and environment. To overcome these obstacles, the local school cannot be "average" in performance or design. Rather, it must be highly innovative and focused on the particular circumstances that hold back disadvantaged children. Such innovation tends to arise in other environments characterized by choice and competition. Competition makes service providers attentive to the particular needs of their customers. In the case of education, this can help generate a sense of understanding and mutual responsibility between the schools and the families they serve.

In bringing about needed changes in Chicago, we have found that partnerships and persistence can make a difference. The disadvantages that many of our students face were not created overnight—nor were the institutions and governance structures that we are attempting to refashion. And determined resistance to reforms, such as those inherent in Renaissance 2010, continues every step of the way. But change is coming, even if it is not coming about through sudden upheaval. Gains are not coming in profound leaps, but real gains are being achieved.

As we look back over 15 to 20 years of reform in Chicago, we can see that each stage of reform contributed to the progress we have made in addressing our educational challenges. We have made changes in the governance of our schools, and we have introduced some choice and competition. In some cases, this progress was achieved by learning from our mistakes. We must continue to learn which reforms work and which do not. And as we strive to make further progress, we must persist in our efforts to bring about the full range of educational opportunities for the younger generation.

ENDNOTE

[1] The views presented here are my own and not necessarily those of the Federal Open Market Committee or the Federal Reserve System.

Choice, Charters, and Public-School Competition

Eric A. Hanushek

When historians review the changes made concerning schools over the course of the twentieth century, two things are likely to stand out: the dramatic consolidation of school districts, leaving fewer, and significantly larger, districts, and the rise in unionization of schools. Historians may not immediately see the interaction of these two things, but it is precisely this interaction that increases the importance of new forms of competition among schools. The improvement of our schools in the twenty-first century is likely to rest on developing forms of school choice—vouchers, charters, and other institutions—that counteract the forces of the twentieth century.

School choice comes in a variety of forms ranging from home-location decisions to home schooling. This paper considers the underlying concepts behind choice and then concentrates on the alternative forms of public-school choice that have developed—contrasting open-enrollment programs with charter schools.

All consideration of school choice is, of course, complicated by the politics of the situation.[1] This discussion focuses on the outcomes of choice and not the underlying politics of implementation.

THE CONCEPT OF SCHOOL CHOICE

The expansion of schooling during the twentieth century dramatically changed the nature of discussions about schooling in the United States. The United States, which led the world's schooling transformation, went from a small, elite system to one that was significantly changed in breadth and depth. Universal schooling with progressively older students became the norm throughout the country.

There was also a dramatic consolidation of school districts. In 1937, there were 119,000 separate public-school districts. Today, there are fewer than 15,000. Over the same period, funding also changed dramatically. In 1930, less than ½ percent of revenues for elementary and secondary schools came from the federal government and less than 20 percent came from states, leaving 80 percent to be raised locally. By 2000, the local share was down to 43 percent, with both federal and state shares rising.

Why is this important? It is reasonable to presume that parents of school children were much closer to what was going on in the schools 75 years ago than they are today. Small districts that were supported by local funds almost certainly must pay attention to the needs and desires of their students. But just the opposite is likely today. A limited number of large districts effectively moves the decisionmaking and management of school districts away from the local population. Moreover, larger districts mean parents have more diverse preferences concerning what they want in their schools. Thus, the choice of any district is necessarily a compromise among various interests.

Another aspect of the changes in government revenue and support has been the overall centralization of decisionmaking. As states have become more prominent in funding schools, they have also moved toward more centralization of decisions about operating them. This is natural because, if they are going to fund schools, they do not want their state (or federal) funds to be wasted. But again, the result is that school decisions have migrated away from the parents and local voters and toward state bureaucracies.

The small school districts found at the beginning of the last century show one way in which schools can be responsive to their constituencies. If the schools deal with a limited number of parents and if the parents directly control the funding of the schools, parents can exert some influence on what the school does.

The responsiveness of districts would not require direct consultation with all of the parents. Tiebout (1956) suggested that parents could satisfy their desires for local governmental services by shopping for the jurisdiction that provided the best level of services for their individual desires. Thus, by sorting out across places, parents could group together to ensure more homogeneous demands. Moreover, since one aspect of schools is how effectively they use their resources, competition for consumers could pressure schools to improve their performance and efficiency.

This view of shopping across alternative jurisdictions does, however, have limitations. For a variety of reasons, the public schools might not look too different from each other. The central state restrictions, the limited viewpoints of school personnel, and other factors could lead schools to be quite similar in approach, curriculum, and goals.

The contraction of choices of different school districts, along with the other choice aspects of home location, thus put natural limits on the Tiebout choice that can go on in many areas of the country. Restoring the ability of parents to enter easily into the schooling process will depend crucially on developing and sustaining new ways for them to exercise choices.

Expanded choice in schools was first promoted in Milton Friedman's *Capitalism and Freedom* (1962). He argued that government may want to intervene in the area of education for a variety of reasons, but none of the potential reasons, including ensuring a minimal level of education by the population or enabling the children of the poor to attend schools, requires government actually to run the schools. The now-obvious alternative identified by Friedman is providing vouchers to parents. These vouchers would transfer funding to the school that a parent chooses, allowing an alternative to the Tiebout choice of schools.

The fundamental idea, underlying either form of choice, is that freeing up consumer demand can have a variety of beneficial effects. Consumers can select the alternative that best meets their interests and desires. Importantly, since few consumers like overpriced goods, such demand pressure could lead to efficiency and innovation in education. If one school did not provide good value, it would tend to lose students to a competitor that offered more for the level of spending. And it is precisely these incentives that are most important in assessing school choice.

With some exceptions due to special circumstances such as the Cleveland situation, the voucher idea has yet to be met with much policy success. Perhaps the most obvious factor is the rise of teachers' unions. At the time of the original suggestions of vouchers and the related significant changes in schools, unions were not pervasive. Their subsequent rise and increase in power has forever changed the ability to introduce any radical policy in schools. Specifically, a fundamental precept and implication of competition in schools is that the job security of some current personnel would be threatened. This result is anathema to unions, which have vigorously attacked any hint of even experimenting with choice. They have been very effective at resisting any such change, mounting powerful media campaigns to prevent citizen referenda on vouchers from being adopted.

A particularly effective argument in the public-relations war over vouchers states that giving money to private schools would harm public schools and that we should instead be working to improve public schools. A second argument states that private schools are not under the control of the government and are not accountable for the government funds they receive. The following sound-bite summary has been the mantra of a number of people: "I favor choice, but it should be restricted to public-school choice." This position has been particularly popular among politicians who want to protect the existing public schools from any competitive pressures while still seeming open to more fundamental school reforms.

Yet citizen sentiment for expanded choice has generally increased over time, and this has led to a variety of innovations in school choice that fit the notion of public-school choice. Importantly, they are not all the same, and they have very different incentive effects. Two quite different kinds of choice stand out: open-enrollment or magnet school plans and charter schools. It is useful to review these in terms of outcomes and incentives.[2]

PUBLIC-SCHOOL CHOICE

A particularly popular version of public-school choice involves an open-enrollment plan. For example, any student could apply to a school in his or her district

other than the one to which he or she was originally assigned. Or, in a more expansive version, no initial assignment is made, and students apply to an ordered set of district schools. A common version of this has been the use of magnet schools that offer a specialized focus such as college preparatory or the arts.

Forms of open-enrollment plans were the response of a number of Southern districts to the desegregation orders flowing from *Brown v. Board of Education*. In general, simple open-enrollment plans did not satisfy the court requirements for desegregation, but magnet schools (with racial-balance restrictions) became a reasonably common policy approach (Armor 1995). In 2001–02, 3 percent of all students attended a magnet school (Hoffman 2003).

As a general rule, open-enrollment plans produce few of the incentives that lie behind voucher plans. The flow of students is heavily controlled by the common restrictions that space must be available and that other requirements, such as racial balance, must be met. Most importantly, however, these plans seldom have much effect on incentives in the schools. Under open enrollment, personnel in undersubscribed schools generally still have employment rights and would simply move to another school with more students. Extending open enrollment across districts conceptually provides stronger incentives but unattractive funding, and the "if there is space at the school" clause generally stops all but some token movement.

A different development—charter schools—appears to offer stronger choice incentives. These schools differ dramatically by state, but their essential feature is that they are public schools allowed to operate to varying degrees outside of the standard public schools. They are schools of choice, surviving through their ability to attract sufficient numbers of students.

Charter schools can offer true competition to the regular public schools because they can draw students away from poorly performing regular publics. Employment rights typically do not transfer between charters and regular publics, so personnel in charter schools could be under pressure to attract students. The pressure on regular public schools comes from the potential loss of students, which would lesson the demand for public schools and their teachers.

Since the nation's first charter-school legislation was enacted in Minnesota in 1991, 41 states and the District of Columbia have passed legislation that provides for charter schools, although some had yet to open any schools by 2004. For the nation as a whole, charter schools increased from a handful in 1991 to nearly 3,200 schools serving almost 800,000 students, or over 1.5 percent of the public-school population, in 2004. In some places, charters have become quite significant. For example, in the 2003–04 school year, almost 17 percent of students in the District of Columbia, 8 percent in Arizona, and 4 percent in Michigan attended charter schools.[3]

To date, studies of the outcomes of charters have been limited by some serious analytical difficulties. Because the students voluntarily choose these schools, it is always difficult to infer the impact of the school as distinct from the characteristics of the students it attracts. Additionally, because charter schools are largely new, most are still going through a start-up phase, and it takes large inferences to know what they will look like in the steady state.

My own work provides some preliminary estimates of the performance of charters in Texas (Hanushek, Kain, Rivkin, and Branch 2005). The simplest design for dealing with selection problems is to compare the average learning growth for individual students when in the regular public schools with their own performance in the charters. In this way, charter students become their own control group.

Three things come out of this in terms of quality indicators. First, on average, charter schools perform very similarly to the standard public schools. But second, start-up problems are real, and new charters do not perform as well as more established charters, (those over two years in age), which, on average, outperform the standard public schools of Texas. Third, there is a significant distribution of performance across both regular public and private schools. The good are good, and the bad are truly bad.

These findings are consistent with much of the other recent work, although there are some remaining uncertainties. The average North Carolina charter appears less effective than the average traditional public school (Bifulco and Ladd 2004), while the average Florida charter is on par with the regular public schools after a start-up phase (Sass 2005). On

the other hand, relying upon comparisons between charter applicants in Chicago who were randomly accepted or randomly denied admission, Hoxby and Rockoff (2004) conclude that the three charter schools they observed significantly outperformed their standard-school counterparts. But these results await both the general maturation of more charter schools and the investigation of their performance in different settings.

Another important aspect of competitive markets is the enforcement of discipline on the other participants—in this case, the regular public schools. Is there any evidence that the regular public schools respond to the pressures of competition? Again, it is very early in the development of charters, but Hoxby (2003) introduces preliminary evidence that there are competitive improvements.

Our Texas study also provides information on competition. If we look at the behavior of parents, we find that they are significantly more likely to withdraw their children from a poorly performing charter as compared to a well-performing charter (Hanushek, Kain, Rivkin, and Branch 2005). This finding is particularly important because parents are not given information on the value-added of their charter school. The behavior of parents shows, however, that they are good consumers and that they can use the performance data that are available to infer the quality of the school. An early and continual criticism of the voucher idea is that parents are not good consumers, an assertion belied by the data.

CONCLUSIONS

The idea of school choice is a natural extension of arguments about the benefits of competition to education. The clearest form, advocated originally by Milton Friedman and picked up by a wide variety of other people, is to give parents vouchers that allow them to shop for schools. While special circumstances have led to the use of vouchers in Milwaukee, Cleveland, and the District of Columbia, their growth has been slow and erratic.

On the other hand, alternative forms of choice—under the banner of public-school choice—have become more pervasive. Some, but not all, of these alternatives offer benefits that are similar to vouchers. Most notably, charter schools offer students and parents the possibility of options that have the ancillary advantages of introducing competitive incentives for schools.

Charter schools are difficult to evaluate. Because students self-select into these schools, it is difficult to separate the quality of the students from the quality of the charter school. Moreover, most charter schools started very recently, making it difficult to see how they will evolve as they age.

Nonetheless, the best available evidence available indicates that, after a start-up period, charters have as much value-added as regular public schools—if not more. As with regular public schools, however, there is a wide range of quality in charters. But, importantly, parents appear able to recognize the quality of charter schools and to act upon that information by exiting low-quality charters at significantly higher rates than higher-quality charters.

Current personnel in the regular public schools resist expansion of charters, which they consider undesirable competition. This resistance takes a variety of forms. In some states there are strong pressures to limit the number of charter schools. In others, arguments that all schools should have a “level playing field” are used to justify increasingly stringent restrictions on the operations of charters. If we are to obtain the benefits of choice and competition, these pressures should be resisted.

ENDNOTES

[1] For example, the teachers' unions, as part of their resistance to competition, gained national publicity for their simple comparison of scores for students in charter schools versus those in regular public schools (Nelson, Rosenberg, and Van Meter 2004). More serious work, however, has concentrated on adjusting for the special populations that opt for charter schools and other choice schools.

[2] One very different option not covered here is home schooling. A significant number of parents have simply withdrawn their children from the regular public schools and have taken personal responsibility for their education (but with no governmental financial support). Some estimates put the numbers of home schoolers between 1.5 percent and 2 percent of all school children, although there is uncertainty even about the numbers involved (Henke, Kaufman, Broughman, and Chandler 2000). Little is known about this in terms of movements in and out or of performance, and the incentive effects for most existing public schools appear small.

[3] Data from the Common Core of Data of the National Center for Education Statistics (*http://nces.ed.gov/ccd/bat/index.asp*).

REFERENCES

Armor, David J. 1995. *Forced Justice: School Desegregation and the Law*. New York: Oxford University Press.

Bifulco, Robert, and Helen F. Ladd. 2004. "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina." Working Paper Series SAN04-01. Duke University: Terry Sanford Institute of Public Policy, August.

Friedman, Milton. 1962. *Capitalism and Freedom*. Chicago: University of Chicago Press.

Hanushek, Eric A., John F. Kain, Steve G. Rivkin, and Gregory F. Branch. 2005. "Charter School Quality and Parental Decision Making with School Choice." Working Paper Series W11252. Cambridge, Mass.: National Bureau of Economic Research, April.

Henke, Robin R., Phillip Kaufman, Stephen P. Broughman, and Kathryn Chandler. 2000. *Issues Related to Estimating the Home-Schooled Population in the United States with National Household Survey Data*. Washington, DC: National Center for Education Statistics, September.

Hoffman, Lee McGraw. 2003. *Overview of Public Elementary and Secondary Schools: School Year 2001–02*. Washington, DC: National Center for Education Statistics, May.

Hoxby, Caroline Minter. 2003. "School Choice and School Productivity (or Could School Choice Be a Tide That Lifts All Boats?)," in Caroline Minter Hoxby, ed., *The Economics of School Choice*. Chicago: University of Chicago Press.

Hoxby, Caroline Minter, and Jonah E. Rockoff. 2004. "The Impact of Charter Schools on Student Achievement." mimeo, November.

Nelson, F. Howard, Bella Rosenberg, and Nancy Van Meter. 2004. "Charter School Achievement on the 2003 National Assessment of Educational Progress." American Federation of Teachers, August.

Sass, Tim R. 2005. "Charter Schools and Student Achievement in Florida," in American Economic Association Annual Meeting Papers. Philadelphia.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures," *Journal of Political Economy* 64 (5): 416–24.

Evaluating HOPE-Style Merit Scholarships [1]

Christopher Cornwell
David B. Mustard

Since the early 1990s, there has been a proliferation of state-sponsored, merit-based college scholarships. Eligibility for the awards typically involves satisfying a grade-point-average requirement in high school, and retention usually depends on meeting a similar GPA standard in college. Another standard feature of these scholarships is that they do not impose means tests. The model for these new state programs has been Georgia's HOPE (Helping Outstanding Pupils Educationally) Scholarship.

State policymakers commonly defend "HOPE-style" merit aid by arguing that it will increase access to colleges and universities, keep the best students from attending college out of state, and encourage academic achievement in high school and college. We evaluate the policymakers' arguments in terms of the empirical support from Georgia for each of these effects. We also examine whether merit scholarships like HOPE increase college stratification by student quality.

THE GEORGIA MODEL

Georgia's HOPE Program was introduced in 1993, financed by a state lottery. The program distributes two types of financial aid—the merit-based scholarship and the HOPE Grant. To qualify for the scholarship, a student must graduate from a Georgia high school with at least a B average in core-curriculum courses. The scholarship covers tuition and fees and provides a $300 book allowance at degree-granting public institutions. Currently, the value of the award is about $4,600 at the state's top universities, accounting for over 40 percent of the total cost of attendance. HOPE scholars at degree-granting private institutions receive a fixed payment of $3,000. Once in college, students must maintain at least a B average with a minimum number of credits to retain the award. In contrast, the HOPE Grant has no merit requirements, but its coverage is limited to tuition and fees associated with nondegree programs offered by two-year and technical schools. Since its inception, the program has paid out more than $3.5 billion in financial aid to over 900,000 students. Forty-five percent of all awards and 60 percent of total aid go to scholarship recipients attending four-year colleges and universities.

As discussed in Cornwell and Mustard (forthcoming), lottery sales far outpaced early projections, leading to a significant expansion of the HOPE program in terms of coverage and generosity. The most important changes were the elimination of the income cap and of the Pell "offset." Initially, the scholarship was restricted to students from households with incomes less than $66,000, but the income cap was raised to $100,000 in 1994 and removed entirely in 1995. In the beginning, HOPE payments were also reduced dollar-for-dollar by any Pell aid the student received. This offset ended in 2001; now students who qualify for both Pell and HOPE can "stack" their awards, providing an even more powerful incentive to attend a Georgia college or university. Most of the states with recently established HOPE-style merit scholarships have generally followed Georgia in leaving out means tests and allowing merit awards to be stacked with Pell aid.

ENROLLMENTS AND THE "BRAIN DRAIN"

Cornwell, Mustard, and Sridhar (forthcoming) examine HOPE's effect on enrollments and the "brain drain" by comparing enrollments in Georgia with those in other southeastern states before and after the program's introduction. Using Integrated Postsecondary Education Data System (IPEDS) data from the National Center for Education Statistics covering the period 1988–97, they show that HOPE raised enrollment in Georgia's colleges and universities and reduced the number of students leaving the state to attend college elsewhere.

Table 1 reports the percentage increases in freshman enrollment attributable to HOPE, broken down by race and institution type, as estimated by Cornwell et al. First, the overall enrollment effect is 5.9 percent, which translates into almost 2,900 extra students per year. Second, the gains are concentrated heavily in four-year schools, with the greater percentage gain in private colleges. Indeed, the magnitude of private-school increase is surprising. One explanation is that the small, moderately selective liberal arts colleges, which comprise a large fraction of the state's private schools, face relatively elastic demand because many similar substitutes operate in proximity to Georgia. Third, the percentage increases of blacks exceed those of whites, with the greater enrollment response for blacks appearing in four-year public colleges. The black enrollment gains are accounted for in part by the students who, instead of leaving the state, chose from the many relatively large, historically black colleges and universities in Georgia, which has the fourth-largest black population and population share in the United States. There is also a program-induced rise in technical school enrollment for blacks (where there is none for whites).

Analyzing the IPEDS student residency and migration data, Cornwell, Mustard, and Sridhar find that HOPE reduced the number of Georgians attending college out of state by about 560 per year. This is a pure scholarship effect because the migration data cover only freshmen in four-year schools who recently graduated from high school. The reduction in "leavers" from the state makes up roughly two-thirds of the total enrollment effect for this group, which accounts for almost 77.5 percent of all first-time freshmen in Georgia's four-year colleges. However, recently graduated freshmen represent only roughly 40 percent of the total four-year-school enrollment rise, implying that the greater enrollment response occurred among freshmen who delayed matriculation for a year after high school graduation.

Finally, the overall enrollment increase reported by Cornwell, Mustard, and Sridhar represents only 15 percent of freshmen scholarship recipients between 1993 and 1997 and an even smaller fraction of all first-year program (scholarship plus grant) beneficiaries. This should not be surprising because programs like HOPE primarily affect where, not whether, a student goes to college.

COLLEGE STRATIFICATION BY STUDENT ABILITY

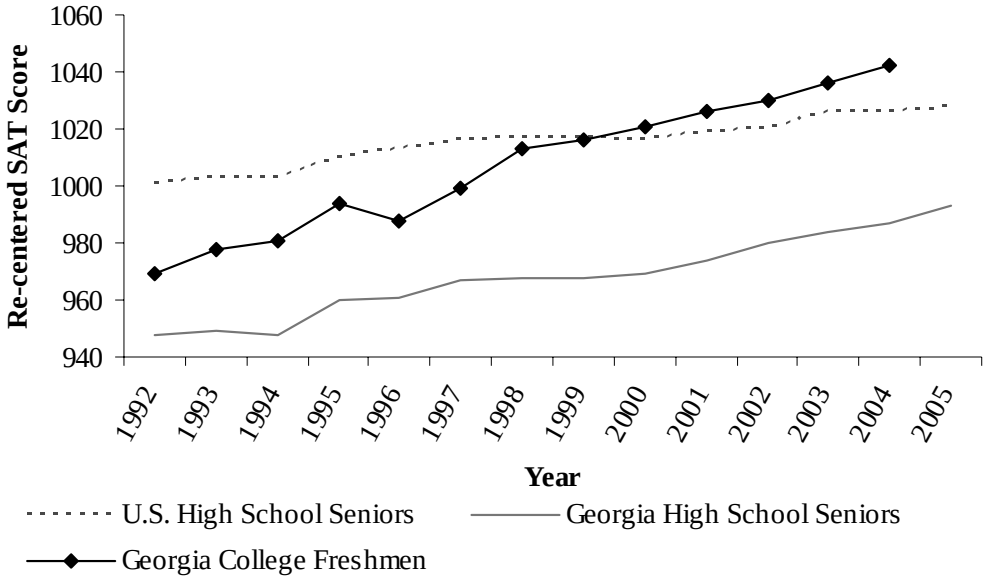
In addition to reducing the number of leavers, HOPE has changed their composition. Figure 1 plots the SAT series for freshmen enrolled in Georgia's public colleges and universities, Georgia high school seniors, and U.S. high school seniors. Since HOPE began, Georgia's freshman SAT scores have increased by a remarkable 60 points. By comparison, the scores of Georgia and U.S. high school seniors rose by only 30 and 20 points, respectively. Further, between 1993 and 2000, Georgia's retention rate of students with SAT scores greater than 1500 climbed from 23 percent to 76 percent.

Table 1: Percentage Increases in Freshmen Enrollments Attributable to HOPE by Race and Institution Type, 1988–97

Group	Overall	4–Year Publics	4–Year Privates	2–Year Publics	2–Year Publics + Techs
All	5.9	9.0	13.0	ns	ns
Whites	3.6	4.4	9.2	ns	ns
Blacks	15.8	26.0	16.8	ns	11.6

Note: 'ns' indicates the estimated effect is not statistically significant.
Source: Cornwell, Mustard, and Sridhar (forthcoming).

Figure 1: SAT Scores of Georgia College Freshmen vs. U.S. High School Seniors and Georgia High School Seniors, 1990–2003



Source: Cornwell and Mustard (forthcoming).

The gains depicted in figure 1 obscure how students are sorted across institution types. Resources available early in life are important in determining a student's prospects for admission to a selective college or university. Thus, to the extent that merit is correlated with household income, programs like the HOPE Scholarship will further stratify higher education by student ability.

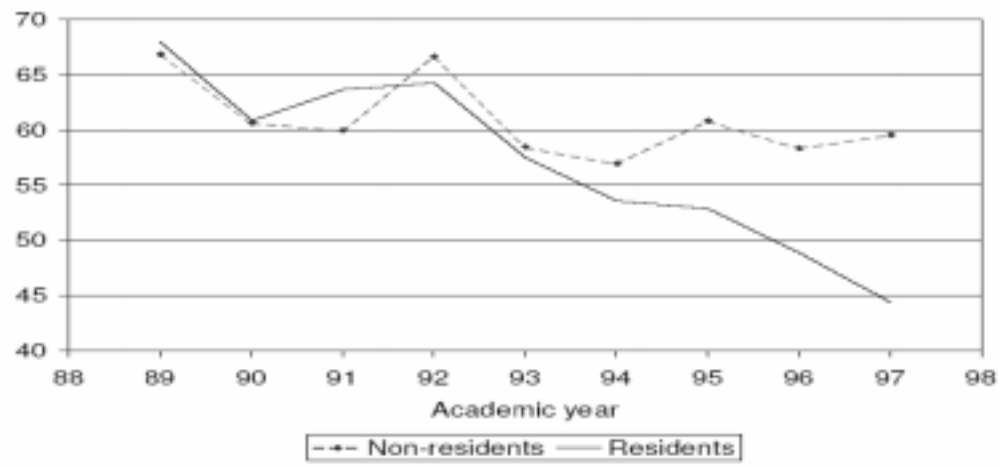
Cornwell and Mustard (2005) address the stratification question using data covering the period 1989–2001, obtained from *Peterson's Guide to College*, to compare student quality in Georgia colleges with that of their southeastern U.S. counterparts. First, we find that in the state's most selective universities, SAT verbal and math scores jumped by 14.3 and 9.4 points because of HOPE. The scholarship also increased these schools' share of students from the top 10 percent of their high school class by 7.6 percentage points. In contrast, the least-selective schools experienced no statistically significant effect from HOPE on any measure of student quality. Second, we show that HOPE reduced the variance of SAT math and verbal scores in the most-selective institutions, but had no impact on the variances at any other institution type. Taken together, these results strongly suggest that HOPE has exacerbated the stratification of enrollment by student quality.

ELIGIBILITY REQUIREMENTS AND ACADEMIC ACHIEVEMENT

The requirements for HOPE eligibility and retention effectively put a premium on maintaining a 3.0 GPA in high school and college. Does this promote academic achievement or encourage other choices that can hinder learning? To the extent the GPA standards for eligibility and retention increase effort and time spent on schoolwork, they enhance learning. If, on the other hand, they cause students to enroll in fewer classes, withdraw from classes more frequently, select courses with higher expected grades, or choose certain majors, their salutary effect on learning may be seriously weakened.

Using data from the longitudinal records of all undergraduates who enrolled at the University of Georgia between 1989 and 1997, Cornwell, Lee, and Mustard (2005a) estimate the effects of HOPE on course-load adjustments by comparing the decisions of in-state and out-of-state students before and after HOPE was implemented. Nonresidents cannot receive the scholarship and therefore constitute a control group. This approach is illustrated in figure 2, which shows the trends in the percentage of resident and nonresident freshmen completing full course loads. From 1992 (the year before HOPE started) to 1997, the resident full-load completion rate dropped almost 20 percentage

Figure 2: Percentage of Freshmen Completing a Full Load Residents vs. Non-Residents, 1989–97



points, while the nonresident rate remained fairly stable around 60 percent.

The broader findings of the Cornwell, Lee, and Mustard study can be summarized as follows: First, HOPE decreased full-load enrollment and increased course withdrawals among resident freshmen. The combined result of these responses is a 9.3 percent lower probability of full-load completion and an almost one-credit reduction in annual course credits completed. Consequently, between 1993 and 1997, resident freshmen completed more than 3,100 fewer courses because of the HOPE Scholarship. Second, the scholarship has the greatest influence on the course-taking behavior of students who are just meeting or falling below the GPA requirement. Third, the scholarship's impact has grown with the lifting of the income cap; by 1995, virtually all resident freshmen entered as HOPE Scholars, while only 35 percent did in 1993. Fourth, HOPE caused Georgia residents to divert course-taking from the regular academic year to the summer, when grades are generally higher, even though the typical summer-school enrollee has a lower SAT score and high school GPA. After HOPE was introduced, summer-course credits increased by an average of 63 and 44 percent in the first two summers following matriculation.

Cornwell, Lee, and Mustard (2005b) go beyond course-load adjustment to examine HOPE's effect on course and major selection. Using the same University of Georgia student-record data, they show that resident freshmen and sophomores completed roughly 1.2 fewer math and science core-curriculum credits because of the scholarship. In addition, they present evidence suggesting HOPE increased the likelihood of a typical resident freshman choosing an education major by 1.2 percentage points, with an even greater impact on women and whites. The scholarship's influence on declared majors is potentially costly because earnings are so closely tied to that choice.

The average GPA of University of Georgia resident freshmen rose 5 percent relative to their out-of-state counterparts during the HOPE period. The results of the Cornwell, Lee, and Mustard study suggest that more than just greater effort or time spent studying may be at work. Rather, HOPE's grade-based retention requirements lead to behavioral responses that partially undermine the scholarship's objective to promote academic achievement.

CONCLUDING REMARKS

Our findings concerning the effects of HOPE-style merit aid are obviously confined to the Georgia experience. The degree to which they generalize to the other states that have adopted similar programs depends on how closely they have followed the HOPE model. Those that have will likely see enrollment effects that largely involve college choice rather than access (with its implication for stratification), because such merit awards target students who will probably attend college anyway. The pattern of enrollment gains will be a function of the number and quality of its schools, notably its four-year institutions. This is particularly important for reducing the "brain drain" because students do not typically leave

the state to attend two-year colleges. As far as academic achievement is concerned, relying heavily on grade-based eligibility and retention criteria will lead to student responses that undermine that objective. Many of the scholarships started in the mid-1990s have this characteristic, although the most recent limit the number of semesters or academic years they can be used, reducing the incentive to lower per-semester course loads.

Finally, we speculate that the proliferation of HOPE-like scholarships, especially in the southeast, may take on the characteristics of an "arms race." In the limit, each state competes to retain its best students, with the students allocated to schools that would, if not for the scholarship, be less attractive to them.

ENDNOTE

[1] Christopher Cornwell based his remarks at the conference on this paper.

REFERENCES

Cornwell, Christopher M., Kyung Hee Lee, and David B. Mustard. 2005a. "The Effects of Merit-Based Financial Aid on Course Enrollment, Withdrawal, and Completion in College," *Journal of Human Resources* 40, 895–917.

Cornwell, Christopher M., Kyung Hee Lee, and David B. Mustard. 2005b. "The Effects of State-Sponsored Merit Scholarships on Course Selection and Major Choice in College," University of Georgia Department of Economics Working Paper.

Cornwell, Christopher M. and David B. Mustard. 2005. "The Effects of HOPE-style Scholarships on College Stratification by Ability, Race, and Gender," University of Georgia Department of Economics Working Paper.

Cornwell, Christopher M. and David B. Mustard. Forthcoming. "Assessing Public Higher Education in Georgia at the Start of the 21st Century," in Ronald Ehrenberg, ed., *What's Happening to Public Higher Education?* Greenwood Press: American Council for Education.

Cornwell, Christopher M., David B. Mustard, and Deepa Sridhar. Forthcoming. "The Enrollment Effects of Merit-Based Financial Aid: Evidence from Georgia's HOPE Scholarship," *Journal of Labor Economics*.

Borrowing Constraints on Families with Young Children [1]

Elizabeth M. Caucutt
Lance Lochner

Raw differences in educational attainment by family income are dramatic. For example, Ellwood and Kane (2000) find that among the high school graduating class of 1992, those with parents in the top income quartile have a 30 percentage point higher probability of attending a post-secondary institution than those from the bottom income quartile. However, most of this gap disappears after controlling for adolescent achievement or cognitive test scores and family background (Cameron and Heckman 1998, 2001; Carneiro and Heckman 2002; Ellwood and Kane 2000). This has led many economists to conclude that short-term borrowing constraints at college ages are not an important determinant of college attendance and completion decisions. That is, increasing credit access for college students is unlikely to affect college enrollment rates by more than a few percentage points and will do little to reduce the attendance gaps by family income or race. By the end of high school, it may be too late to help many youths from disadvantaged backgrounds. It appears that, because they are ill-prepared to do so, poor and minority youths choose to attend college at lower rates than wealthier white youths.[2]

Indeed, there is a wealth of indirect evidence that skill investments are complementary over the life cycle and that a failure to make adequate investments during early childhood years reduces incentives in later years to invest through college attendance or high school completion (Cunha et al., forthcoming). The importance of cognitive-achievement scores in determining high-school-dropout or college-attendance rates is one indicator of the complementarity of early and late investments. Using data on children from the National Longitudinal Survey of Youth (NLSY), figures 1 and 2 show the importance of adolescent achievement scores, as measured by a combined math and reading Peabody Individual Achievement Test (PIAT) score at ages 13–14, in determining high-school-dropout and college-attendance rates. After controlling for family background and family income (measured when children were ages 15–18), a one-standard-deviation increase in PIAT scores at ages 13–14 is associated with a 12 percentage point decline in high-school-dropout rates and a 15 percentage point increase in college attendance.

This study investigates the role of family income and borrowing constraints in determining early investments in children and youth achievement scores. As figure 3 shows, youths raised in families in the bottom third of the income distribution are much less likely to be among the highest PIAT test scorers (at ages 13–14) than are those in middle- and high-income groups.[3] While more than 50 percent of all 13- to 14-year-olds in the top tercile of the income distribution are in the top third of the test-score distribution, fewer than 20 percent of those in the bottom income tercile managed such scores. These findings raise the natural question: To what extent do family borrowing constraints during early childhood and adolescence influence early investments in children, cognitive achievement levels, and ultimately college attendance and completion?

While a number of studies have recently examined the effects of credit constraints on college-going behavior (see Carniero and Heckman 2002 for a summary of the empirical literature), very little attention has been paid to the role of borrowing constraints when children are younger.[4] Yet it seems possible that constraints at early ages play a more important role in determining investment decisions for a number of reasons. First, most empirical studies indicate that early investments in children produce high long-term payoffs (see Karoly et al., 1998 or Blau and Currie, forthcoming, and references therein). Randomized studies of early intervention

Figure 1: Fraction of Youth Who Drop Out of High School by Adolescent Achievement Scores

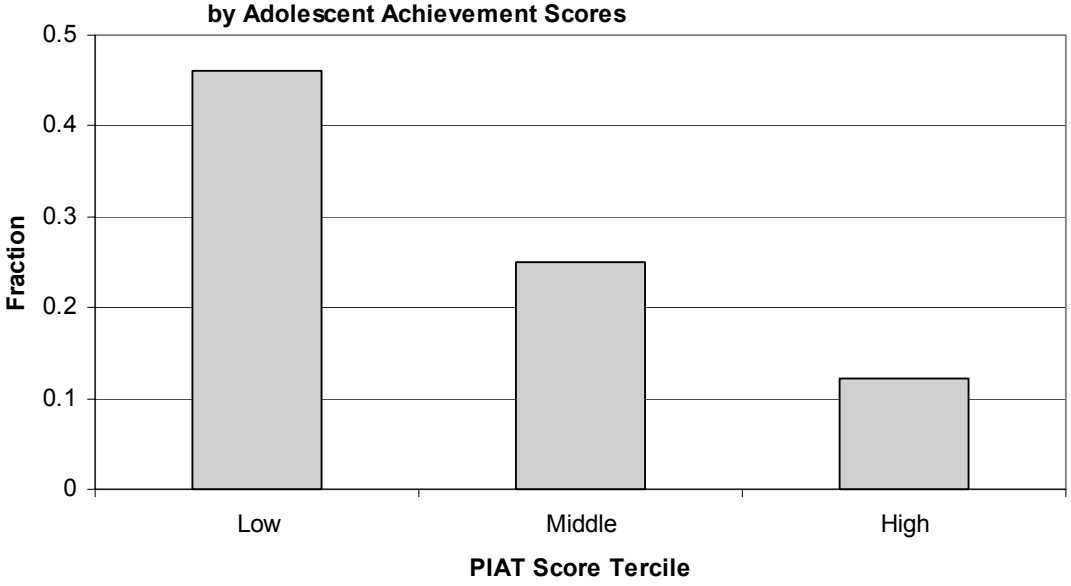
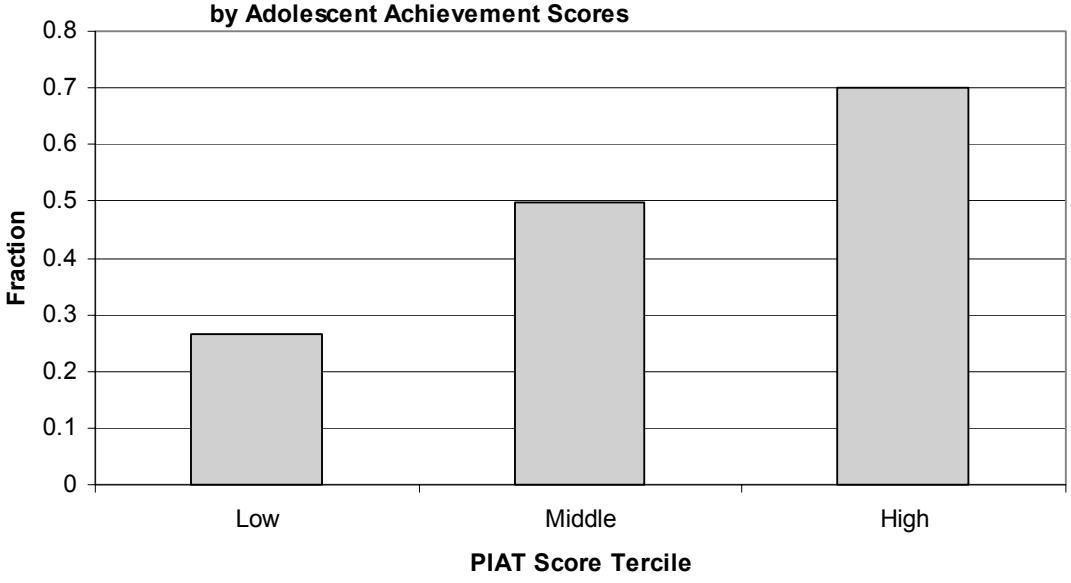
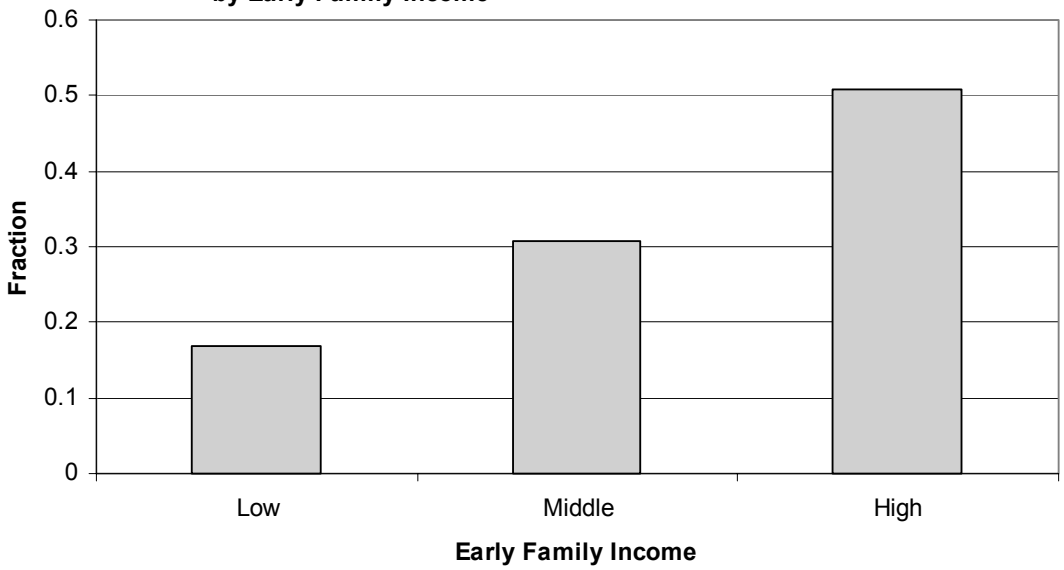


Figure 2: Fraction of Youth Who Attend College by Adolescent Achievement Scores



and preschool programs for disadvantaged children have estimated large benefit–cost ratios, suggesting that many poor families are not making investments in their children even though those investments would more than pay for themselves in the long run. Not only are children able to learn quickly when they are young, but early learning begets later learning as emphasized by Cunha et al. (forthcoming). Second, recent studies have shown that increases in family income lead to increases in the test scores of adolescents and young children (e.g., Blau 1999; Duncan and Brooks-Gunn 1997; Levy and Duncan 1999; and Dahl and Lochner 2005). A few of these studies suggest that family income is more important at earlier ages (e.g., Duncan and Brooks-Gunn 1997; Levy and Duncan 1999). Third, parents age with their children and, as both children and parents age, parental resources tend to increase with the accumulation of human capital and the associated rise in earnings. Fourth, despite generous government

Figure 3: Fraction of Youth in Top PIAT Test Score Tercile by Early Family Income



student-loan programs for college-age students and their families in the United States and other developed countries, governments have not traditionally offered loans to parents of young children to help finance earlier human-capital investments.

While the direct costs of public elementary and secondary education are fully subsidized, a good education through high school is not free. In many U.S. communities, parents must choose between sending their child to poor public schools and paying for their child to attend better private schools. Alternatively, parents may choose between high-cost communities with good public schools and low-cost neighborhoods with poor ones. Other investments in young children can also be costly. Preschool programs in the United States can cost as much as attendance at a top university. While the government does not fully neglect poor preschool-age children, the quality of publicly provided preschool programs (e.g., Head Start) is far below what it could be (Zigler 1994; Blau and Currie, forthcoming). Expenditures on computers and books also add up. Finally, parental time is an important, yet costly, input that poor parents may be unable to afford.

In order to better understand the role of borrowing constraints, we distinguish between *intergenerational* borrowing constraints, which would prevent parents from borrowing against their children's future earnings, and *intragenerational* borrowing constraints, which would prevent individuals from borrowing against their own future income. Intuitively, the former implies that only the present value of lifetime parental income affects child success, while the latter implies that the timing of parental income (over the life cycle) also matters. We use data on children from the NLSY to test for the latter form of constraint, estimating the effects of family income at different ages of the child on the math and reading achievement of those children at ages 5–14. We interpret evidence that the timing of income matters as evidence that *intragenerational* borrowing constraints distort investment decisions.[5]

INTERGENERATIONAL AND INTRAGENERATIONAL BORROWING CONSTRAINTS

Parents may be *intergenerationally* constrained from borrowing against their children's future income; that is, parents may not be able to pass on debts to their children. This is more like a generational budget constraint than a borrowing constraint, but it does imply that family income may affect child achievement. Importantly, this type of constraint, if binding, suggests that the present discounted value of lifetime family income will be an important determinant of child achievement. However, if parents can save and borrow against their own future earnings, the timing of their income should not be important.

The second, stronger constraint—an *intragenerational* one—more closely matches the standard idea of a borrowing constraint. Parents may be unable to borrow against their own future earnings in order to smooth consumption or to make investments in their children at young ages. If parents are borrowing constrained in this way, the timing of their income will be important. That is, constrained parents who earn a smaller share of their lifetime income at younger ages will tend to invest less in their children when they are young than will parents who earn a larger share of their income early on.

The first type of constraint is, in theory, easy to test. Do children from wealthier families perform better than children from poorer families? In practice, however, this type of test is difficult to implement since innate abilities of children may be correlated with the abilities and lifetime earnings of their parents. Additionally, rich parents and poor parents may differ in many unobservable ways that have little to do with income. This problem has plagued past work on credit constraints and college-going behavior as discussed in Carniero and Heckman (2002). So, while we think that intergenerational constraints are almost certainly important, we focus attention on the effects of intragenerational constraints.

Testing for intragenerational borrowing constraints relies on examining how the timing of income matters conditional on the discounted value of lifetime income. This amounts to comparing the test scores of children in families with the same total lifetime earnings but with different income profiles. If families are unaffected by borrowing constraints, children from families who earn a larger share of their income early on should perform as well as children from families who earn more of their income late in their careers. If borrowing constraints are binding, the first set of children should perform better than the second. For example, in the absence of intragenerational constraints, a child raised in a family that earns $20,000 for the first 10 years of a child's life and then $40,000 for the next 10 years should perform as well as a child raised in a family earning $40,000 for the first ten years and $20,000 the next ten (ignoring inflation and discounting). If intragenerational constraints are important, the latter child should perform better. By holding lifetime income constant, we can reduce concerns that a positive correlation between parents' and children's abilities will bias our results.[6]

It is important to recognize that income earned earlier should be worth more in a present-value sense. That is, income earned today is worth $1 + r$ times income earned next year (in the absence of credit constraints), where r is the annual interest rate. It is necessary to account for this when examining the effects of income timing. Therefore, we discount income earned by the family so that it is measured in present-value terms as of the child's birth year.[7] After this adjustment, the absence of intragenerational borrowing constraints suggests that the timing of income should be irrelevant conditional on the discounted present value of lifetime family income.

We examine three types of evidence on the importance of income timing. First, we test whether the slope of a family's income profile significantly affects test scores conditional on the discounted present value of lifetime family income over a long time span. Second, we estimate the effects of past income on children's achievement test scores to see whether income earned at a child's earlier ages has a different effect on test scores than does income earned at later ages. Third, we estimate whether future income has the same effect on test scores as do past and current income. The first test examines the role of income timing over the past, present, and future; the second compares the effects of income earned at different points in the past; and the third compares the effects of past income with future income. As we next show, these tests point to the existence of intragenerational borrowing constraints (as well as intergenerational constraints): family income earned at a child's earlier ages has significantly larger effects on a child's test scores than does income earned at later ages.

EVIDENCE ON BORROWING CONSTRAINTS AND CHILD ACHIEVEMENT

The National Longitudinal Survey of Youth

We use data on children from the NLSY and the main NLSY sample of mothers. These data are ideal for studying the effects of family income on children because they enable us to link children to their mothers and follow families over many years. For children, biannual measures of family background and cognitive and behavioral assessments are

available from 1986 to 2000. Detailed longitudinal demographic, educational, and labor market information for the mothers is available annually from 1979 through 1994 and biannually thereafter. Additionally, a widely used measure of cognitive ability—the Armed Forces Qualifying Test (AFQT)—is available for mothers in the NLSY. The NLSY oversamples minority and poor white families, which provides a larger sample of the families more likely to face borrowing constraints. We use data drawn from more than 7,000 interviewed children born to over 3,500 interviewed mothers.

Since family-income measures are available annually from 1979 to 1994 and biannually thereafter, for children born after 1979, we were able to compile an income history for almost every year since birth. Our empirical analysis first uses the Consumer Price Index for Urban Consumers (CPI-U) to deflate all income measures to year-2000 dollars. We then created two measures of family income that we use for our empirical analysis. The first simply discounts income at each age of the child back to the first year of that child's life, using a 5 percent interest rate. This puts all income measures for a child on an equal basis for comparison—a dollar at age 10 should be just as valuable as a dollar at age one after discounting in this way. The second income measure we created is a measure of average "lifetime income," which is simply the average of all discounted income measures for all observed periods from birth to the final survey date. This measure is fixed for each child and does not vary over time.

We analyze PIAT math and reading scores, collected biannually from 1986 to 2000, for children ages 5–14. The assessments measure mathematics ability, oral-reading ability, and the ability to derive meaning from printed words. To simplify our exposition, we focus on a combined math and reading test score, which places a weight of 50 percent on the math score and 25 percent each on of the reading-comprehension and reading-recognition scores. To make these scores more easily interpretable, we created standardized test scores with a mean of zero and a standard deviation of one.[8]

Empirical Tests of Intragenerational Borrowing Constraints

In the absence of credit constraints, the slope of a family's life-cycle-income profile should have no affect on child achievement after controlling for the discounted present value of lifetime income. The presence of intragenerational borrowing constraints suggests that children raised in families that obtain more of their income earlier will perform better; thus, the slope of income profiles should be negatively related to achievement. We test this proposition by regressing PIAT scores at all observed ages on the discounted present value of lifetime family income and the slope of family-income profiles (as well as a

Table 1: Effects of Lifetime Income and the Slope of Family-Income Profiles on PIAT Scores

	1	2	3
Slope of Family-Income Profiles	**–0.1689**	**–0.1187**	**–0.1629**
	(0.0838)	**(0.0796)**	**(0.0793)**
Average "Lifetime Income"	0.2199	0.0964	0.0624
	(0.0062)	(0.0065)	(0.0070)
Controls:			
Basic Background Measures		X	X
Additional Controls			X

Notes:
(1) Income is measured in $10,000 and is discounted at an annual rate of $r = 0.05$.
(2) All specifications control for the child's age, race, and gender and for the mother's age.
(3) Average "lifetime income" refers to the discounted present value of all available income measures after birth of the child.
(4) Basic background measures include the mother's education, AFQT score, and family background (foreign born, rural residence at age 14, living with both parents at age 14); the highest grade completed by the child's grandparents; and year dummies.
(5) Additional controls include the number of adults and children/dependents in the household, current marital status of the mother, age and education of the spouse, and whether or not the mother is currently living with her parents.

variety of family background characteristics). These results are presented in table 1 for a few different specifications. Consistent with intragenerational credit constraints, we generally find negative effects of the slope on children; although, the estimates are not always statistically significant.[9] To interpret these estimates, it is useful to note that incomes grow about $2,300 more per year for those with high earnings growth (90^{th} percentile) relative to those with low earnings growth (10^{th} percentile). The most conservative estimates in table 1 suggest that this difference is associated with a 0.027 standard deviation difference in PIAT test scores.

We next estimate the effects of family income earned at all earlier ages (allowing those effects to vary by age) on PIAT scores, controlling for a variety of individual and family background characteristics as well as average discounted lifetime income. Here, we are interested in determining whether the effect of income on subsequent PIAT scores depends on the age at which that income was received. In the absence of intragenerational borrowing constraints, income earned at all ages (once appropriately discounted) should have the same effect on child achievement. By contrast, when borrowing constraints are binding for some families, income received at earlier ages should have a greater effect than income received at later ages.

We explore the effects of income at different ages assuming that the effects of income on a child's test score depend linearly on the age at which income is earned and the age at which the achievement test is taken. Specifically, we assume that the effects of family income for child i earned at age j, $I_{i,j}$, on an achievement test score at age a, $T_{i,a}$ is given by

$$\frac{\partial T_{i,a}}{\partial I_{i,j}} = \alpha_1 + \alpha_2 j + \alpha_3 a \,.$$

In this case, α_2 tells us how income earned at different ages affects subsequent child test scores, while α_3 tells us at what ages test scores respond most to changes in past income. Table 2 reports

Table 2: Effects of Past Income on PIAT Scores

	1	2	3	4
Long-Term Effects:				
α_1	–0.0140	–0.0102	–0.0101	0.0260
	(0.0031)	(0.0038)	(0.0064)	(0.0098)
α_2	**–0.0013**	**–0.0011**	**–0.0011**	**–0.0020**
	(0.0006)	**(0.0008)**	**(0.0008)**	**(0.0008)**
α_3	0.0027	0.0025	0.0025	–0.0019
	(0.0004)	(0.0005)	(0.0006)	(0.0011)
Temporary Effect of Current Income		–0.0075	–0.0075	–0.0044
		(0.0136)	(0.0139)	(0.0139)
Avg. "Lifetime Income"			–0.0006	–0.3610
			(0.0272)	(0.0785)
Avg. "Lifetime Income" * Current Age				0.0476
				(0.0097)

Notes:
(1) Long-term effect of income at age J on child PIAT score at age A is $\alpha_1 + \alpha_2 J + \alpha_3 A$.
(2) Income is measured in $10,000 and is discounted at an annual rate of $r = 0.05$.
(3) Average "lifetime income" refers to the discounted present value of all available income measures after birth of the child.
(4) All specifications control for the child's age, race, and gender; mother's age, AFQT score, education, family background (foreign born, rural residence at age 14, living with both parents at age 14), marital status, and the age and education of her spouse; the highest grade completed by the child's grandparents; number of adults and children/dependents in the household; whether or not the mother is currently living with her parents; and year dummies.

Table 3: Effects of Past, Current, and Future Income on PIAT Scores

	1	2	3
Past and Current Income	0.1698	0.0850	0.0619
	(0.0118)	(0.0114)	(0.0116)
Future Income	0.0634	0.0221	0.0111
	(0.0075)	(0.0072)	(0.0072)
Controls:			
Basic Background Measures		X	X
Additional Controls			X

Notes:
(1) Income is measured in $10,000 and is discounted at an annual rate of $r = 0.05$.
(2) All specifications control for age, race, and gender of the child and age of the mother.
(3) Basic background measures include the mother's education, AFQT score, and family background (foreign born, rural residence at age 14, living with both parents at age 14); the highest grade completed by the child's grandparents; and year dummies.
(4) Additional controls include the number of adults and children/dependents in the household, current marital status of the mother, age and education of her spouse, and whether or not the mother is currently living with her parents.

estimates of α_1, α_2, and α_3 using the NLSY data. Note that specifications 2–4 also account for an additional affect of current income on contemporaneous test scores, while specifications 3 and 4 incorporate an independent effect of discounted lifetime income.[10] Focusing on the estimates of α_2, which are all negative, we see that income earned at later ages has a smaller effect on subsequent child achievement scores than does income earned at earlier ages. These estimates are statistically significant at the 0.05 level in specifications 1 and 4. Consistent with intragenerational borrowing constraints, the estimates suggest that shifting $10,000 in family income from age ten to the first year of a child's life would increase subsequent test scores by .01 to .02 standard deviations. More generally, the estimates suggest that past income has a positive effect on test scores at ages when the math and reading tests were administered (recall that tests were not administered to children before age five). The final specification suggests that lifetime income has growing effects as a child ages. That is, children from wealthier families (as measured by lifetime family income) perform increasingly well over time relative to children from less-fortunate families.

The permanent-income hypothesis does not distinguish between income earned in the past and that earned in the future. If individuals are reasonably certain about their future income prospects and unaffected by intragenerational borrowing constraints, income earned in the past, present, and future should all affect child test scores equally. Because the NLSY offers panel data over a long time period, it is possible to observe family income measured both before and after some tests are taken by children. Table 3 reports estimates from regressions of children's test scores on the present value of past and current income as well as the present value of all future income. Columns 2 and 3 control for additional family-background variables as described earlier. All of these estimates suggest that past and current income have a significantly greater effect on test scores than does future income. In the final column, the coefficient estimate for future income is not statistically different from zero, while the effect of past income is significantly positive at 0.06. This result is consistent with both intergenerational and intragenerational borrowing constraints.

An obvious concern with this approach is uncertainty about future earnings. If individuals are completely uncertain about future income, actual realizations of that income process should have no effect on current decisions or outcomes, even if individuals are not borrowing constrained. However, since future income primarily represents income earned over the next one to five years, it seems unlikely that it is all that uncertain for most families. To examine the role of uncertainty more formally, we ask whether the results

of table 3 hold for families with fairly predictable income profiles. To measure the predictability of income, we estimate family-specific log income regressions on age and age-squared, using all income observations after the child's birth. We can then compute two potential measures of uncertainty (or variability, at least): (i) R^2 statistics that measure the fraction of the variance in log income that can be explained by age and age-squared alone, and (ii) the square root of the mean squared error (RMSE) from the regression. We separate individuals according to these measures of ability and separately run regressions of children's test scores on past/current and future family income. If the insignificant coefficient on future income in table 3 is due primarily to uncertainty in future earnings, we expect that future income should have effects similar to past/current income among those with predictable earnings profiles (i.e., a high R^2 statistic or a low RMSE). This is not the case. Using either measure of predictability, we find that past/current income has a significantly greater effect than future income among those with highly predictable income profiles. In a combined measure that takes only those individuals with an R^2 above 0.75 and those within the lowest quartile of RMSE, we find a more dramatic difference in coefficients on past/current and future income than we observe in table 3.[11] While predictability in an *ex post* sense (as implied by our measures) does not necessarily imply a high degree of predictability in an *ex ante* sense (i.e., low uncertainty), these results are at least consistent with a more important role for intragenerational borrowing constraints than for uncertainty.

CONCLUSIONS

While none of these "tests" for intragenerational borrowing constraints are perfect, we view the combination of all three sets of results as convincing evidence that some families with young children are constrained. At the very least, this evidence suggests that a better understanding of the role credit constraints play in determining families' investments in children is warranted.

The existence of intragenerational borrowing constraints suggests a positive role for government policy. An inexpensive way to address such concerns may be to expand borrowing opportunities for families with younger children. Simply allowing young parents to borrow against their future earnings in order to pay for early-childhood-development programs or to finance private-school tuition should help alleviate intragenerational constraints. For example, a program modeled on the federal PLUS loan program for parents of college students could be extended to qualifying parents of younger children. It is important to note, however, that expanding borrowing opportunities in this way is not likely to address intergenerational borrowing constraints, which are also likely to be important. Dealing with poor parents' inability to borrow against their children's future earnings prospects is a more complicated problem, which is likely to require a redistribution of resources from wealthier families to poorer families. Subsidies for early childhood programs and private schooling can alleviate problems with borrowing constraints, but they come at a sizeable cost. To the extent that subsidies are granted to all children, wealthier families will tend to respond to such incentives by overspending on childhood investments at high costs to taxpayers. Thus, a more efficient approach may require targeting subsidies to lower-income families that are most likely to be affected by borrowing constraints of one form or the other. But this requires redistribution from the middle and upper classes to the less fortunate. Such redistribution is not necessary to address intragenerational borrowing constraints since they can be alleviated by expanded borrowing opportunities for younger parents.

ENDNOTES

[1] Lance Lochner based his remarks at the conference on this paper.

[2] See Carneiro and Heckman (2002) or Cunha et al. (forthcoming) for a detailed discussion of these issues.

[3] Income measures are based on the discounted present value of family income from the child's birth through age 12. Income is deflated using the Consumer Price Index for Urban Consumers and discounted at an annual rate of 5 percent. See the data discussion in the "Evidence on Borrowing Constraints and Child Achievement" section for further details.

[4] Only recently have economists (e.g., Restuccia and Urrutia 2004; Caucutt and Lochner 2005; and Cunha and Heckman 2005) begun to consider multiple investment periods at young ages. However, only Caucutt and Lochner (2005) examine the role of early versus late borrowing constraints. Restuccia and Urrutia (2004) abstract from financial asset accumulation, while Cunha and Heckman (2005) shut down late borrowing altogether and do not focus on early borrowing constraints.

[5] If investments are perfectly substitutable over time, then the timing of income may not matter, even if a family is borrowing constrained. Strictly speaking, our tests offer a joint test against perfect substitutability and borrowing constraints.

[6] One might expect more able parents to earn more of their income later because they invest more in their human capital early on. Then, if abler parents have abler children (a problem in the test of the intergenerational constraint), children in families earning more of their income at later ages should be innately abler. This suggests that, in the absence of credit constraints, children from families earning their income earlier should perform worse, on average—the opposite of the prediction based on credit constraints. Thus, a positive intergenerational correlation in ability may make it difficult to find evidence of intragenerational borrowing constraints.

[7] We use a discount rate of $r = 0.05$; however, other reasonable rates yield similar conclusions.

[8] See Caucutt and Lochner (2005) for a detailed discussion of the data.

[9] Estimates of the coefficient on the slope of income profiles are likely to be biased upward for two reasons. First, since we only use income over a limited number of years in computing average lifetime income, this number will tend to be too low for those with a steeper slope relative to those with a flatter income profile. Because average lifetime income has a positive effect on children, this mismeasurement will tend to bias estimates of the coefficient on the slope of income profiles upward. However, this is unlikely to have much effect on our estimates since income far into the future is heavily discounted. A second potential source of bias would arise if unobserved differences across families are related to the slope of family income profiles. If family income profiles are rising because parents are accumulating human capital, then parents with the steepest earnings profiles will tend to accumulate the most human capital. If those parents also tend to invest more in their children or if their investments are more productive, this will produce upwardly biased estimates of the coefficient on the slope of the income profile.

[10] These effects are estimated based on the following regression:

$$T_{i,a} = X_{i,a}\gamma_x + \bar{I}_i\gamma_I + \alpha_0 I_{i,a} + \alpha_1\left(\sum_{j=0}^{a} I_{i,j}\right) + \alpha_2\left(\sum_{j=0}^{a} jI_{i,j}\right) + \alpha_3\left(\sum_{j=0}^{a} aI_{i,j}\right) + \varepsilon_{i,a},$$

where $X_{i,a}$ represents background characteristics for child i at age a, and $\bar{I}_i$ represents average family income from the child's birth through the final period of observation. The inclusion of $I_{i,a}$ allows for an additional temporary effect of contemporaneous income on test scores, so

$$\frac{\partial T_{i,a}}{\partial I_{i,a}} = \alpha_0 + \alpha_1 + (\alpha_2 + \alpha_3)a.$$

[11] For these individuals, the coefficient on past and current income is 0.149 (standard error of 0.053), while the coefficient on future income is –0.035 (standard error of 0.030).

REFERENCES

Blau, D., and J. Currie. Forthcoming. "Who's Minding the Kids? Preschool, Day Care, and After-School Care," in E. Hanushek and F. Welch, eds., *Handbook of Education Economics*. North-Holland, Amsterdam.

Cameron, S., and J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy* 106(2): 262–333.

Cameron, S., and J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy* 109(3): 455–99

Carneiro, P., and J. Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal* 112: 705–34.

Caucutt, E., and L. Lochner. 2005. "Early and Late Human Capital Investments, Borrowing Constraints, and the Family," Working Paper.

Cunha, F., and J. Heckman. 2005. "The Technology of Skill Formation," Working Paper.

Cunha, F., J. Heckman, L. Lochner, and D. Masterov. Forthcoming. "Interpreting the Evidence on Lifecycle Skill Formation," in E. Hanushek and F. Welch, eds., *Handbook of Education Economics.* North-Holland, Amsterdam.

Dahl, G., and L. Lochner. 2005. "The Impact of Family Income on Child Achievement," Working Paper.

Duncan, G., and J. Brooks-Gunn. 1997. *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

Ellwood, D., and T. Kane. 2000. *Who is Getting a College Education? Family Background and the Growing Gaps in Enrollment.* New York: Russell Sage Foundation.

Karoly, L., et al. 1998. *Investing in Our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Interventions.* Santa Monica: The RAND Corporation.

Levy, D., and G. Duncan. 1999. "Using Sibling Samples to Assess the Effect of Childhood Family Income on Completed Schooling," Working Paper.

Restuccia, D., and C. Urrutia. 2004. "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review* 94(5): 1354–78.

Zigler, E. 1994. "Reshaping Early Childhood Intervention to be a More Effective Weapon against Poverty," *American Journal of Community Psychology* 22(1): 37–47.

Early Childhood Development on a Large Scale [1]

Rob Grunewald
Art Rolnick

For well over 20 years, government leaders at the state and local levels have been deeply engaged in efforts to promote economic development. Their concerns are understandable: A stronger economy enables citizens to better engage in "the pursuit of happiness" that our founding fathers hoped to guarantee. To the degree that public leaders can assist economic growth, they facilitate that pursuit.

Unfortunately, many economic-development schemes using public dollars are at best a zero-sum game. In the name of economic development and creating new jobs, virtually every state in the union has a history of trying to lure new companies with public subsidies. Previous studies have shown that the case for these so-called bidding wars is shortsighted and fundamentally flawed (Burstein and Rolnick 1995). From a national perspective, jobs are not created—they are only relocated; the public return is at most zero. And the economic gains that seem apparent at state and local levels are also suspect because they would likely have been realized without the subsidies. In other words, what often passes for economic development and sound public investment is neither.

Persuasive economic research indicates that there is a far more promising approach to economic development with government assistance. It rests not on an externally oriented strategy of offering subsidies to attract private companies, but rather on government support of those much closer to home—quite literally: our youngest children. This research shows that by investing in early childhood education, governments—in partnership with private firms and nonprofit foundations—can reap extraordinarily high economic returns, benefits that are low risk and long lived.

In this essay, we put forth a pragmatic proposal for economic development at the state and local levels that capitalizes on the high returns investment in early childhood education can yield. Our proposal envisions a private/public endowment that would fund early childhood development scholarships for all at-risk children. The scholarships would cover the expense of parent mentoring as well as tuition for children to attend qualified ECD programs. Government support of the endowment would provide the assurance of long-term commitment, and the market-based nature of the ECD and mentor programs would promise innovation, outcome accountability, and quality improvement.

We don't pretend to have all the answers to economic development, but we're quite certain that investing in early childhood education is more likely to create a vibrant economy than using public funds to lure a sports team by building a new stadium or to attract an automaker by providing tax breaks. Investing in the education of children in their earliest years makes sense as an economic development strategy precisely because the returns are large, reliable, and reaped by both the individuals involved and the general public. As economists, we are trained to be skeptical of policies that interfere with market forces. But when it comes to early childhood education, we're confident that this is one policy with a high public return.

EXTRAORDINARY RETURNS

Policymakers rarely view early childhood development as economic development. They should. Careful academic research demonstrates that tax dollars spent on ECD provide extraordinary returns compared with investments in the public, and even private, sector. The potential return from a focused, high-quality ECD program is as high as 16 percent per year. Some of these benefits are private gains for the children involved in the form of higher wages later in life. But the broader economy also benefits because individuals who participate in high-quality ECD programs have greater skills than they otherwise would, and they're able to

contribute productively to their local economies. Thus, it's estimated, the annual *public* return to good ECD programs is 12 percent.

The promise of ECD programs is based on fundamental facts about early human development. A child's quality of life and the contributions that child makes to society as an adult can be traced to his or her first years of life. From birth until about five years old, a child undergoes tremendous development. If this period of life includes support for growth in language, motor skills, adaptive abilities, and social-emotional functioning, the child is more likely to succeed in school and to later contribute to society. Conversely, without support during these early years, a child is more likely to drop out of school, depend on welfare benefits, and commit crime—thereby imposing significant costs on society. ECD programs recognize this potential—and this risk—and seek to nurture healthy development from the earliest years.

In a previous essay, we reviewed several longitudinal evaluations that all reached essentially the same conclusion: The return to ECD programs that focus on at-risk families far exceeds the return to other projects that are funded as economic development (Grunewald and Rolnick 2003). Cost-benefit analyses of the Perry Preschool Program, the Abecedarian Project, the Chicago Child-Parent Centers and the Elmira Prenatal/Early Infancy Project showed returns ranging from $3 to almost $9 for every dollar invested. This implies an annual rate of return, adjusted for inflation, between 7 percent and 16 percent.

A more recent analysis suggests that these figures may actually understate the true returns: The November 2004 follow-up study on the Perry Preschool Program 40 years after its inception calculates the total benefit-cost ratio at $17 for every dollar invested, confirming that the benefits of ECD continue well into adulthood. Other recent studies of ECD programs in Michigan, New Jersey, Oklahoma, and elsewhere provide additional evidence that investments to help young children prepare for school and beyond pay large dividends to society.

ESSENTIAL ELEMENTS FOR LARGE-SCALE ECD

These findings, promising though they are, pose a challenge: Small-scale ECD programs have been shown to work, but can their success be reproduced on a much larger scale? There are reasons to be skeptical; some recent attempts at scaling up ECD have been disappointing. However, based on a careful review of past and current programs—those that have failed as well as those that thrive—we believe that large-scale efforts can succeed if they incorporate four key features: careful focus, parental involvement, outcome orientation, and long-term commitment. We further believe that to achieve these characteristics, large-scale ECD programs must be structured so as to blend the benefits of market incentives and long-term government support. In the discussion that follows, we describe the importance of these features and explain how a hybrid structure can achieve them.

Focus on At-Risk Children

Without doubt, all children benefit from investment in early childhood development. Given the inherent limits of tax revenue, however, we suggest that government resources for ECD programs be focused on those children at highest risk for developmental deficits. Conditions that can indicate whether a child is at risk include low family income, violence or neglect in the home, low parent-education levels, low birth weight, and parent chemical addiction.

Children from economically advantaged families are likely to thrive without additional government resources. But children from low-income families need additional support. Hence, to maximize the impact of scarce public dollars, large-scale ECD programs should focus on at-risk children.

Encourage Parental Involvement

Research shows unequivocally that parental involvement is a crucial ingredient in the success of ECD programs. When parents receive training in why and how to nurture their children's development, they're better able to nurture their children at home, outside of ECD program hours. Comprehensive ECD programs should therefore be designed to encourage parents to participate.

Assess Outcomes Regularly

ECD programs succeed when their goals are clear, explicit, and carefully monitored. Since their primary goal is the improved functioning of the children in their care, these programs should perform regular assessments of cognitive and social-emotional outcomes. And the programs themselves should be oriented toward achieving constant progress for each child. Outcome assessments allow for individual progress reviews, for curriculum improvements, and for staff and program accountability.

Provide Long-Term Commitment

Children thrive in secure, consistent environments. Similarly, programs designed to expand the cognitive and psychological abilities of children need the security of long-term commitment. This is not to say that such programs shouldn't be challenged to improve continuously, but children, parents, and ECD providers will benefit if they're assured of financial backing and institutional support as long as specified standards and outcomes are met.

AN ENDOWMENT WITH A MARKET ORIENTATION

Achieving these characteristics in large-scale ECD programs requires the flexibility, innovation, and incentives that are inherent to markets and the long-term assurance and stability that government backing provides. To establish a successful, large-scale ECD program, therefore, we propose a permanent scholarship fund for all families with at-risk children. Similar to endowments in higher education, earnings from an endowment for ECD would be used to provide scholarships for children in low-income families who aren't able to afford a quality ECD program. The program would be financed and managed as follows: A state or local government, in partnership with the private sector and the federal government, would create an ECD endowment to fund the scholarships. The scholarships would cover child tuition to qualified ECD programs *plus* the cost of parent mentoring to ensure parental involvement. Scholarships would be outcome based, meaning that they would include incentives for achieving significant progress toward the life and learning skills needed to succeed in school.

Tuition-Plus Scholarships

The central component of our market-oriented approach to ECD is tuition-plus scholarships. A tuition-plus scholarship would cover tuition for the at-risk child to a qualified ECD program plus the cost of high-quality parent mentoring and home visits. Parent mentoring would include parent education; information about available financial, health and human services; and guidance on selecting an ECD program.

Through parent decisions and provider responses, the market would determine the structure of the ECD industry. Market participants would include ECD providers from the public and private sectors, which represent a mix of preschools, child-care providers, and home-visiting programs. The market structure, however, would be influenced by standards set by an executive board that manages the ECD endowment. ECD providers would have to comply with these standards in order to register the scholarship children. The standards would be consistent with the cognitive and social-emotional development needed to succeed in school. We envision a diverse mix of providers competing to serve at-risk children, leveraging the existing ECD infrastructure, and opening the door for new providers.

To encourage ECD providers to compete for the most severely at-risk children, scholarships would be based on initial conditions. To this end, the scholarship amount would be highest for a child with multiple risk factors. This would create an incentive for providers to register children who require more costly resources.

We should note several additional features of the scholarships. First, a partial scholarship could be layered on top of existing funding streams that providers currently receive. Second, the scholarship provided to the family would be for qualified ECD services only; actual payments would flow from the endowment directly to the family-chosen provider. And third, the scholarships would include financial incentives to providers based on accountability measures.

The Mentoring Program

Home visits by qualified mentors are among the best ways to achieve a high degree of parental involvement. To this end, as noted, the scholarships would provide funds for qualified mentors. Mentor qualifications would include ECD training and parent training and

counseling on issues related to health as well as education. Mentors would help parents decide which of the qualified ECD providers best meets the family's needs and would advise parents throughout the program.

Research shows that reaching children with multiple risk factors as early as possible is essential; even three years old may be too late. Therefore, we suggest that while scholarships would pay tuition for a child to attend an ECD program beginning at age three, the parent mentoring program could start much earlier.

The Value of a Market Orientation

A market-oriented approach would directly involve the parents with their children's education; research shows this is vital. Parents would be empowered to choose among the various providers and select one based on location, hours of service, quality of program and other features, much as they would any other product or service. The process of self-education and provider choice would itself involve the parent.

Furthermore, the approach would be outcome-based, so scholarships would include financial incentives focused on performance and would encourage innovation. While programs would have to meet requirements to accept children with scholarships, providers would have room for innovation in providing services.

Unlike a top-down, planned system, the ECD industry would be shaped by the market, through micro-level decisions by parents and responses by providers. This approach would allow the diverse mix of current providers and new entrants to find the best means to supply high-quality ECD.

THE ADVANTAGES OF AN ECD ENDOWMENT

An endowed fund for ECD represents a permanent commitment and effectively leverages resources by public and private stakeholders. Because the endowment would provide a stable funding source, we would expect the market response to be better than otherwise. A permanent commitment sends a market signal to providers that they can expect a consistent demand for their product. By drawing up a business plan that demonstrates it can successfully attract scholarship children, an ECD provider can leverage funds for capital expansions or improvements from low-interest loan sources and philanthropic organizations; lenders will be reassured by the stability of the ECD endowment.

State governments are well-positioned to provide leadership to build a public/private endowment. Just as they do for capital campaigns for physical buildings, state governments can lead drives to build human capital through ECD. The state can encourage contributions to the fund by matching donations and providing tax credits. A donation of $50,000 to $150,000 would help provide ECD for an at-risk child every year into perpetuity.

As mentioned above, a board of directors with representatives from the public and private sectors would provide oversight for the endowment. Under the board's supervision, the program's executive director would determine the number of families eligible for scholarships, develop a mentoring program that would work with existing organizations, and design incentives for providers to ensure desired outcomes while promoting best practices.

How Much Money Would the Endowment Need?

Based on costs used in previous studies and current programs for at-risk children, we estimate that total resources needed to fund an annual scholarship for a high-quality ECD program for an at risk three- or four-year-old child would be about $10,000 to $15,000 for a full-day program that includes parent mentoring. The scholarship either would cover the full cost of tuition or would be layered on top of existing private and public funds, such as child-care subsidies, to enhance quality features that correlate with school-readiness outcomes.

The endowment board could vary the amount of the scholarship to reach children in families just over the poverty line on a sliding scale or increase the amount of the scholarship for children facing multiple risk factors. The board could also consider providing scholarships for families that don't qualify based on income, but whose children are identified with risk factors other than living in poverty.

To derive an approximate dollar amount for the endowment, therefore, a state would have to estimate the number of children to be covered, multiply that by the average scholarship, and calculate the investment return for the interest derived from

investing the endowment funds in low-risk government or corporate bonds.

In Minnesota, for example, we estimate that in order to ensure that all three- and four-year-old children living below poverty receive high-quality ECD, the state needs about an additional $90 million annually. For children who aren't already involved in an ECD program, the scholarship would give them access. For children who are enrolled in a child-care center or preschool, the scholarship would ensure that the quality is at the necessary level to meet school-readiness goals. A one-time outlay of about $1.5 billion would create an endowment that could provide scholarships to the families of children living below poverty on an annual basis. With the endowment's funds invested in corporate AAA bonds, earning about 6 percent to 7 percent per year, we estimate that $90 million in annual earnings would cover the costs of scholarships, pay for program monitoring and assessments, and supplement existing revenue sources as needed for early childhood screening and teacher training reimbursement programs.

CONCLUSION

The evidence is clear that investments in ECD for at-risk children pay a high public return. Helping our youngest children develop their life and learning skills results in better citizens and more productive workers. Compared with the billions of dollars spent each year on high-risk economic development schemes, an investment in ECD is a far better and far more secure economic development tool. Now is the time to capitalize on this knowledge.

We argue that a market-oriented approach to ECD has several strong features. The present ECD landscape includes a variety of providers from the public and private sectors; a market-oriented approach would help improve the access and quality of ECD without creating additional bureaucracy. Focusing on at-risk children and encouraging direct parental involvement would help reach those children and families with the greatest need for ECD programs. Providers would receive incentives for successful outcomes and make local decisions on how to best achieve strong results. Finally, with a long-term, demand-side commitment through the creation of state-level private/public endowments, we expect a strong response from the supply side of the ECD market.

This essay outlines a market-oriented approach to ECD, and we acknowledge that the proposal should be tested in pilot projects to learn from practical experience. For example, a pilot project that distributes 200 or 300 scholarships over a five-year period would provide experience and lessons about implementing a scholarship system. With this information, researchers, policymakers, and practitioners could convene to make informed recommendations.

In our view, the case is closed for why we must invest in ECD. Now it is time to design and implement a system that will help society realize on a large scale the extraordinary returns that high-quality ECD programs have shown they can deliver.

ENDNOTE

[1] Art Rolnick based his remarks at the conference on this article (*The Region* 2005), which itself is an abbreviated version of "A Proposal for Achieving High Returns on Early Childhood Development" by Rob Grunewald and Art Rolnick (Federal Reserve Bank of Minneapolis, Prepared for "Building the Economic Case for Investments in Preschool," Washington, D.C., December 3, 2004. Convened by the Committee for Economic Development, with support from The Pew Charitable Trusts and PNC Financial Services Group.)

REFERENCES

Burstein, Melvin L., and Arthur J. Rolnick. 1995. "Congress Should End the Economic War among the States," Federal Reserve Bank of Minneapolis Annual Report, *The Region* 9.

Grunewald, Rob, and Art Rolnick. 2003. "Early Childhood Development: Economic Development with a High Public Return," *The Region* 17, December Supplement.

How Much Does Studying Matter? [1]

Ralph Stinebrickner
Todd R. Stinebrickner

Understanding the impact of most potential education-policy changes is made difficult by the reality that the large majority of variation in student outcomes is unexplained by traditionally observable individual and school characteristics. Thus, it is important that while a very large amount of recent attention has been paid to understanding the determinants of educational outcomes, knowledge of the causal impact of the most fundamental input in the education production function—students' own study time and effort—has remained essentially non-existent.

One primary reason for the current void in our understanding is that standard data sources have not traditionally collected information about how much time students spend studying. The very small amount of existing work that has provided direct evidence about the relationship between studying and academic performance focused on collecting measures of study effort and obtained estimates of the (conditional) correlation between the number of hours that a person studies and his or her academic performance. In the first of this work, Schuman et al. (1985), over the course of a ten-year period, took four different measurement approaches in an explicit attempt to "produce a positive relation between amount of study and GPA" at the University of Michigan and found that none of the approaches were "very successful in yielding the hypothesized substantial association." Similar replication results at different schools by Hill (1991) and Rau and Durand (2000) produced generally similar results.

The bias associated with viewing the descriptive relationships in previous work as estimates of the causal role that studying plays in the grade-production process arises, in part, because students who spend more time studying may be different in unobserved ways related to, say, ability than those who spend less time studying. However, further confounding this "endogeneity" problem is the possibility that individuals who receive bad grade shocks or have difficult classes during a particular semester may react by changing their effort during that semester. Not only is it not possible to know the size of the bias that is present if one views the correlations found in previous papers as estimates of the causal effect, but it is also not possible to know the direction of the bias. Thus, given the central policy importance of effort and the reality that no previous work has addressed the endogeneity problem that may very well be present, it should perhaps be disconcerting that a recent review of the current evidence led Schuman (2000) to write that "for now, we can conclude that the amount of studying has some but not a great deal to do with students' achievement as measured by grades, especially GPA."

Ideal for learning about the importance of studying would be a random experiment in which two groups of students who are identical in all respects at the beginning of school are forced to study different amounts during school but continue to behave identically in all other ways (class attendance, sleeping, drinking, study efficiency, paid employment, etc.) that could influence the outcome of interest. In this case, one can learn about the causal impact of studying simply by examining whether grade performance differs between the two groups. In this paper, we examine the effect of studying on college grade performance by taking advantage of a "natural experiment" that we find closely resembles this ideal experiment. More specifically, our approach takes advantage of the fact that some students in our sample are randomly assigned roommates who bring video games with them at the beginning of the school year, while the remainder of the students in our sample are randomly assigned roommates who do not.

In the next section, we describe a unique survey project at Berea College that provides our data. In the "Results" section, we describe the approach we take to identify the causal effect of studying in more detail then present results. And in the conclusion, we discuss the importance of this work for policymakers, including the fact that it provides perhaps the first direct evidence about an underlying avenue through which peer effects operate.

A GENERAL OVERVIEW OF THE BEREA PANEL STUDY

Located in central Kentucky, where the "bluegrass meets the foothills of the Appalachian Mountains," Berea College is a liberal arts college that operates under a mission of providing educational opportunities to students of "great promise but limited economic resources." The survey data used in this paper are part of the Berea Panel Study (BPS) that Todd Stinebrickner and Ralph Stinebrickner (hereafter referred to as S&S) started with the explicit objective of collecting the type of detailed information that is necessary to provide a comprehensive view of the decisionmaking process of students from low-income families. The BPS involves surveying two cohorts of students approximately 12 times each year.

Of direct relevance for the analysis in this paper, a sequence of time-use surveys was administered at multiple times during each academic year. Also of relevance, the baseline and follow-up surveys collected substantial information about friends, roommates, and other information related to studying and grade performance. Student identifiers allow the survey data to be merged with Berea College's administrative data.

RESULTS

We are interested in estimating the causal effect of studying on first-semester grade point average. During the first semester, daily study effort was collected on four different weekdays using 24-hour-time diaries. We create a variable, $STUDY_i$, by averaging the number of hours that person i studies per day over the subset of days during the semester that his or her study effort is observed.[2] This variable serves as a proxy for the average amount that a person studies per day across all days during the first semester.

The intuition behind our approach for identifying the causal effect of studying is as follows: We say that a student is "treated" if the student is randomly assigned a roommate who brings a video game at the beginning of the year and "untreated" if the student is randomly assigned a roommate who does not. In the first section below, we use information about whether a student is treated to divide our sample into two groups—treated and untreated. In the second section below, we show that the presence of a video game causes students in the former group to study less, on average, than students in the latter group. In the third section below, we use the fact that roommates are randomly assigned, along with additional unique information from the BPS, to argue that it is plausible to believe that the students in the two groups are very similar in all (nonstudy) dimensions that influence grade performance. And in the last section below, we take advantage of the fact that, if this is the case, then differences in average grade performance between the groups can be attributed to differences in average study effort between the groups. Our intuitively appealing identification approach is formalized by an instrumental-variables estimation procedure.

Dividing the Sample Using the Treatment Variable

We focus on the cohort of students who entered Berea College in the fall of 2001, since this group completed a survey question that elicits whether their roommates brought video games to school. Slightly more than one-third of students at Berea either live off campus or request a specific roommate. The sample used in this paper contains information about 210 students who lived on campus and were randomly assigned roommates. Fifty-three percent of males and 24 percent of females in our sample have roommates who bring some sort of video games to school.

Does the Instrument Influence Study Decisions?

We find that, for both males and females, study effort differs in a quantitatively important manner between students in the sample whose roommates bring video games to school and students in the sample whose roommates do not bring video games to school. Specifically, the sample average of STUDY is .667 of an hour per day lower (2.924 vs. 3.591) for males who receive the video-game treatment than for males who do not receive the treatment. The sample average of STUDY is .467 of an hour per day lower (3.226 vs. 3.693) for females who receive the video-game

treatment than for females who do not receive the treatment. It is not possible to reject the null hypothesis that the effect of the treatment is the same for males as it is for females.

Pooling the male and female observations, we regress $STUDY_i$ on whether the person received the treatment, whether the person is male, the person's score on the American College Test (ACT), and whether the person is black. We find an estimate (std. error) of -.565 (.241), which indicates that, consistent with the unpooled results in the previous paragraph, the treatment reduces study time by over one-half of an hour per day. Given that students in the sample study 3.48 hours per day on average, the estimated effect is quantitatively important, and a test of the null hypothesis that the treatment has no effect on study effort is rejected at all levels of statistical significance greater than .02.

Are the Two Groups the Same in All Other (Nonstudy) Dimensions?

In order for our approach to be valid, it must be the case that the treatment's only influence on a student's grade performance comes through its effect on the student's study effort. There are two avenues through which this requirement could be violated. First, it would be violated if whether a student receives the treatment is related to the student's unobserved characteristics at the time of college entrance, in which case the treated and untreated groups would not be identical at the time of college entrance. Second, it would be violated if, in addition to affecting decisions about study time, the treatment also affects other behaviors that take place during the first semester and that influence grade performance. Roommates who bring video games to school may be different in observable and unobservable ways than those who do not bring them. As a result, in thinking about these two avenues through which this requirement could be violated, it is necessary to take into account that the treatment involves both the physical presence of the video game(s) and the presence of whatever type of roommate accompanies the game(s). However, it is important to note at this point that while it is perhaps tempting a priori to view students who bring video games as types who will tend to encourage a variety of harmful behaviors in their peers, this does not seem to be the case. Specifically, as detailed in S&S (2005), we find no evidence that students at Berea who bring video games have harmful sleep habits, are of lower observed ability, are less likely to attend class, or are more likely to drink alcohol.

The First Avenue: Student Characteristics at the Time of College Entrance

The random assignment of roommates in our sample plays the key role in ensuring that the condition that the groups are the same in all nonstudy dimensions is not violated by the first avenue described in the previous paragraph. If students were choosing roommates, they would also (perhaps quite indirectly) be choosing whether roommates bring video games. In this case, the amount that a student intends to study and other factors such as the student's ability could be related to whether his roommate brings a video game. The random assignment of roommates guarantees that, conditional on a student's sex, students in the sample who receive the treatment come from the same population distribution as students in the sample who do not receive the treatment.

The Second Avenue: Student Behaviors during College Other Than Study Effort

With respect to whether the condition that the two groups are the same in all nonstudy dimensions could be violated through the second avenue described above, there seem to be two general possibilities. One possibility is that, in addition to reducing the amount of time spent studying, students who receive the treatment also reduce time spent in other activities that influence grade performance directly. Seemingly most important among these other activities is class attendance, which is unique in that it directly influences the amount of course material to which a person is exposed. However, also potentially important are other activities that influence how rested or clear-thinking a person is at the time he or she is studying or attending class. The activities that seem most likely to fit this description are sleeping, drinking/partying, and paid employment. In S&S (2004), we carefully examine whether differences in class attendance, sleeping, drinking/partying, and paid employment exist between the treated and untreated groups. We find no evidence that this is the case.

The other way that the condition that the two groups are the same in all nonstudy dimensions could be violated through the second avenue is if, in addition to reducing the amount that a student studies, the

treatment also causes a student to study less efficiently. Unique questions in the Berea Panel Study allow us to examine this possibility directly. For example, in S&S (2004) we find no evidence that the presence of a video game implies that a student changes where he or she studies or that the student is studying with the television on. We also find no evidence that treated students are more likely to be assisted with their coursework by their roommates than untreated students.

While it would never be possible to empirically establish with full certainty that our two groups are identical in all ways other than study effort, the random assignment of roommates ensures that students in treated and untreated groups are identical in the population at the time of entrance, and the unique features of our survey-collection efforts allow us to credibly examine the remaining reasons that this might not be true. Thus, in the remainder of the paper we assume that the two groups are the same in all ways other than study effort.

Estimates of the Causal Effect of Studying on Grade Performance

As described earlier, the intuition underlying our identification strategy is straightforward with the binary treatment variable. Our results from the previous section suggest that, conditional on sex, factors other than study effort are similar for treated and nontreated students in the population. Thus, if studying has no effect on grade performance, grade performance would be identical (conditional on sex) for the treated and untreated groups, even though study effort is different between the groups. We find that males in the sample who receive the treatment have grades that are .239 lower than males who do not receive the treatment, and females in the sample who receive the treatment have grades that are .128 lower than females who do not receive the treatment. Estimates of the causal effect of studying must take into account the differences in average study effort that led to these differences in average grades. So, for example, given that the treatment reduces study effort by .667 of an hour for males, an estimate of the effect of studying on GPA obtained from the sample of males would be .239/.667=.358. Similarly, an estimate of the effect of studying on GPA obtained from the sample of females would be .128/.467=.274.

The use of an instrumental-variable estimator allows us to pool all male and female observations in our sample, to condition on a variety of observable characteristics other than study effort, and to conduct inference that recognizes that we would obtain different estimates if we were to use a different sample of the same size. We find an instrumental-variable estimate of .356, which indicates that an additional hour of studying per day causes first-semester grade point average to increase by .356. We note that the standard error associated with the estimate is large, which implies that the estimate would vary considerably from sample to sample. Thus, although a test of the null hypothesis that studying has no effect on grade performance produces a t statistic of 1.748 and the test is rejected at significance levels greater than .08, it is important to keep in mind that nontrivial uncertainty exists about the size of the true causal effect in the population of all students. Nonetheless, unlike the previous work in this literature, the instrumental-variable estimate suggests that studying may play a very important role in academic performance. In the conclusion, we discuss the quantitative importance of the estimated effect in the context of a brief policy discussion.

CONCLUSION

To the best of our knowledge, this work represents the only evidence about the causal relationship between study effort and grade production. Many policy decisions depend on the extent to which college outcomes of interest are driven by decisions that take place after students arrive at college rather than by background factors that influence students before they arrive at college. Thus, it is important that our estimates suggest that human-capital accumulation may be far from predetermined at the time of college entrance. For example, an increase in study effort of one hour per day (an increase of approximately .67 of a standard deviation in our sample) is estimated to have the same effect on grades as a 5.74 point increase in ACT scores (an increase of 1.54 standard deviations in our sample and 1.21 standard deviations among all ACT test takers). In addition, the reduced form effect of being assigned a roommate with a video game is estimated to have the same effect on grades as a 3.10 point increase in ACT scores (an increase of .83 of a standard deviation in our sample and .65 of a standard deviation among all ACT test takers).

While not its primary focus, this paper also makes an important contribution to the peer-effects literature in general and particularly to the peer-effects literature that achieves identification by using college roommates. The goal of the empirical peer-effects literature has been to look for empirical evidence that documents that peer effects can matter. This paper provides depth to that literature by not only providing strong evidence that peer effects can matter, but also by providing perhaps the first direct evidence about an avenue (time-use) through which peer effects operate in a particular educational context (higher education). This paper also makes a contribution to a substantial literature outside of economics by establishing that video games can have a large causal effect on academic outcomes.

Finally, it is worth noting that, because of sampling variation, a considerable amount of uncertainty exists about the population parameter of interest. Nonetheless, unlike results from the small amount of earlier work that only examined the correlation between studying and academic performance, our results indicate that the effect of studying may be very substantial. Certainly more work in this area is warranted, and our findings strongly suggest that other surveys that focus on students of school age should seriously consider collecting information about this very fundamental input in the human-capital production process.

ACKNOWLEDGEMENTS

This work was made possible by generous funding from The Mellon Foundation, The National Science Foundation, The Social Science Humanities Research Council, and support from Berea College. We are very thankful to Anne Kee, Lori Scafidi, Dianne Stinebrickner, Pam Thomas, and Albert Conley, who have played invaluable roles in the collection and organization of the data from the Berea Panel Study. The authors would like to thank John Bound, Dan Black, Brian Jacob, Lance Lochner, Jeff Smith, and seminar participants at Northwestern, Maryland, Syracuse, the University of British Columbia, the University of Toronto, and the NBER.

ENDNOTES

[1] Todd R. Stinebrickner based his remarks at the conference on this paper.

[2] Response rates were relatively high on these surveys; the median person in our sample answered all four surveys, and the average number of responses was 3.11.

REFERENCES

Hill, Lester, Jr. 1991. "Effort and Reward in College: A Replication of Some Puzzling Findings," in James W. Neuliep, ed., *Replication Research in the Social Sciences*. Newbury Park, CA: Sage, 139–56.

Rau, William, and Ann Durand. 2000. "The Academic Ethic and College Grades: Does Hard Work Help Students to 'Make the Grade'?" *Sociology of Education* 73: 19–38.

Schuman, Howard, Edward Walsh, Camille Olson, and Barbara Etheridge. 1985. "Effort and Reward: The Assumption that College Grades are Affected by the Quantity of Study," *Social Forces* 63: 945–66.

Schuman, Howard. 2000. "Comment: Students' Efforts and Reward in College Settings," *Sociology of Education* 74 (1): 73–4.

Stinebrickner, Todd, and Ralph Stinebrickner. 2004. "The Causal Effect of Studying on Academic Performance," Working Paper, The University of Western Ontario.

Motivation Matters: Merit Scholarships and Student Achievement

Michael Kremer

Merit scholarships, largely diminished in the United States in the 1960s and 1970s in favor of need-based scholarships, are now making their way back into education-policy debate. While merit scholarships remain controversial, evidence from a program in Kenya suggests these scholarships can be an effective tool to raise both students' and teachers' efforts and to boost academic achievement.

Here in the United States, a growing number of states, including Georgia, Michigan, New York, and Massachusetts, now offer merit scholarships to college-bound students who perform well academically. Many other countries have similar programs, some of which target younger students. However, some educationalists oppose these kinds of scholarships on equity grounds, fearing that the benefits would mainly go to students from better-off families. Others argue that offering cash rewards for academic performance could weaken students' intrinsic motivation to learn or cause them to focus on prepping for tests at the expense of other dimensions of learning.

Unfortunately, the nature of most existing U.S. merit scholarship programs makes it hard to find reliable evidence on how exactly these programs affect students' learning. Often it is difficult to identify an appropriate group of students who was not eligible for the program to compare with a group who was. Without a credible comparison group, we cannot easily differentiate between program effects and other confounding factors that may influence achievement in education.

For example, one of the best existing sources of evidence on merit scholarships in the United States is Georgia's HOPE (Helping Outstanding Pupils Educationally) program, which awards in-state college scholarships to high school students who graduate with at least a B average. After the program was introduced in 1993, the average SAT score for Georgia high school seniors rose almost 40 points. But since all students in the state were eligible, there was no reliable way of determining with certainty whether factors other than the scholarship also contributed to the rise.

In order to approach this question more systematically, Edward Miguel, Rebecca Thornton, and I examined evidence from a merit scholarship program for primary-school girls in Kenya. In contrast to most U.S. programs, this one was phased into a number of schools in random order, allowing us to compare schools that were eligible for the program with similar schools where the scholarship had not yet been introduced. That way, the differences in educational outcomes between the two groups of students could be attributed solely to the effect of the scholarship. Our survey included information on test scores, attendance, study habits, and students' attitudes toward learning.

EDUCATION IN KENYA

Various school fees—levied to cover nonteacher costs such as textbooks, chalk, classroom repair, and other school expenses—have historically created a barrier to education in Kenya. When the program we studied was introduced in 2001–02, primary-school fees averaged approximately $6.40 per year per family (in 2003, the government abolished these fees). Families spent another $6 or so to provide each student in the household with a school uniform and other school supplies. In western Kenya—where annual per capita income is less than $1 a day—these are substantial expenses.

The Kenya Certificate of Primary Education (KCPE), given in grade eight (the end of primary school), tests students' knowledge in five subject areas: Swahili, English, geography and history, mathematics, and science. The results of the exam determine whether students are

admitted to secondary school and, if so, which schools will admit them—much as the ACT and SAT tests affect admission to tertiary education in the United States In order to prepare for the KCPE, students typically take standardized exams at the end of each school year in grades four through eight. These preparation tests carry a financial cost; students pay roughly $1–$2, depending on the year, to sit for the exams.

School-fee problems and the challenge of the placement exam both contribute to the low number of students passing from primary school to secondary school. In the part of western Kenya we examined, drop-out rates climbed precipitously in grades five through seven; only one-third of the enrolled students ever finished primary school. Drop-out rates were especially high among teenage girls.

THE GIRLS' SCHOLARSHIP PROGRAM

In 2001, a Dutch nonprofit organization called International Child Support (ICS) Africa began awarding scholarships to high-achieving, grade-six girls from Busia and Teso, two rural districts in western Kenya, for the next two academic years—that is, through the end of primary school. In order to win the award, girls had to score in the top 15 percent of the year-end, grade-six exams within their district. Each winning girl received (a) a grant of $6.40, paid to the girl's school to cover fees; (b) a grant of $12.80 for school supplies, paid directly to the girl's family; and (c) public recognition at a school awards assembly.

The competition for scholarships took place across a large number of schools and among a large number of students, making it less likely that the program would undermine cooperation between students within schools and classrooms. During the first year of the scholarship, roughly 57 percent of the 63 program schools had at least one winner, with an average of 5.6 winners in each of those schools. During the second year, 70 percent of program schools had at least one winner.

Through several unannounced attendance checks each school year, ICS personnel administered questionnaires to students in grades five through seven, collecting information on study effort, habits, and attitudes toward school and schoolwork. These surveys also confirmed that most students were aware of the scholarship and understood who was eligible to receive the award.

In both Busia and Teso, ICS invited selected schools to participate in the program through random draws, similar to a lottery. In Busia, all schools invited to participate did so and, hence, the characteristics of program schools and comparison schools were balanced. In Teso, some schools and individuals chose not to participate, and so the random allocation of invitations was not sufficient to ensure balanced program and comparison groups, making it harder to draw inferences about the impact of the program. We focus below on the effects in Busia.

RESULTS

The randomized selection of treatment and control schools made it relatively easy and straightforward to get a reliable measure of the scholarship program's impact. We simply needed to compare students' test scores and other educational outcomes across the two groups of schools before and after the introduction of the scholarship. We made these comparisons over the two-year span of the program, first comparing students who were in grade six in 2001, and then students who were in grade six a year later, in 2002.

The test-score impact of the scholarship program was large and tended to spill over to other students who had little or no chance of winning the award. During both years of the program in Busia, girls' test scores in scholarship schools improved markedly when compared to girls in control schools; test scores increased by 0.29 standard deviations among grade-six girls in 2001, and they increased by 0.21 standard deviations among grade-six girls in 2002. These improvements are roughly equivalent to 0.2 grades of extra primary schooling. As expected, girls scoring just below the winning threshold on baseline exams showed the largest test-score gains. But there were also large improvements among girls who scored poorly on baseline exams and who were therefore not likely to win the award from the outset.

Test scores among boys, who were definitely ineligible to compete for the award, also improved. Boys' test scores in Busia program schools increased over the two years by 0.13–0.21 standard deviations. As with the girls, boys at all levels of the original baseline distribution improved their scores, although the gains at the top of the distribution were somewhat more pronounced.

Finally, these test-score gains also appear to have extended beyond the time students were eligible to win the scholarship. Data collected in 2002 from the original cohort of girls, then in grade seven, suggest that these gains were lasting and were not due to extra preparation sessions or cheating on the exam in an effort to win the award.

One potential explanation for the broad improvement in test scores was the jump in student and teacher attendance rates in program schools. Student attendance increased by as much as 5 percentage points for both girls and boys in Busia, equivalent to reducing absenteeism by almost one-third. At the same time, teacher absentee rates dropped by about 6 percentage points, an effect roughly as large as the attendance gains among students. The improvements in attendance were evenly distributed across each school year, indicating that these gains were not due to extra study sessions just before the exams.

The large attendance gains in program schools among boys and girls with low baseline test scores immediately suggests that the rise in student effort (as measured by attendance) was not simply due to test preparation. Moreover, the higher teacher-attendance rates provide a plausible explanation for the positive spillover effects experienced by boys in Busia program schools: Any increase in teacher effort caused by the scholarship benefited the class as a whole.

Moreover, we found little evidence to support the common criticisms of merit scholarships. Students did not appear to have spent more time cramming for exams or otherwise focusing on them at the expense of other aspects of learning. Nor did we find evidence to support the argument that external rewards like merit scholarships interfere with a student's self-esteem or motivation to learn. According to our survey results, students' attitudes toward school and school work remained similar between program and comparison schools and between girls and boys.

On the other hand, we did find that scholarship winners came from somewhat more advantaged families than the other students in the sample. Parents of scholarship winners, for example, had nearly three more years of schooling than parents of nonwinners (7.7 years compared to 4.8 years). However, there was no notable difference between winners and nonwinners in terms of important household assets, such as iron roofs or latrines, and so there was no evidence that children from wealthier households were more likely to win.

Finally, when compared to other randomized interventions conducted in the same region, it appears that the scholarship program was more cost-effective than alternative projects that supplied textbooks, flip-charts, school uniforms, or offered performance-based incentives to teachers. In terms of effectiveness, the average test-score gain in merit scholarship program schools for female and male students in both Busia and Teso over the two years of the program was roughly 0.12 standard deviations; the comparable gain for schools participating in the teacher incentive program over two years was just 0.07 standard deviations. The average gain for the textbook-program schools was only 0.4 standard deviations, while the flip-chart and the child-sponsorship programs (which provided the school uniforms) did not produce any statistically significant effects on test scores.

In dollar terms, although the picture is a little less clear, the scholarship program was still the least expensive way to improve test scores. Using the average program impact for both Busia and Teso, the per-pupil cost of increasing test scores by 0.1 standard deviation was $1.41, the comparable cost of the teacher incentive program was $1.36, and the textbook program $5.61. However, if we limit our analysis to Busia, where the girls' scholarship program was well received, the per-pupil cost of the program was only $0.75, far lower than any other program aimed at improving test scores.

REFINING THE DEBATE

While merit-based scholarships have re-emerged in recent years in the United States and elsewhere, we still have little evidence on precisely how these programs affect students' learning. Critics have argued that such programs benefit only those students with a certain background, or that they improve education outcomes only over the short run, or that they cause students and teachers to concentrate on prepping for achievement tests rather than other important aspects of learning.

However, evidence from the randomized evaluation we conducted in Kenya suggests that such programs can produce an environment where teachers and students from all skill levels increase their effort, resulting in higher student academic achievement that

is long-lasting. Introducing the scholarship in Kenya, moreover, did not appear to have any significant negative impact on students' attitudes or desire to learn.

Although we found evidence that students from more advantaged backgrounds (in this case, those whose parents had more schooling) gained most from the program, there may be ways to spread those benefits more widely. In the United States, one way to achieve this might be to limit merit scholarships for tertiary education to students from poor, or poorly performing, areas where the more disadvantaged students would have a chance at winning an award.

Finally, while most research on education focuses on the impact of additional material resources on student performance, the results of our study suggest that student motivation is a critical variable.

ENDNOTE

[1] For more information on Kenya's program, see "Incentives to Learn" by Michael Kremer, Edward Miguel, and Rebecca Thornton, National Bureau of Economic Research Working Paper. No. 10971, December 2004. *www.nber.org/papers/w10971*.

Contributors

Clive R. Belfield is an assistant professor of economics at Queens College, City University of New York, and associate director for the National Center for the Study of Privatization in Education at Teachers College, Columbia University. His most recent book is *Privatizing Educational Choices in America's Schools* (Paradigm Publishers, 2005). Dr. Belfield has conducted economic analyses of preschooling programs for the states of New York, Massachusetts, Ohio, and Wisconsin. He is also the editor of a forthcoming collection of articles on the economics of early childhood education.

Christopher Berry is an assistant professor in the Harris School at the University of Chicago. His research focuses on the political economy of American local government, education policy, and economic development. Dr. Berry is currently engaged in several projects examining how the political organization of state and local government influences fiscal policy and economic performance. His other research has addressed the effect of test scores on school-board elections, zoning, and residential segregation, and the use of alternative financial service institutions by low-income households. Dr. Berry received a PhD from the Department of Political Science at the University of Chicago.

Eric Bettinger is an assistant professor of economics at the Weatherhead School of Management at Case Western Reserve University and a faculty research fellow in the Education Program at the National Bureau of Economic Research. His research focuses on how higher-education policies affect students' college success. He has studied the effectiveness of college remediation, the relationship between college instructor characteristics and student outcomes, and the effects of Pell grants on persistence. In recent years, Dr. Bettinger has studied educational voucher programs in both the United States and developing countries. He completed his PhD in economics at the Massachusetts Institute of Technology.

Christopher Cornwell is a professor of economics in the Terry College of Business at the University of Georgia. He has been a visiting professor at Universität Erlangen-Nürnberg in Nürnberg, Monash University in Melbourne, and Central European University in Budapest, as well as a visiting scholar at the Federal Reserve Bank of Cleveland. Dr. Cornwell's research interests include applied econometrics, labor economics, and the economics of education. He is currently engaged in a long-term project examining the economic effects of merit-based financial aid programs modeled after Georgia's HOPE Scholarship. He has a PhD in economics from Michigan State University.

Eric A. Hanushek is the Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University. He is also chairman of the Executive Committee for the Texas Schools Project at The University of Texas at Dallas, a research associate for the National Bureau of Economic Research, and a member of the Koret Task Force on K-12 Education. He is a leading expert on education policy, specializing in the economics and finance of schools. Dr. Hanushek's ongoing research includes the impacts of high-stakes accountability and of class-size reduction, and the importance of teacher quality. He completed his PhD in economics at the Massachusetts Institute of Technology.

Michael Kremer is a professor of economics at Harvard University and senior fellow at the Brookings Institution. Dr. Kremer is an expert on AIDS and infectious diseases in developing countries, economics of developing countries, education and development, and mechanisms for encouraging research and development. He serves as associate editor for the *Quarterly Journal of Economics* and the *Journal of Development Economics* and is the coauthor of several books. Dr. Kremer previously served as a teacher in Kenya. He founded and was the first executive director of WorldTeach, a nonprofit organization that places volunteer teachers in developing countries. He received his PhD from Harvard University.

Lance Lochner is an associate professor of economics and a research fellow at the CIBC Centre in Human Capital and Productivity at the University of Western Ontario. He serves as a research associate for the Federal Reserve Bank of Cleveland and a faculty research fellow in the Labor Studies' and the Children's Research programs of the National Bureau of Economic Research. Dr. Lochner spent five years in the Economics Department at the University of Rochester and one year as a national fellow at Stanford University's Hoover Institution. He received his PhD in economics from the University of Chicago.

Michael H. Moskow is the president and chief executive officer of the Federal Reserve Bank of Chicago. He is a trustee of Lafayette College, a member of the Advisory Board to the Kellogg School of Management at Northwestern University, and a member of the Visiting Committee of the Irving B. Harris Graduate School of Public Policy Studies at the University of Chicago. He is active in numerous professional and civic organizations, including serving as chairman of the National Bureau of Economic Research and as vice chairman of the Chicago Council on Foreign Relations. Dr. Moskow received a PhD in business and applied economics from the University of Pennsylvania.

Sandra Pianalto is the president and chief executive officer of the Federal Reserve Bank of Cleveland. Before joining the Bank, Ms. Pianalto was an economist at the Federal Reserve Board of Governors and served on the staff of the Budget Committee of the U.S. House of Representatives. She currently serves on several boards of directors including the Cleveland Foundation and the Northeast Ohio Council on Higher Education. Ms. Pianalto earned a master's degree in economics from the George Washington University. She is a graduate of the Advanced Management Program at Duke University's Fuqua School of Business and holds honorary doctor of humane letters degrees from The University of Akron, Baldwin-Wallace College, Kent State University, and Ursuline College.

Arthur Rolnick is senior vice president and director of research at the Federal Reserve Bank of Minneapolis and an associate economist with the Federal Open Market Committee. His research interests include banking and financial economics, monetary policy, monetary history, the economics of federalism, and the economics of education. He serves on several nonprofit boards, including the Minnesota Council on Economic Education and Greater Twin Cities United Way. He is also on the *Minneapolis Star Tribune*'s Board of Economists, and is a member of Minnesota's Council of Economic Advisors. Dr. Rolnick received a doctorate in economics from the University of Minnesota.

Todd R. Stinebrickner is an associate professor and graduate director in the Department of Economics at the University of Western Ontario. He is the cofounder and codirector of the Berea Panel Study, which was designed to provide more detailed information about college students than data available from other longitudinal data sources. His work in the area of higher education includes the examination of issues related to direct costs, liquidity constraints, time-use, and peer effects. He has also examined teacher decisions in elementary and secondary schools that have substantial numbers of low-income and minority students. Dr. Stinebrickner received a PhD in economics from the University of Virginia.

Participants

David Altig
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Michael Anderson
West Side Ecumenical Ministry
Cleveland, Ohio

Joshua Angrist
Massachusetts Institute of Technology
Cambridge, Massachusetts

Joanne Arhar
Kent State University
Kent, Ohio

Rich Arnold
Bloomberg News
Cleveland, Ohio

Mitchell Balk
The Mt. Sinai Health Care Foundation
Cleveland, Ohio

Kelly Banks
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Thomas Barrett
Hanna Perkins Center
Shaker Heights, Ohio

Patricia Barto
Cleveland Municipal School District
Cleveland, Ohio

Paul Bauer
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Clive R. Belfield
Queens College, City University of New York
Flushing, New York

Christopher Berry
University of Chicago
Chicago, Illinois

Eric Bettinger
Case Western Reserve University
Cleveland, Ohio

Kenneth Boris
Cuyahoga County Department of Children and Family Services
Cleveland, Ohio

Joyce Bresler
Starting Point
Cleveland, Ohio

Michael Bryan
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Heather Burton
The Center for Community Solutions
Cleveland, Ohio

Kevin Burtzlaff
Cleveland Municipal School District
Cleveland, Ohio

Louise Caldrone
Partnership for Education in Ashtabula County
Ashtabula, Ohio

John Carlson
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Fred Carr
The University of Akron
Akron, Ohio

Charles Carsltrom
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Robert Cheren
Case Western Reserve University
Cleveland, Ohio

Akia Churn
Girl Scouts of Lake Erie Council
Cleveland, Ohio

Robecca Cohen
The Center for Community Solutions
Cleveland, Ohio

Karina Colbenson
City Year Cleveland
Cleveland, Ohio

Christopher Cornwell
University of Georgia
Athens, Georgia

Ben Craig
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Richard Crepage
Ashland County Educational Service Center
Jefferson, Ohio

Christine Cross
Cuyahoga County Department of Children
and Family Services
Cleveland, Ohio

Cheryl Davis
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Erin Davis
Cuyahoga County Employment and Family Services
Cleveland, Ohio

Patricia DeMaioribus
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Paolo DeMaria
Ohio Department of Education
Columbus, Ohio

Bill Doll
Bill Doll and Company
Cleveland Heights, Ohio

Jillian Driscol
Family and Children First Council
Cleveland, Ohio

Mary Ellen Edwards
The University of Toledo
Toledo, Ohio

David Ellis
The Center for Community Solutions
Cleveland, Ohio

Clyde Evans
Ohio House of Representatives
Columbus, Ohio

Frank Fecser
Positive Education Program
Cleveland, Ohio

Maria Ferreyra
Carnegie Mellon University
Pittsburgh, Pennsylvania

Trisha Fomby
The YWCA of Greater Cleveland
Cleveland, Ohio

Mark Freeman
Shaker Heights City Schools
Shaker Heights, Ohio

Ruth Friedman
Beachwood Board of Education
Beachwood, Ohio

June Gates
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Marty Gelfand
Office of U.S. Representative Dennis Kucinich
Cleveland, Ohio

Amanda Gibson
Federal Reserve Bank of Richmond
Richmond, Virginia

Melissa Gibson
The YWCA of Greater Cleveland
East Cleveland, Ohio

Robyn Gibson
Children's Hunger Alliance
Cleveland, Ohio

Magda Gomez
Cuyahoga County Office of Early Childhood
Cleveland, Ohio

Colleen Grady
Ohio State Board of Education
Strongsville, Ohio

Mark Greenlee
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Emily Hagan
Voices for Children of Greater Cleveland
Cleveland, Ohio

Amy Hanauer
Policy Matters Ohio
Cleveland, Ohio

Eric A. Hanushek
Stanford University
Stanford, California

Barbara Henshaw
Federal Reserve Bank of Cleveland,
Cincinnati Branch Office
Cincinnati, Ohio

Karen Herpel
Serenity Research
Cleveland Heights, Ohio

Charles Hickman
Northeast Ohio Council on Higher Education
Fairlawn, Ohio

David Huskins
The University of Akron
Akron, Ohio

Robert Jaeckin
West Side Ecumenical Ministry
Cleveland, Ohio

Bernice Jefferis
Cleveland Heights–University Heights School Board
Cleveland Heights, Ohio

Jill Jefferis
Harvard Graduate School of Education
Cleveland Heights, Ohio

Connie Jones
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Brad Jupp
Denver Public Schools
Denver, Colorado

Michelle Katona
Office of Early Childhood, Invest in Children
Cleveland, Ohio

Maggie Keenan
Neighborhood Centers Association
Cleveland, Ohio

Katie Kelly
The Center for Community Solutions
Cleveland, Ohio

Ellen Kleinerman
The Plain Dealer
Cleveland, Ohio

Stephanie Klupinski
Catalyst Cleveland
Cleveland, Ohio

Kara Koenig
The Center for Community Solutions
Cleveland, Ohio

Mary Beth Koss
Achievement Centers for Children
Highland Hills, Ohio

Connie Krauss
Greater Akron Chamber
Akron, Ohio

Sandy Kreisman
University Circle, Inc.
Cleveland, Ohio

Michael Kremer
Harvard University
Cambridge, Massachusetts

Elizabeth Kucinich
Office of U.S. Representative Dennis Kucinich
Lakewood, Ohio

Barbara Kurtz
Early Childhood Options
Cleveland, Ohio

Laura Kyzour
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Helen Ladd
Duke University
Durham, North Carolina

Dennis LaRue
Youngstown Business Journal
Youngstown, Ohio

Wendy Leatherberry
The Center for Community Solutions
Cleveland, Ohio

Yoonsoo Lee
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Lance Lochner
The University of Western Ontario
London, Ontario, Canada

Joy Loube
Catalyst Cleveland
Cleveland, Ohio

Stuart Love
Meyer Hill Lynch Corporation
Maumee, Ohio

Ralph Lundberg
Neighborhood Centers Association
Cleveland, Ohio

Tammy Macek
Lorain County Community College
Elyria, Ohio

Steve Malone
Federal Reserve Bank of Richmond
Richmond, Virginia

Megan Marsick
Early Childhood Options
Cleveland, Ohio

Leaura Materassi
Cleveland Municipal School District
Cleveland, Ohio

Steve Matthews
Bloomberg News
Cleveland, Ohio

C. J. Matthews
Mount Sinai Baptist Church
Cleveland, Ohio

John McGovern
Cleveland State University
Cleveland, Ohio

Connie McHenry
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Mark McMillian
North Central Ohio Regional Training Center
Cleveland, Ohio

Sara Mierke
Foundation Management Services
Cleveland, Ohio

Willetta Milam
Cleveland Municipal School District
Cleveland, Ohio

Janet Miller
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Ann Millin
The George Gund Foundation
Cleveland, Ohio

Tom Mooney
Ohio Federation of Teachers
Columbus, Ohio

John Morris
Economics Center for Education and Research
Cincinnati, Ohio

Michael H. Moskow
Federal Reserve Bank of Chicago
Chicago, Illinois

Chuck Murr
Associated Press/Dow Jones
Cleveland, Ohio

Janet Narten
Foundation Management Services
Cleveland, Ohio

F. Howard Nelson
American Federation of Teachers
Washington, D.C.

Howard O'Cull
West Virginia School Boards Association
Charleston, West Virginia

Kim Palmer
Reuters News Service
Cleveland, Ohio

Elijah Palmer
The University of Toledo
Toledo, Ohio

Judith Peters
West Side Ecumenical Ministry
Cleveland, Ohio

Kriss Plumer
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Eric Price
Quality Dayton
Beavercreek, Ohio

Linda Prosak
Cleveland Muncipal School District
Cleveland, Ohio

Jennifer Ransom
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Allison Rand
Foundation Management Services
Cleveland, Ohio

Joseph Regano
Solon City Schools
Solon, Ohio

John Richard
Perry Local Schools
Massillon, Ohio

Diana Robbins
University Circle, Inc.
Cleveland, Ohio

Jim Rohn
Employment and Family Services
Cleveland, Ohio

Arthur Rolnick
Federal Reserve Bank of Minneapolis
Minneapolis, Minnesota

Brian Rudick
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Peter Rupert
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Erkin Sahinoz
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Zach Schiller
Policy Matters Ohio
Cleveland, Ohio

Mark Schweitzer
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Kim Sebaly
Kent State University
Kent, Ohio

Don Slocum
Neighborhood Leadership Institute
Cleveland, Ohio

Mark Sniderman
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Doris Stedmire
Mount Sinai Baptist Church
Cleveland, Ohio

Todd R. Stinebrickner
The University of Western Ontario
London, Ontario, Canada

Elaine Sutton
The Center for Community Solutions
Cleveland, Ohio

William Testa
Federal Reserve Bank of Chicago
Chicago, Illinois

Maria Thompson
Federal Reserve Bank of Cleveland
Cleveland, Ohio

James Thomson
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Piet Van Lier
Catalyst Cleveland
Cleveland, Ohio

Nancy Van Meter
American Federation of Teachers
Washington, D.C.

Jean VanNess
GAR Foundation
Akron, Ohio

Carol VanSickle
Toledo Area Chamber of Commerce
Toledo, Ohio

Paula Warren
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Stephanie Warsmith
Akron Beacon Journal
Akron, Ohio

Bradley Whitehead
The Cleveland Foundation
Cleveland, Ohio

Helen Williams
The Cleveland Foundation
Cleveland, Ohio

Alaina Wolfe
The Center for Community Solutions
Cleveland, Ohio

Charlyn Wood
CONNECT
Cleveland, Ohio

Alexander Wrege
The University of Toledo Foundation of Education
Toledo, Ohio

Shadya Yazback
Federal Reserve Bank of Cleveland
Cleveland, Ohio

Natasha Yovich
DSAS/Grandparent–Kinship Care Program
Cleveland, Ohio

Acknowledgments
Conference Program and Proceedings Volume

Conference Management Team
David Altig, Vice President and Associate Director of Research, Research; Kelly Banks, Assistant Vice President and Public Information Officer, Public Information; Michael Bryan, Vice President and Economist, Research; Cheryl Davis, Vice President, Office of the Corporate Secretary; Peter Rupert, Economic Advisor, Research; Mark Schweitzer, Assistant Vice President, Research; Mark Sniderman, Senior Vice President and Director of Research, Research.

Conference Program Coordination, Support, and Media Relations
Patricia DeMaioribus, Communications Coordinator, Public Information; June Gates, Media Relations Coordinator, Public Information; Virginia Hopley, Communications Coordinator, Public Information; Connie Jones, Administrative Assistant, Research; Laura Kyzour, Administrative Assistant, Public Information; Janet Miller, Research Assistant, Research; Kriss Plumer, Public Information Representative, Learning Center and Money Museum; Jennifer Ransom, Communications Coordinator, Learning Center and Money Museum; Brian Rudick, Research Assistant, Research; Erkin Sahinoz, Research Analyst, Research; Angie Tenney, Conference and Meeting Planner, Facilities.

Proceedings Volume Production and Web Support
Michael Galka, Manager, Communications Support; Liz Hanna, Web Designer, Web Services; Amy Koehnen, Editor and Designer, Communications Support; Michele Lachman, Editor, Communications Support; David Toth, Manager, Web Services.