Mr. Chairman and Members of the Committee:

My name is Caroline Hoxby. I am the Scott and Donya Bommer Professor of Economics at Stanford University and the Director of Economics of Education at the National Bureau of Economic Research, the nation's leading nonprofit economic research organization. I served for several years as a presidential appointee to the National Board for Education Sciences. Over my career, first at Harvard and recently at Stanford, I have conducted research on a wide array of topics in elementary, secondary, and higher education including class size, charter schools, college tuition, school finance, and bilingual education. There is a common theme in my research and the research of the many Ph.D.s I have trained: we attempt to answer questions in education by applying the most reliable, most advanced, most scientific methods to the best available data.

Thank you for the invitation to testify. It is an honor to address you, and I believe that today's topics are absolutely key to improving education in the United States.

The United States faces a very bleak future if we do not figure out how to quickly and continuously improve the education of our population. The American industries that are still growing, thriving, and exporting are the industries that are most dependent on educated workers. If our economy is to grow fast enough to help solve our fiscal crisis, we must have a smarter, more productive education sector, not one that is simply more costly.

If this sounds like an insurmountable challenge, it is only because Americans can point to so little educational improvement over the past four decades that we, as a nation, have begun to believe that very little improvement is possible. Contrast this with medicine or almost any other field of applied knowledge. If we were offered the choice between a medical procedure that relied on today's knowledge versus the knowledge of 1970, we would--all of us--choose today's. We would probably be ambivalent about today's schools versus the schools of 1970.

The difference between education and medicine is not that improvement is impossible in education but possible in medicine. It is not that all children are difficult to manage and all patients are easy to manage. The difference is that education has not, until recently, benefitted from rigorous, scientific research.

The Education Sciences Reform Act (ESRA) of 2002 greatly transformed education research, moving it much closer to the successful models used by the National Science Foundation and the National Institutes of Health. ESRA stated unequivocally that the Institute for Education Sciences (IES) should facilitate research that met high, scientific standards in order that it produce reliable results. This was the crucial statement. Until ESRA, much of the U.S. Department of Education's budget for research was wasted on studies that were widely recognized to be unreliable. Not only was taxpayer money wasted, but the Department unintentionally endorsed and promoted poor research methods by funding low-standard studies.

Prior to ESRA, there were two particularly acute problems with Department of Education-funded

studies. The first was that they often employed subjective measures of what schools did and what students achieved. If a study relies on subjective measures, a researcher's ideology often dictates what the data says. The second and more pervasive problem was that Department-funded studies often made bold causal claims despite the fact that they used methods that could not possibly support such claims. Claims of causation--such as "stricter teacher licensure rules raise student achievement"--were made when the study showed nothing more than a correlation. For instance, in my example, schools with a higher percentage of teachers who are licensed are schools that serve students who come from more advantaged backgrounds. These students tend to have higher achievement regardless of how their teachers are licensed. It turns out that the correlation between teacher licensure and achievement tells us literally nothing about the causal effect of teacher licensure on achievement. In short, prior to ESRA, Department of Education-funded research routinely provided misinformation to American families and schools.

I support the recommendations that the National Board for Education Sciences has already made regarding the reauthorization of ESRA. Those recommendations, however, are necessarily detail-oriented. In my remaining time, I wish to provide a "big picture" perspective on ESRA, IES, and--more broadly--the role of the federal government in education research.

I have three main points.
1. IES has greatly improved education research since the enactment of ESRA, but vigilance and continued improvements are needed. We cannot afford to relax standards now. Rather, even higher scientific standards should be the goal.
2. The federal government, universities, and philanthropic organizations should share the responsibility for supporting education research. This mixed model, somewhat peculiar to the U.S., is essentially the right model. Each entity plays an important and distinct role.
3. The data collection and research support functions of the U.S. Department of Education should be the functions on which people with diverse political views can agree. This is because no market functions better in the absence of information on which parents, students, and schools can make choices. Also, truly scientific research in education is probably our best hope for improving the skills of Americans quickly, with the expenditures we are already making.

Again, my first point is that IES has greatly improved education but that now is the time to further raise, not relax, the scientific standards that are the crucial contribution of ESRA. We are not yet in the situation where high, scientific standards are so ingrained in the education research community that IES can take its "foot off the gas." Since its creation, IES has consistently promoted scientific methods by favoring studies that employ experimental and quasi-experimental methods such as randomized controlled trials, randomization built into pilot programs, and regression discontinuity. These methods produce reliable results when used properly. That is why they are also used in fields such as medicine and social program evaluation. Vigilance is needed, however, because even the best experiment is not "dummy proof." IES should continue to raise the bar, insisting on even better training in issues like attrition and measurement that arise in experiments. Also, not all important questions can be answered with experimental or quasi-experimental methods, and IES therefore needs to develop greater expertise in other evaluation methods, methods that produce reliable results only when

they are applied by researchers who are very highly trained.

Expert review panels are the key means by which IES gains access to expert opinion, maintains high research standards, and improves its own staff's knowledge of the latest methods and research. The Department of Education's expert panels have improved greatly since the enactment of ESRA. They now contain a sufficient percentage of well-trained experts that the panel process can be said to, very often but not always, fund research that produces reliable results. While IES reviewing is not yet equal in quality to the NSF reviewing I have experienced, IES has made remarkable progress. The Institute is only able to convene top experts and attract high quality proposals because researchers believe that the Department turned the corner with ESRA and now promotes scientifically-grounded research. Top experts only participate in review processes in which they believe. Top researchers, who can devote themselves to issues other than education, only submit proposals to reviewers who are expert enough to judge proposals well. In other words, IES is currently in a virtuous cycle: higher scientific standards induce participation by more expert reviewers. This leads better researchers to submit higher quality proposals, and the cycle continues. Vigilance is necessary, however: the virtuous cycle can easily break down and become a vicious cycle in which poor standards lead to poor participation, at which point the review process attracts only poor proposals.

Another thing that IES is doing well but that requires vigilance is data collection. IES, through its National Center for Education Statistics, has been collecting survey data on students and schools for decades. These data tend to be well-respected--this is one function of the Department's research arm that was high quality prior to ESRA. However, top-notch education research has migrated away from survey data and towards detailed administrative data. About 75 percent of studies published by top applied journals now rely on administrative data--datasets based on schools recording what a student does, what teachers and policies and classrooms he encounters, and what outcomes he attains, both in the short-term (test scores) and long-term (college graduation, earnings, and so on). The reason that research is migrating from survey to administrative data is that modern scientific methods that produce reliable estimates often require the large scale and completeness of administrative data. While the U.S. continues to have some of the world's best survey data on education, our country has fallen far behind the frontier in administrative data on education. Currently, most northern European countries and some South American countries have substantially better administrative data than the U.S. This *matters* because top researchers are motivated just as much by the availability of data that allow them to write excellent studies as they are motivated by funding. Thus, researchers are increasingly drifting away from the analysis of U.S. education policies and toward the analysis of other countries' education policies. To be concrete, I could now write a study of English, Dutch, or Swedish school choice reforms using better data than are available to me in the U.S. IES is making valiant efforts, which I praise, to create and sponsor stronger administrative databases, but this is another area in which continued exertion is needed. Integrating states' data and data from its own agencies (like the National Student Loan system) is probably the cheapest and quickest way for IES to improve education research.

A final thing that IES has done well under ESRA is courageously contract for rigorous studies of

high profile programs and programs on which the federal government already spends substantial money. I would cite, as examples, the evaluation of the D.C. Opportunity Scholarship Program, the evaluation of the 21st Century Community Learning Centers, and the evaluations of Professional Development programs in mathematics and reading. It simply does not make sense for U.S. taxpayers to fund programs year after year in the absence of scientific evidence of their effects, and findings from such rigorous studies should play an important role in any debate about their future. You may have observed that I said these contracts were courageous. They were. When one conducts a study using strong, scientific methods, one cannot know how it will turn out. It is always possible that some constituency will be angered by the results, but--then--that is the entire point of doing research. If we could accurately choose education programs simply by knowing "in our hearts" that they were right, we would already have very successful schools.

There are a few areas in which IES has great intentions but is not having the effect for which it hopes. The Regional Education Laboratories and the What Works Clearinghouse are examples.

My second point is that support and responsibility for education research should be shared by the federal government, universities, and philanthropic organizations. In the U.S., we have a successful model in which each of these entities plays an important and distinct role. While I would never argue that our model is perfect, I am routinely struck by how well it functions when I am abroad and experience other countries' education research. A similar mixed model of support is used for medical research.

The federal government should play a few roles in education research. First, and most obviously, it should collect and make available accurate data on all aspects of education that can be measured: expenditures, revenues, achievement, personnel, curriculum, school policy, and so on. Because there are enormous economies of scale and scope in data collection and because cross-state comparisons are so important to research, it is important that the *federal* government and not just state governments collect data and make it available in a timely way.

Second, the federal government should publish descriptive reports on American education. The word *descriptive* is important because such reports are part of the government's duty to disseminate data, rather than a duty to do causal research. A report that describes where English Learners enroll is descriptive. This must be distinguished from research that attempts to test a causal hypothesis such as whether bilingual education raises English Leaners' achievement. The federal government is *not* in a good position to conduct causal research itself. This is because such research requires methods that need expert review, and the government cannot both convene the reviewers and be the entity that is reviewed. In the same way, we would not want an accused person to convene his own jury. A good review process requires independence.

Third, the federal government should contract for highly reliable evaluations of the education programs it supports. These evaluations cost only a small fraction of what is spent on the programs themselves. For this small expenditure, a good evaluation can save taxpayers vast amounts of money, either by providing the evidence that improves a partially-successful program

or by providing the evidence that gives Congress the grounds for abolishing an unsuccessful program. The federal government should be prepared to fund evaluations of its programs with little financial help from universities or philanthropies. This is because the goal of such evaluations is *not* to be innovative or to explore new questions. The goal is to produce clear answers to well-specified questions regarding established programs. The ideal evaluation should employ methods that are well-validated that the evaluation is boring in every way *except for the results*. Fortunately, in the U.S., we have active competition for such contracts among a good number of organizations: Mathematica, Abt, Rand, Westat, AIR, MDRC, and so on.

Fourth, the federal government should share in the support of (but not be the exclusive supporter of) research by university-based and similar scholars. These are the people who develop new methods, who ask questions that are still somewhat speculative, and who conduct "basic research" in education. I will return to this point.

Philanthropic institutions also play a vital role in education research. In some ways, their role parallels the federal role *except* that philanthropies should focus more on trial programs that are innovative and less on established programs funded by the government. This is because the government uses money that taxpayers are obliged to pay while philanthropic organizations use money that their donors freely give. If a philanthropy spends money on a speculative educational program that does not succeed, the consequences fall on its donors--people who are affluent enough to accept this risk in return for the prospect of developing exciting new programs that benefit society. Philanthropies can obtain reliable evaluations by contracting with the same organizations that contract with the federal government. And, like the government, philanthropies should share in supporting research by university-based and similar scholars.

Let me now turn to the role of university-based researchers. As I mentioned, university-based researchers are primarily responsible not only for developing new and more scientific methods of evaluation, but also testing them, validating them in an array of applications, and training people to use them. For instance, university researchers developed the cutting-edge methods to deal with attrition and non-compliance in randomized controlled trials. They also developed the quasi-experimental methods that are currently the workhorses of evaluation. In addition, university-based researchers are almost entirely responsible for conducting basic research--research that has no immediate policy relevance but that provides fundamental information on which policies should be ultimately based. For example, I study peer effects--how students' achievement is affected by the other students who share the classroom with them. This basic research is a fundamental that we need to evaluate policies like school choice that affect which students are in each school. Another good example of basic research is the recent spate of studies that show (a) that different teachers have very different effects on achievement and (b) that a teacher's effect is *not* related to her credentials. This basic research is a fundamental we need for thinking about teacher pay incentives, teacher training, teacher tenure, and policies that affect which teachers end up in which schools. Finally, university researchers should be primarily responsible for investigating educational programs that are speculative, still under development, or implemented on a purely trial basis. University researchers must also do the uncomfortable work of analyzing programs that are currently unpopular with the administration

and/or philanthropies. As an example of a purely trial program designed and investigated by university researchers, I would point to the recent study that shows that students are more likely to enroll in college if their family can automatically file the Free Application for Federal Student Aid when it files its taxes. As an example of unpopular research, I would point to studies of school choice from the 1990s. Researchers who worked on such topics did not win many friends in the education establishment, but we are now glad that the studies exist because they inform us about how to structure choice policies.

I have said that the federal government and philanthropies should share in the support of university-based education research. Why? If the government and philanthropies do not have "skin in the game", they will not attract university researchers to study the policies or develop the methods that are important to them (the government and philanthropies). They will not attract top experts to review the contract-based studies they support. They will not learn about cutting-edge research and cutting-edge methods in real time. It is the nature of cutting-edge work that you cannot learn about it just by reading an article after the fact. You need to interact with researchers--ask them questions, pose alternatives.

Universities themselves should also share in supporting education research. Why? If we want university-based researchers to invent better methods and conduct basic research, they need to be rewarded for these activities. No one is better at generating these rewards than universities themselves. This is because universities' constituents give them incentives to create knowledge that is *original* and *a public good*, as all basic research is.

By sharing in the support for educational research, the federal government, universities, and philanthropists also share in setting the research agenda. This is a good thing. Innovation never benefits from one entity having a monopoly on what questions are interesting.

My third and final point is brief. The data collection and research support functions of the U.S. Department of Education should be the functions on which people can most easily agree. Americans tend to disagree on the degree to which the federal government should mandate educational standards and impose policies on schools. Many Americans believe that families and local communities should make education choices for themselves. But, it is hard to argue that anyone--families, communities, schools, or federal policy makers--will make better choices if they have less access to reliable information. As I stated at the outset, Americans badly need to be better educated--and soon--because our economic growth and well-being depend on this. I truly believe that our best hope is to improve education by spending smarter--using scientific methods to identify which programs and policies are effective and which are counterproductive or just a waste of money.