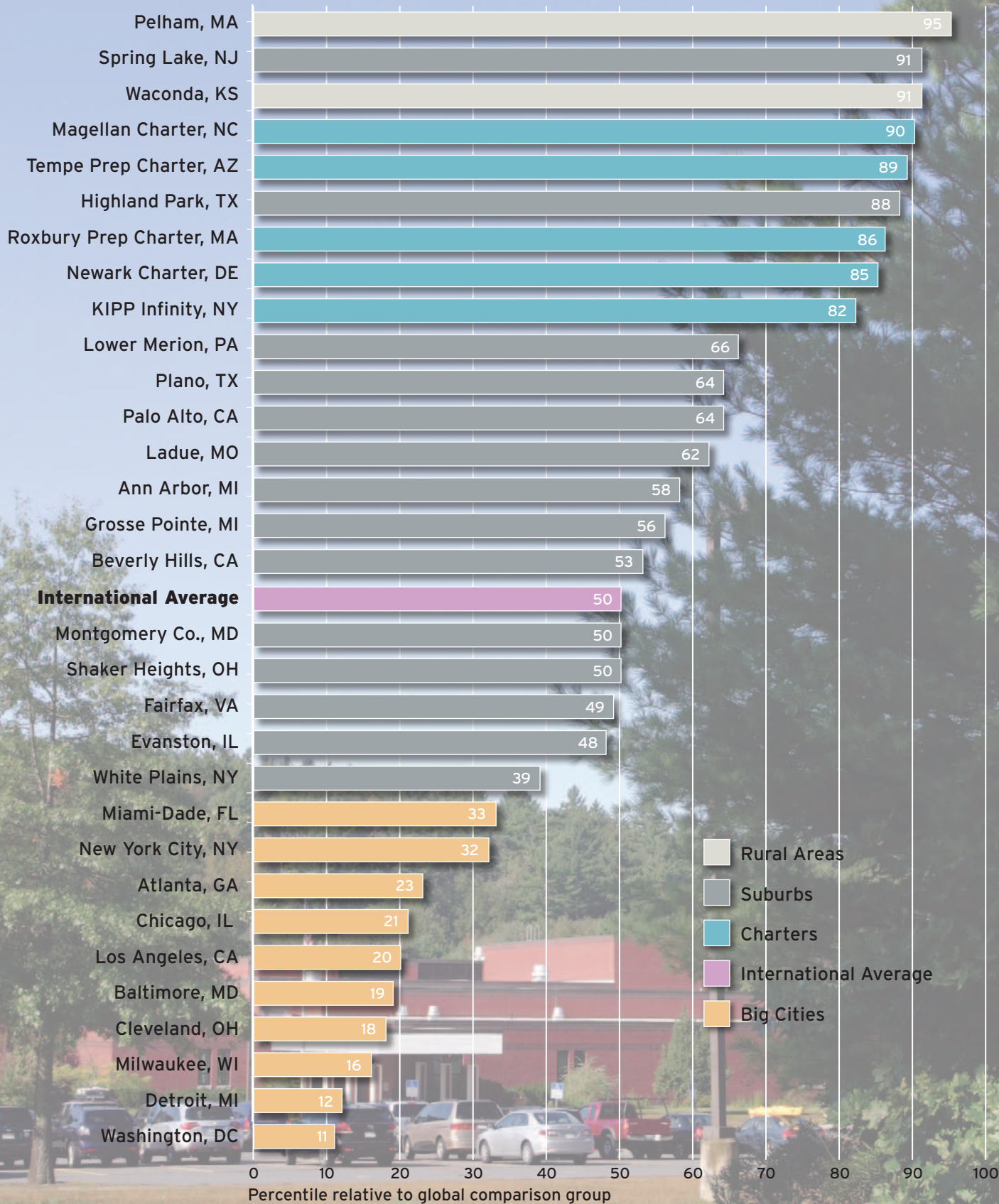# Barely Above Average (Figure 1)

Affluent communities in the United States often perform close to the international median (50th percentile), while only a few rural communities, small suburbs, and charter schools perform at the highest levels.

District	Percentile
Pelham, MA	95
Spring Lake, NJ	91
Waconda, KS	91
Magellan Charter, NC	90
Tempe Prep Charter, AZ	89
Highland Park, TX	88
Roxbury Prep Charter, MA	86
Newark Charter, DE	85
KIPP Infinity, NY	82
Lower Merion, PA	66
Plano, TX	64
Palo Alto, CA	64
Ladue, MO	62
Ann Arbor, MI	58
Grosse Pointe, MI	56
Beverly Hills, CA	53
International Average	50
Montgomery Co., MD	50
Shaker Heights, OH	50
Fairfax, VA	49
Evanston, IL	48
White Plains, NY	39
Miami-Dade, FL	33
New York City, NY	32
Atlanta, GA	23
Chicago, IL	21
Los Angeles, CA	20
Baltimore, MD	19
Cleveland, OH	18
Milwaukee, WI	16
Detroit, MI	12
Washington, DC	11

Percentile relative to global comparison group

Legend:
- Rural Areas
- Suburbs
- Charters
- International Average
- Big Cities

ILLUSTRATED CHART / BRUCE SANDERS DESIGN

Note: The bars show the percentile at which the average student in selected districts performed in math in 2007 relative to students in a set of 25 other developed countries. For example, the average student in Shaker Heights, Ohio, has math achievement that would place her at the 50th percentile among her peers in our comparison group.

SOURCE: Authors' calculations

When the Best Is Mediocre

American education has problems, almost everyone is willing to concede, but many think those problems are mostly concentrated in our large urban school districts. In the elite suburbs, where wealthy and politically influential people tend to live, the schools are assumed to be world-class.

Unfortunately, what everyone knows is wrong. Even the most elite suburban school districts often produce results that are mediocre when compared with those of our international peers. Our best school districts may look excellent alongside large urban districts, the comparison state accountability systems encourage, but that measure provides false comfort. America's elite suburban students are increasingly competing with students outside the United States for economic opportunities, and a meaningful assessment of student achievement requires a global, not a local, comparison.

We developed the Global Report Card (GRC) to facilitate such a comparison. The GRC enables users to compare academic achievement in math and reading between 2004 and 2007 for virtually every public school district in the United States with the average achievement in a set of 25 other countries with developed economies that might be considered our economic peers and sometime competitors. The main results are reported as percentiles of a distribution, which indicates how the average student in a district performs relative to students throughout the advanced industrialized world. A percentile of 60 means that the average student in a district is achieving better than 59.9 percent of the students in our global comparison group. (Readers can find all of the results of the Global Report Card at **www.globalreportcard.org.** The web site contains a full description of the method by which we calculated the results. For a summary, see the methodology sidebar.)

For the purposes of this article, we focus on the 2007 math results, although the GRC contains information for both math and reading between 2004 and 2007. We focus on 2007 because it is the most recent data set, and we focus on math because it is the subject that provides

Developed countries far outperform our most affluent suburbs

By JAY P. GREENE and JOSH B. McGEE

the best comparison across countries and is most closely correlated with economic growth. Readers should feel free to consult the GRC web site to find reading results as well as results for other years.

The Example of Beverly Hills

It is critically important to compare exclusive suburban districts against the performance of students in other developed countries, as these districts are generally thought to be high-performing. The most wealthy and politically powerful families have often sought refuge from the ills of our education system by moving to suburban school districts. Problems exist in large urban districts and in low-income rural areas, elites often concede, but they have convinced themselves that at least their own children are receiving an excellent education in their affluent suburban districts.

Methodology

The Global Report Card (GRC) builds on state accountability test results for the 13,636 school districts included in the American Institutes for Research (AIR) data set. The AIR data set is remarkably comprehensive inasmuch as the total number of school districts in the United States is estimated to be in the neighborhood of 14,000 districts. Given that AIR is a reputable research organization, we assume the data to be accurate.

Using the AIR data, we compute a student-weighted average across all grades of student performance on state accountability tests (under federal law, districts must test in grades 3-8, and once in high school). We place that average achievement in each district on a normal distribution of achievement relative to other districts in each state.

Then, using results from the U.S. Department of Education's National Assessment of Educational Progress (NAEP), we locate the center of each state's distribution of achievement in math and reading relative to the average performance in the United States. The districts within states with averages that trail the U.S. average are shifted down by the amount that their state lags the national average, and the opposite is done for districts in states with averages that exceed the national one.

An international test of math and reading performance administered by the Organisation for Economic Co-operation and Development (OECD), Programme for International Student Assessment (PISA), allows us to shift every district up or down relative to the results from the set of countries with developed economies. The results are expressed as a percentile, indicating where the average student in each district would be ranked in academic performance among the set of global peers. A percentile ranking of 60 means that the average student in a district performed better than 59.9 percent of students in the global comparison group.

To be included in this comparison group, countries had to have a 2007 per capita gross domestic product (GDP) of at least $24,000 and a population of at least 2 million, not be a member of OPEC, and have test results from PISA. Twenty-five countries met these criteria (see Table 1). Twenty-three countries had per-capita GDPs that significantly trailed the $45,597 of the United States. Some, such as Slovenia ($27,868) and Greece ($29,483), were roughly half as wealthy as the U.S. Only Norway ($53,968) and Singapore ($48,490) have higher per-capita wealth than the U.S. Overall, the countries with which we compare U.S. students are our major economic competitors. The performance of the comparison group was computed as the average of those 25 countries.

Although our estimates are the best available and provide good approximations of relative student performance across districts, states and countries, they are not exact. We are comparing the performance of students who took different tests, in different grades, and sometimes in different years. We have to assume that the results on all tests are normally distributed and that achievement can be compared by shifting those entire distributions up or down in sync with the over- or underperformance of each district relative to U.S. and global averages. But since test performance correlates highly across tests and standardized achievement levels of groups of students change only slightly from one grade to the next and one year to the next, the assumptions we make are not particularly restrictive. Any particular school district may have dramatically improved—or slid dramatically backward—over a short period of time, but those instances are likely to be exceptional, as overall U. S. performance has changed only slightly in recent years.

Unfortunately, student achievement in many affluent suburban districts is worse than parents may think, especially when compared with student achievement in other developed countries. Take for example Beverly Hills, California. The city has a median family income of $102,611 as of 2000, which places it among the top 100 wealthiest places in the United States with at least 1,000 households. The Beverly Hills population is 85.1 percent white, 7.1 percent Asian, and only 1.8 percent black and 4.6 percent Hispanic.

The city is virtually synonymous with luxury. A long-running television show featured the wealth and advantages of Beverly Hills high-school students (as well as their overly dramatic personal lives). If Beverly Hills is not the refuge from the ills of the education system that elite families are seeking, it's not clear what would be.

But when we look at the Global Report Card results for the Beverly Hills Unified School District, we don't see top-notch performance. The math achievement of the average student in Beverly Hills is at the 53rd percentile relative to our international comparison group. That is, one of our most elite districts produces students with math achievement that is no better than that of the typical student in the average developed country. If Beverly Hills were relocated to Canada, it would be at the 46th percentile in math achievement, a below-average district. If the city were in Singapore, the average student in Beverly Hills would only be at the 34th percentile in math performance.

Of course, people don't think of Beverly Hills as a school district with mediocre student achievement. This is partly because people assume that affluent suburbs must be high achieving and partly because state accountability results inflate achievement by comparing affluent suburban school districts with large urban ones. According to California's state accountability results, the average student in Beverly Hills is at the 76th percentile in math achievement relative to other students in the state. But outperforming students in Los Angeles, which is only at the 20th percentile in math relative to a global comparison group, should provide little comfort to Beverly Hills parents.

Los Angeles Unified is not the main source of competitors for Beverly Hills students, so the state accountability system encourages the wrong comparison. If Beverly Hills graduates are to have the kinds of jobs and lifestyles that their parents hope for them, they will have to compete with students from Canada, Singapore, and everywhere else. Beverly Hills students have to be toward the top of achievement globally if they expect to get top jobs and earn top incomes.

Fair Comparison (Table 1)

Nearly all of the comparison countries have per-capita GDPs that are lower than that of the United States.

Country	Population (000s)	GDP per Capita
Australia	20,750	$39,694
Austria	8,200	$38,303
Belgium	10,392	$35,953
Canada	32,936	$39,089
Denmark	5,468	$36,198
Finland	5,238	$33,912
France	63,682	$31,447
Germany	82,401	$33,181
Greece	10,706	$29,483
Hong Kong	6,980	$45,446
Ireland	4,109	$43,351
Israel	6,990	$25,302
Italy	58,148	$30,505
Japan	127,433	$32,063
Korea	48,250	$24,950
Netherlands	16,571	$36,394
New Zealand	4,132	$27,440
Norway	4,628	$53,968
Singapore	4,553	$48,490
Slovenia	2,009	$27,868
Spain	40,448	$33,616
Sweden	9,031	$35,271
Switzerland	7,555	$39,161
Taiwan	22,829	$27,884
United Kingdom	60,776	$34,320
United States	301,279	$45,597

SOURCE: University of Pennsylvania World Tables

Results from Affluent Suburbs Nationwide

We can repeat the story of Beverly Hills all across the country. Affluent suburban districts may be outperforming their large urban neighbors, but they fail to achieve near the top of international comparisons (see Figure 1). White Plains, New York, in suburban Westchester County, is only at the 39th percentile

in math relative to our global comparison group. Grosse Point, Michigan, outside of Detroit, is at the 56th percentile. Evanston, Illinois, the home of Northwestern University outside of Chicago, is at the 48th percentile in math. The average student in Montgomery County, Maryland, where many of the national government leaders send their children to school, is at the 50th percentile in math relative to students in other developed countries. The average student in Fairfax, Virginia, another suburban refuge for government leaders, is at the 49th percentile. Shaker Heights, Ohio, outside of Cleveland, is at the 50th percentile in math. The average student in Lower Merion, Pennsylvania, near Philadelphia, is at the 66th percentile. Ladue, Missouri, a wealthy suburb of St. Louis, is at the 62nd percentile. And the average student in Plano, Texas, near Dallas, is at the 64th percentile in math relative to our global comparison group.

All of these communities are among the wealthiest in the United States. All are overwhelmingly white in their population. All of them are thought of as refuges from the dysfunction of our public school system. But the sad reality is that in none of them is the average student in the upper third of math achievement relative to students in other developed countries. Most of them are barely keeping pace with the average student in other developed countries, despite the fact that the comparison is to *all* students in the other countries, some of which have a per-capita gross domestic product that is almost half that of the United States. In short, many of what we imagine as our best school districts are mediocre compared with the education systems serving students in other developed countries.

Pockets of Excellence

While many affluent suburban districts have lower achievement than we might expect, some districts are producing very high achievement even when compared with that of students in other developed countries. For example, the average student in the Pelham school district in Massachusetts is at the 95th percentile in math. That means that if we were to relocate Pelham to another developed country in our comparison group, the average student in Pelham would outperform 95 percent of the students in math. That's very impressive.

Of course, Pelham is a small district that is home to Amherst College, among other institutions of higher learning, and serves a rather select group of students. But not all college-town school districts are equally high achieving. As

we have already seen, Evanston, Illinois, is at the 48th percentile in math in a global comparison. Palo Alto, California, the home of Stanford University, is at the 64th percentile. And the average student in Ann Arbor, Michigan, home to the University of Michigan, is at the 58th percentile in math relative to students in other developed countries. So, the 95th percentile math achievement in Pelham is outstanding, even for college towns.

Spring Lake, New Jersey, has a similarly impressive record of having the average student at the 91st percentile in math. It is a very small and affluent community on the New Jersey shore that has somehow escaped the influence of Snooki and The Situation. Waconda, Kansas, a small rural community, also is at the 91st percentile. Highland Park, Texas, an affluent community near Dallas, is at the 88th percentile.

Interestingly, of the top 20 U.S. public-school districts in math achievement, 7 are charter schools (some states treat charter schools as separate public-school districts). And most of the 13 traditional districts remaining are in rural communities rather than in a large suburban "refuge" from urban education ills.

Pools of Failure

In total, only 820 of the 13,636 public-school districts for which we have 2007 math results had average student achievement that would be among the top third of student performance in other developed countries. That is, 94 percent of all U.S. school districts have average math achievement below the 67th percentile. There aren't that many truly excellent districts out there.

Of the 13,636 districts, 9,339, or 68 percent, have average student math achievement that is below the 50th percentile compared with that of the average student in other developed countries. Most of our large school districts are

Many of what we imagine to be our best public school districts are mediocre when compared with the education systems serving students in other developed countries.

well below the 50th percentile. This is especially alarming, because these lower-performing large districts comprise a much greater share of the total student population than do the relatively small higher-performing districts.

The average student in the Washington, D.C., school district is at the 11th percentile in math relative to students

Of the top 20 U.S. public-school districts in math achievement, 7 are charter schools (some states treat charter schools as separate districts). Most of the 13 districts remaining are in rural communities.

in other developed countries. In Detroit, the average student is at the 12th percentile. In Milwaukee, the average student is at the 16th percentile. Cleveland is at the 18th percentile. The average student in Baltimore is at the 19th percentile in math relative to students in other developed countries. In Los Angeles, the average student is at the 20th percentile. The average student in Chicago is at the 21st percentile in math. Atlanta is at the 23rd percentile. The average student in New York City is at the 32nd percentile in math. And in Miami-Dade County, the average student is at the 33rd percentile in math.

Not 1 of the largest 20 school districts is above the 50th percentile in math relative to other developed countries. Those districts contain almost 5.2 million students or more than 10 percent of the country's schoolchildren. The rare and small pockets of excellence in charter schools and rural communities are overwhelmed by large pools of failure.

Previous Research

The Global Report Card is not the first analysis to compare the performance of U.S. students to international peers. Eric A. Hanushek, Paul E. Peterson, and Ludger Woessmann (see "Teaching Math to the Talented," *features*, Winter 2011) used a very similar method to compare the performance of students in each state to students in other countries and arrived at similarly gloomy conclusions. Using state NAEP results for 8th-grade students and PISA results for 15-year-olds internationally, the researchers focused on the percentage of students performing at an advanced level in math. In almost every state, they found that we had far fewer advanced students than most of the countries taking PISA. They also narrowed the comparison to white students in the U.S. and to students whose parents had a college education to show that even advantaged students in the U.S. failed to achieve at an advanced level in math relative to their international peers. More recently, Hanushek et al. updated their analysis to examine the percentage of students in each state and across countries performing at the proficient level in math and reading. The results were similarly disappointing.

The main difference between the GRC and the Hanushek et al. analyses is that in our study we push the comparison down to the district level. By focusing on white students and children of college-educated parents, Hanushek et al. clearly mean to convey that even students in elite suburban districts have mediocre achievement. Our contribution with the GRC is to name the districts so that people do not indulge the fantasy that their suburb's record is somehow different from the disappointing performance of others with advantaged students in their state.

There are other important differences between the GRC and the Hanushek et al. analyses. We incorporate test results for U.S. students in all available grades (typically grades 3 through 8 and grade 10) rather than focusing on the grade closest to the 15-year-olds in the PISA sample. We could have focused only on 8th-grade results, as Hanushek et al. did, but in doing so we would have greatly reduced the number of test results on which we were doing the calculations for school districts. We preferred to gain precision in estimating the achievement in each district by increasing our sample size rather than restricting the sample to 8th graders in order to gain comparability in the age of the students under review.

The GRC analysis also differs from those of Hanushek et al. in that the latter focus on students performing at the advanced or proficient level, while we focused on the average student performance in both math and reading. Hanushek et al. concentrated on advanced or proficient performance because they were trying to compare our best students with the best abroad to show that even our best are mediocre. We did the same by highlighting the results for elite suburban school districts. Focusing on the average also avoids any dispute about how "advanced" or "proficient" are defined across different tests.

Gary Phillips at the American Institutes for Research has also conducted a series of analyses comparing state achievement on NAEP to international performance on a different international test, the Trends in International Mathematics and Science Study (TIMSS). Phillips arrives at somewhat less gloomy conclusions about U.S. performance, but that is because the countries included in TIMSS differ from those covered by PISA. Hanushek et al. rightly note that PISA provides a much more appropriate comparison for the U.S.: "Put starkly, if one drops from a survey countries such as Canada, Denmark, Finland, France, Germany, and New Zealand, and includes instead such countries as Botswana, Ghana, Iran, and Lebanon, the average international performance will drop, and the United States will look better relative to the countries with which it is being compared."

Apart from charter school districts, in over half of the states there are no more than three districts in which the average achievement would be in the upper third.

This has sparked a debate among researchers about whether TIMSS or PISA provides a better set of countries against which we should compare the U.S. The Global Report Card circumvents this dispute by developing its own set of countries against which we compare U.S. students. The comparisons provided by TIMSS and PISA depend on which countries decide to take each test each time it is administered. And PISA scales its scores against the results for members of the OECD, which excludes countries like Singapore while including countries like Mexico. Our comparison group depends on PISA results, but it is also based on objective criteria, like per-capita GDP, to identify a set of developed economies that can be reasonably compared with that of the U.S. Our comparison group is a significant improvement on the self-selection of countries that choose to take a test as well as an improvement upon arbitrary membership in an organization like the OECD.

No Refuge

The elites, the wealthy families that have a disproportionate influence on politics, clearly recognize the dysfunction of large urban school districts and have sought refuge in affluent suburban districts for their own children. But the reality is that there are relatively few pockets of excellence to which these families can flee.

In four states, there is not a single traditional district with average student achievement above the 50th percentile in math. In 17 states, there is not a single traditional district with average achievement in the upper third relative to our global comparison group. And apart from charter school districts, in over half of the states, there are no more than three traditional districts in which the average achievement would be in the upper third.

The elites in those states have almost nowhere to find an excellent public education for their children. But state accountability systems and the desire to rationalize the lack of quality options have encouraged the elites to compare their affluent suburban districts to the large urban ones in their state. These inappropriate comparisons have falsely reassured them that their own school districts are doing well.

This false reassurance has also perhaps undermined the desire among the elites to engage in dramatic education reform. As long as the elites hold onto the belief that their own school districts are excellent, they have little desire to push for the kind of significant systemic reforms that might improve their districts as well as the large urban districts. They may wish the urban districts well and hope matters improve, but their taste for bold reform is limited by a false contentment with their own situation.

But the elites should not take comfort from the stronger performance of affluent suburban districts relative to large urban districts. As the Global Report Card reveals, even our best public-school districts are mediocre when compared with the achievement of students in a set of countries with developed economies.

Of course, the Global Report Card does not isolate the extent to which schools add or detract from student performance. Factors from student backgrounds, including their parents, communities, and individual characteristics, have a strong influence on achievement. But the GRC does tell us about the end result for student achievement of all of these factors, schools included. And that end result, even in our best districts, is generally disappointing.

Jay P. Greene is professor of education reform at the University of Arkansas and a fellow at the George W. Bush Institute. Josh B. McGee is vice president for public accountability initiatives at the Laura and John Arnold Foundation.