Getting paid

One would think that economist Michael Podgursky ("Fringe Benefits," Check the Facts, Summer 2003) would analyze teachers' salaries through the lens of supply and demand. Such an analysis would not examine teachers' salaries as they are, but would ask what salaries are necessary to attract highly qualified teachers to the field. Instead, Podgursky insists that teaching is a swell job, using data showing that elementary teachers actually earn a few pennies per hour more than mechanical engineers (even though their salaries fall short of salaries in mechanical engineering by $17,000 per year). What a surprise this must be to students aspiring to engineering degrees!

Podgursky cites "straight-time" hourly pay estimates from the Bureau of Labor Statistics (BLS), which show K–12 teachers having a 38-hour workweek and a 37-week work year (which would end in April). Another BLS survey describes the hours worked by teachers more accurately. The average teacher was under contract to work six and a half hours per day, but teachers actually spent an average of eight and a quarter hours at school each day. The standard time diary methodology suggests that a teacher's working day is almost ten hours long when work outside school hours is factored in.

It's true that teachers work about 40 fewer days per year than private-sector employees. But even if teachers were able to fill the gap with additional work at $30 per hour—an unreasonable expectation for teachers seeking part-time employment—their annual earnings would increase by only $9,600. This still leaves teachers $3,000 per year short of accountants, $17,000 short of computer systems analysts, and $25,000 short of engineers. Teachers also earned only 8 percent more than the average worker in 2001 and 4 percent more than other government workers.

The AFT uses data from the National Association of Colleges and Employers (NACE), which contain information on about 30,000 job offers, not 2,600 as Podgursky reported, to update a 30-year time series for the earnings of new college graduates who found full-time jobs in the private sector. We agree with Podgursky that NACE salary data are higher than the average earnings of new college graduates, many of whom work part time, attend graduate school, or are underemployed.

Podgursky assumes that teachers enjoy generous benefits, but according to the BLS, benefit costs for teachers made up 24 percent of the total compensation package in December 2002—the lowest percentage of any broad occupational category. Benefits compose 28 percent of the average civilian worker's compensation package.

Furthermore, benefit costs for teachers have risen more slowly than the average, not faster, as Podgursky insinuated. Between 1989 and 2001, the benefits portion of the U.S. Department of Labor's employment cost index increased 48 percent for K–12 education workers, compared with 67 percent for all private-sector white-collar employees.

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

Michael Podgursky's analysis of teacher compensation makes thoughtful use of the sometimes incomplete and conflicting data that have been available to us. Regarding the issue of the U.S. Department of Education's reliance on data from the teacher unions on annual teacher salaries, the National Center for Education Statistics (NCES) plans to add teacher salary data items to its Common Core of Data finance collections. Collection of school year 2003–04 finance data would take place during the spring and summer of 2005, and NCES would publish the data in 2006.

It would also be desirable to have detailed information comparing the total compensation packages of teachers with those of other professions. While NCES currently gathers information on salaries and benefits through the quadrennial Schools and Staffing Survey, this survey does not collect information that would permit comparisons with other professions. There is no logical way to redesign the survey to do this. The Bureau of Labor Statistics has a large-scale payroll survey that gathers detailed information by profession, but this survey lacks information on the characteristics of employees, such as years of experience or educational qualifications. This survey also would not lend itself to a major redesign. The monthly Current Population Survey conducted by the Bureau of the Census offers considerable potential for making salary comparisons of
various occupations based on age and educational attainment. The sample size is limited for a number of occupations, including teachers, but NCES staff are working on ways to effectively increase the sample through more complex analysis of multiple waves.

Val Plisko
Associate Commissioner
National Center for Education Statistics
U.S. Department of Education

Michael Podgursky responds: Collective-bargaining agreements negotiated by the AFT in large urban districts typically include language restricting the contractual teaching workday to little longer than the school day for students. Thus the contractual workday in urban districts such as Chicago, Los Angeles, New York City, or San Diego is roughly 6 hours and 45 minutes. (This includes a duty-free lunch as well as a prep period.) The AFT claims that the average public school teacher actually spends 8 hours and 15 minutes in school daily. There are several possible explanations for this discrepancy. First, the self-reported data from teachers may be inflated. Or the self-reported data are correct and the contractual workday is routinely exceeded by teachers. This raises the question of why the AFT and NEA would insist on bargaining unprofessional language into contracts that most of their members then ignore. Finally, it may be that pay gaps between urban and suburban teachers in part reflect an hours gap, with suburban (and rural) teachers putting in longer workdays than their urban counterparts. These are interesting topics for further research.

In my article I avoided comparing the overall fringe benefit rate of teachers to nonteachers for the reason Howard Nelson notes, and then ignores. The overall fringe benefit rate for private-sector professionals on 12-month contracts includes paid vacations. However, in lieu of paid vacations, teachers on 190-day contracts get their summers off. A correct apples-to-apples comparison would either exclude paid vacations for the former or treat teachers as 12-month employees with 66 days of paid vacation. If we take the latter approach, this adds an additional 26 percent to the fringe benefit rate for teachers, and would thus produce a total fringe benefit rate that is far in excess of the private sector.

Comparing the pay and benefits of teachers and nonteachers is complicated and highlights the need for independent, arms-length assessment and high-quality data. The Bureau of Labor Statistics (BLS) performs this valuable function in other industries. Unfortu-
nately, in the area of teacher compensation, the national policy debate is largely framed by data and analysis from the teacher unions. The public interest would be better served if the U.S. Department of Education worked with the BLS to produce objective data on this important topic.

Second career

As a former aerospace engineer in my eighth year of a second career as a public high-school physics teacher, I find that Peter Temin and Richard Vedder pretty much get it right in their discussion of teachers’ compensation (“Are Teachers Underpaid?” Forum, Summer 2003). I offer the following comments.

Incompetence is not the major problem. In a labor-intensive business enterprise, one is always trying to improve the average level of staff ability by replacing the merely adequate with the better than adequate. Failure to do so is a double whammy: You don’t get the services of the very able; worse, your competitor does. I find few of my colleagues outright incompetent, but there are more than a few who wouldn’t last ten seconds in a nontenured work environment.

In public schools, managers are never disciplined for the way they treat staff. This is primarily because seniority rules strongly discourage teachers’ changing employers, which means they can’t use the threat of resigning to curtail management abuse. I have seen very able colleagues treated with a degree of disrespect that in industry would have resulted in the keys being thrown on the boss’s desk.

There is no distinction between adequate, good, and outstanding teachers. Having unique professional experience or winning grants for your school counts for nothing in determining pay or layoffs. It’s all strict seniority.

When I left engineering, sick leave was typically five days a year, and new employees got one week’s vacation in the first year. In my school district, everyone starts out with 15 days’ sick leave, 16 days’ paid vacation, and 11 paid holidays during the school year.

Most of my colleagues care deeply about their work and their charges, but whenever the union’s interest in minutes, money, and benefits collides with students’ interests, the students lose. As a union officer once shouted at me in a moment of exasperation, “The kids aren’t in the contract!”

The biggest disincentive to stay in this profession is not lack of money, but the relentless assault on one’s self-esteem by the corrosive stew of union avarice, management fecklessness, and school committee politics. Apart from that, I love it.

CLARK NEILY
Allston, Massachusetts
Does accountability work?

Margaret Raymond and Eric Hanushek harshly criticize (see “High-Stakes Research,” Feature, Summer 2003) our study of high-stakes testing policies. Before reporting the results from our study, the New York Times journalist obtained feedback from our study’s external reviewers as well as from scholars and advocates who support high-stakes testing. Raymond and Hanushek ask the media, “Why not bring in some outside expertise to review such a report before heralding its arrival?” Actually, the media did.

Our study analyzed data across multiple indicators of academic achievement, not simply the National Assessment of Educational Progress (NAEP). Yet Raymond and Hanushek’s review looked only at the results from our analysis of NAEP, ignoring the effects that high-school graduation exams have had on college-admissions tests like the SAT and on participation and performance in Advanced Placement courses. They also ignored the fact that high-school graduation exams have resulted in increased dropout rates and an increasing use of the General Educational Development, or GED, tests as a substitute for a high-school diploma. Do these consistently negative effects matter when assessing high-stakes testing? We think so.

Raymond and Hanushek discard our findings on the basis that our methods were flawed. All of our findings were derived using one of the strongest designs in empirical research—the archival time-series analysis, a method that some claim is second in quality only to a true controlled experiment. An archival time-series analysis is simple enough that readers do not need a background in statistics to understand the underlying logic. Readers need not get caught up in more-complicated analyses, such as significance testing, effect sizes, and even regression—statistical methods that Raymond and Hanushek criticize us for not using. However, many statistical textbooks recommend against using complicated statistical methods with archival time-series analyses.

Raymond and Hanushek throw the bias card into their critique, writing, “When a report is commissioned by an organization like the Great Lakes Center for Education Research and Practice, a Midwestern group sponsored by six state affiliates of the National Education Association, it would seem to call for a reasonable dose of skepticism.” Not mentioned by Raymond and Hanushek is the fact that the research was originally funded by the Rockefeller Foundation and was published in a peer-reviewed scholarly journal six months before the consortium of teacher unions released this version of the study. The fact that teacher unions backed the study had no impact on its conclusions.

Raymond and Hanushek claim that the “accumulated literature” supports the conclusion that “student performance on the available measures, usually state tests, improves after accountability reforms are introduced.” We believe that is patently false. We conducted a thorough review of the literature on high-stakes testing and found very few articles that would support such a proposition.

AUDREY AMREIN, DAVID BERLINER
Arizona State University
Tempe, Arizona

Margaret Raymond and Eric Hanushek respond: The assertion that “archival time-series analysis” is second in quality only to a true controlled experiment is ludicrous. Long ago, in their classic discussion of research design, Donald Campbell and Julian Stanley said that the time-series design “rarely has accepted status in the enumerations of available experimental designs in the social sciences.” The obvious inability of simplistic historical approaches to establish “experimental isolation”—to rule out other factors that might have influenced the observed outcomes—opens up results from such analyses to significant interpretative questions.

Another problem with Amrein and Berliner’s study is that they did not define an adequate comparison group. Instead, they compared student-performance trends in (some of) the states that adopted high-stakes testing with the average gain among states participating in NAEP—a trend that partially reflects the gains among high-stakes states, thereby corrupting the analysis. Amazingly, they make no attempt to defend this faulty approach. Instead, they trumpet the fact that they reached similar conclusions when they applied Amrein and Berliner’s own time-series methodology to the data (with an appropriate comparison group of
states that have not adopted high-stakes testing), their conclusions were completely reversed. Yet Amrein and Berliner don’t even address this. Their response ignores the egregious errors in implementation that we identified, namely the fact that they threw out a majority of the state observations, mis-coded outcome information, and completely confused the sequence of test introduction and achievement measurement in several states.

We know of no legitimate statistical text that argues it is irrelevant to use tests of statistical significance to guard against random fluctuations in the data—in this case, scores on tests of student performance. Each administration of the NAEP involves a different group of students, a different set of test questions, and a different testing environment. Across test administrations, these differences can lead to random changes in scores that bear little relation to actual changes in students’ knowledge and skills. The purpose of tests of statistical significance is to determine whether results reflect genuine changes in performance or simply random fluctuation.

That four of Amrein and Berliner’s colleagues from education schools approved of their report says more about the quality of the standards for research at too many schools of education than about the validity of this particular study. The disregard for standard scientific principles reveals why so little has been learned about effective educational practices.

Too soon to tell

In “Locked Down” (Feature, Summer 2003), Ronald Brownstein questions the efficacy of the No Child Left Behind Act’s school choice provisions. For one thing, it is a little too soon to draw conclusions based on anecdotal data in the first year of implementation. After all, one of the most chronic problems with school reform is the lack of patience with bold, innovative reforms that take at least a few years to bear fruit. Brownstein ignores the good faith effort that officials in cities like New York put forth once it was made clear that they did not convey the availability of options to parents in a timely fashion. In fact, most states and districts have been open to reforming their systems to meet the requirements of the law once they were notified of problems. Brownstein also fails to point out that districts that repeatedly circumvent the letter of the law will eventually come under more rigorous state sanctions and possibly federal sanctions.

NINA SHOKRAII REES
Deputy Under Secretary
Office of Innovation and Improvement
U.S. Department of Education

Vox pupils

The Koret Task Force on K–12 education (“Are We Still at Risk?” Forum, Spring 2003) left a crucial group out of its assessment of the education system: students. No students were on the task force, and none was even involved in determining their policy recommendations.

Koret is not alone. The Education Commission of the States requires prospective commissioners to “reflect broadly the interests of the member government, higher education, the state education system, local education, lay and professional persons, and public and nonpublic educational leadership.” The commissioners represent the interests of virtually every education-related group except students.

Before we can determine how best to reform education, we must answer a crucial question: Why are young people not learning? Ultimately, teachers, parents, scholars, and lawmakers can answer this question only to a limited degree. They are not the ones rocketing through adolescence while reading Salinger and studying cellular mitosis.

Current students and recent graduates from junior high and high school understand best what obstacles they and their peers encounter. They are able to reflect on their experiences, identify what works and why, and pinpoint what can be improved. Students, for example, may be indifferent to using technology as a teaching tool, but may insist on the efficacy of smaller classes. Students could answer important questions—but no one is asking them.

GANESH SITARAMAN
Cambridge, Massachusetts

Central planning

E.D. Hirsch contends that certain “nationalized, bureaucratic, non-market education systems” such as Japan’s develop higher-order skills not by directly teaching such skills but by paying close attention to the “sequence and coherence of content” (see “Not So Grand a Strategy,” Feature, Spring 2003). The policy implication, Hirsch writes, is that the United States needs a “coherent, specified grade-by-grade elementary curriculum,” not the “local control of curriculum and leaving a hundred flowers bloom.”

Hirsch underestimates the deformities inherent in highly centralized education systems. Centrally planned systems do not encourage ideas from the grassroots, thereby ignoring the nature of knowledge and discovery. Market-like mechanisms, not central planners, are the best way to direct resources toward success and away from failure. They are the only way to nurture sustainable innovation—as opposed to K–12 education’s endless fads.

TOM SHUFORD
Ventura, California

CORRECTION: In the second paragraph of “Philosopher or King?” (Richard Kahlenberg, Feature, Summer 2003), the phrase “of all things” was inadvertently introduced.