Journal of School Choice: International Research and Reform
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/wjsc20

Experiments to Generate New Data About School Choice: Commentary on "Defining Continuous Improvement and Cost Minimization Possibilities Through School Choice Experiments" and Merrifield's Reply

Nathan Berg [a] & John Merrifield [b]
[a] University of Texas at Dallas , Dallas, Texas, USA
[b] University of Texas at San Antonio , San Antonio, Texas, USA
Published online: 14 Dec 2009.

To cite this article: Nathan Berg & John Merrifield (2009) Experiments to Generate New Data About School Choice: Commentary on "Defining Continuous Improvement and Cost Minimization Possibilities Through School Choice Experiments" and Merrifield's Reply, Journal of School Choice: International Research and Reform, 3:4, 397-413, DOI: 10.1080/15582150903425822

To link to this article: http://dx.doi.org/10.1080/15582150903425822

Journal of School Choice, 3:397–413, 2009

ISSN: 1558-2159 print/1558-2167 online
DOI: 10.1080/15582150903425822

Routledge
Taylor & Francis Group

COMMENTARY

Experiments to Generate New Data About School Choice: Commentary on "Defining Continuous Improvement and Cost Minimization Possibilities Through School Choice Experiments" and Merrifield's Reply

NATHAN BERG
University of Texas at Dallas, Dallas, Texas, USA

JOHN MERRIFIELD
University of Texas at San Antonio, San Antonio, Texas, USA

Benefiting from new data provided by experimental economists, behavioral economics is now moving beyond empirical tests of standard behavioral assumptions to the problem of designing improved institutions that are tuned to fit real-world behavior. It is therefore worthwhile to consider the potential for new experiments to advance school choice debates and the daunting task of institutional reform. Deciding which new institutions to introduce and test is an important scientific question. New institutional alternatives are, however, by definition, untested, and empirical arguments to guide this process are difficult to construct without adding more layers of theoretical assumptions. This article analyzes normative criteria that can be used when introducing new institutions for delivering school services. Because of the long timeframes for testing the effects of new institutions and the extraordinarily high stakes for children and families who participate in such experiments, performance metrics that are normally uncontroversial, such as cost minimization, will likely be insufficient for justifying the introduction of new institutions.

Address correspondence to Nathan Berg, PhD, Associate Professor of Economics, School of Economic, Political, and Policy Sciences, University of Texas at Dallas, 800 W. Campbell Road, GR31, Richardson, TX 75083, USA. E-mail: nberg@utdallas.edu

KEYWORDS experiment, behavioral economics, inequality, laissez faire, equilibrium, transition

EXPERIMENTS AND SCHOOL CHOICE DATA

Behavioral/experimental economics is poised to enter a new phase in its relatively brief intellectual history, moving beyond empirical tests of standard behavioral assumptions in the social sciences to the problem of designing improved institutions that are tuned to fit real-world behavior. For this reason and others, it is worthwhile to analyze the potential value that new experiments could provide to the school choice debates by generating new data to settle long-standing questions.

Insofar as behavioral/experimental economics succeeds in delivering results that lead to improved real-world institutions, these improvements will likely rest on two key methodological innovations. First, instead of using standard assumptions about consumers and suppliers responding continuously and optimally to small changes in the costs and benefits they face, simpler behavioral models can produce improved predictions about how non-omniscient individuals with real-world limitations will respond to changes in the institutional environment—for example, following the default, imitating how others within close physical proximity are behaving, or satisficing by aiming simply to do at least as well as a threshold-level of performance, resulting in large regions of the cost/benefit space where there is no behavioral response at all to small changes in costs. The second reason why behavioral/experimental economics can make new contributions to institutional design is its methodological embrace of experimentation and the creation of new empirical data that illuminate substantive problems in economics.

From the standpoint of anyone trying to make our analyses of debates over school choice more of an empirical science and more relevant to the real world, John Merrifield's (2009) call for experiments in the field of school choice is very much welcome. His call for more experiments is based on the widely shared premise that existing data about institutional change provide, at best, a grossly insufficient empirical record on which to base high stakes decisions about the reform of educational institutions. Merrifield's proposals for new forms of data collection are all the more welcome given that he is not an experimental economist. His argument for how we benefit by producing a much stronger empirical record should be widely appealing to researchers working in a number of subfields in the social sciences.

The substance of his argument can be directly translated into the methodological language of experimental economics. I attempt to make this translation here. Hopefully, my backward translation—from Merrifield's arguments exposited in natural language into the jargon of experimental economics—provides additional illumination regarding which approaches

deserve to have the widest support. Additionally, I hope that it provides an evaluative framework for analyzing other claims about which institutional reforms should take priority.

Merrifield proposes mostly real-world experiments—what experimental economists call "field experiments"—to distinguish them from "laboratory experiments." Thus, Merrifield's proposal calls for data collection on the schools that real families actually choose, their educational expenditures, and changes in supply conditions. As far as supply conditions go, perhaps the most important outcomes to be studied are changes in the menu, or choice set, of schooling options available after laws have been changed to encourage new organizations (mostly private companies, one presumes) to offer new kinds of educational services.

Restated in the jargon of experimental economics, the essence of the argument is that the designers of existing field experiments that aimed to provide increased school choice erred in the calibration of treatment variables. The treatment variables (the variables controlled by those designing experiments) used in school choice field experiments have, so far, been set at levels that are virtually indistinguishable from those of the control runs of the experiment.

Thus, it is little surprise that the resulting data generally show little behavioral response to the treatment variables in school choice experiments, such as voucher systems, privatization initiatives, and charter schools. A key reason why the existing data are poor is that insufficient variation in the treatment variable reduces the precision of any estimated effects of treatment variables on the dependent variables of interest, say, test scores.

To see why, recall that, in a bivariate regression, the estimated standard error of the coefficient is proportional to (the square root of) the sum of squared residuals divided by the sample variance of the X variable, where X is interpreted as the treatment variable that experimentalists control. Running experiments where the different values of X are clustered too close to one another produces very small sample variance of X and, consequently, a large standard error on the estimated value of attempting to measure the effect of institutional change on expected levels of performance. Note that if X measures the institutional features of different school systems, then this variable is probably a multidimensional vector rather than a single number indexing a one-dimensional spectrum. Multidimensionality of X only exacerbates the problem of insufficient variation in X with respect to our goal of learning about regularities in the relationship mapping X into expected levels of school performance.

A second reason why timid choices of values of X that fail to cover a wide enough range hurt the quality of resulting data (in a possibly multidimensional space representing different institutional configurations governing the provision of educational services) is reduced capacity for out-of-sample extrapolation. If experimental data cover only a tiny range in institutional space, then the estimated relationships based on those data will not allow for reliable predictions about as-yet only imagined institutional configurations. As any introductory

textbook on linear regression explains, a small range of variation in the independent variable provides little reliable information when extrapolating from measured relationships between, say, institutions and school cost or school performance, to more radical proposals for institutional reform that entail bolder shifts in the supply and demand conditions of school services.

Couched in antiseptic terms that draw on nothing more controversial than regression analysis 101, the premise that existing school choice data provide an insufficient basis for the real policy analysis and reform measures that our scientific discipline—not to mention the citizen voter— needs should enjoy widespread support among economists of nearly every political or ideological stripe. Designing new experiments that produce data capable of revealing more precisely estimated effects, and giving a basis for prediction over a much wider range of shifts away from the institutional status quo, would seem to be a clear scientific priority. One hopes that stakeholders on many fronts will realize the tremendous value of generating new data to inform school reform and come together to join forces in designing a variety of new experiments whose data will provide positive externalities enjoyed by researchers, policymakers, educators, and families of school-age children alike.

WHICH EXPERIMENTS? WHICH *X'* TO TRY NEXT?

The next question concerns which direction to shift X. We know that the status quo, X_0, does not perform well. We know that existing data are insufficient for estimating a regression coefficient measuring the change in expected school performance (or cost) when X is shifted. The question remains, then: how to choose X'.

The question "Which X'?" is a symbolic translation of the question "Which alternative institutional arrangements should be designed into our next rounds of school choice experiments?" Since real families and school children are involved, there is a deep ethical obligation to choose an X' that will have a very good chance at producing improved levels of performance. Without this imperative to make our experiments as valuable as possible for the people taking part in them, an alternative research program might be proposed in which we shift one element of the vector X at a time to radically inferior positions, just so that we might produce new statistical information about the performance response of schools to different kinds of institutional deficiencies. Of course such a proposal is ethically offensive and scientifically absurd. It serves, however, to underscore how important it is that, when we choose new experimental values of X to cover a wider range of variation, we look for values of X' with a high likelihood, based necessarily on theory since there is no direct empirical evidence about worlds in which X' prevails, of improving the average performance metric. Range of variation is not enough. The level of performance matters, perhaps more than anything else.

But it is precisely in our discussion of which alternative institutional arrangements to pursue that we are likely to get bogged down in debate, possibly consuming so much energy and leading to so much distraction that no new experiments are ever conducted. In fact, these two distinct objectives—of raising performance and of accumulating more data by varying institutional variables more—probably accounts for some of the gridlock that has stood in the way of school choice debates advancing. As mentioned, because the extant data are poor, we have only theory and a small amount of suggestive empirics, open to multiple interpretations, to use as we debate what the high-priority choices of *X'* to use in forthcoming school choice experiments should be.

On the question of which experiments would be most fruitful to run, I want to call cautionary attention to the use of efficient market theory in guiding these experimental design choices. My concerns are twofold, relating to the two following quotes. Merrifield writes, "[W]e can discover a cost effective menu of schooling options, and each item's minimum cost, through market accountability experiments." He also writes, "The entry-price system combination has a great economy-wide track record because it harnesses highly dispersed knowledge about schooling practices, and exploits self-interest through opportunities to compete" (Merrifield, 2009, p. 271). I want to take up these points separately, first focusing on cost-minimization (e.g., finding the efficient frontier in the school services production set) and then on market mechanisms more generally in the context of deciding which school choice experiments should take priority.

COST MINIMIZATION?

Given the choice between the current school system (which many observers agree is broken according to various measures of performance) versus a high-performing system that is 20% inefficient, which would you choose? I do not think *efficiency* in the zero to 20% range is anywhere as significant a priority for most parents—nor should it be—than the *level* of performance; that is, the quality of teaching and education that their children receive. I, for one, would be happy to pay 20% too much in order to find for my children a few teachers who are intellectually alive, not bullied by someone else's curriculum planning, not numbed through years of uninspired teacher training, and committed to boldly pursuing objectives that coincide with ones that matter to me.

As education researchers and economists, estimating the cost-minimizing envelope of school services production technology (e.g., using data envelopment analysis, or DEA) is an intellectually interesting problem. And perhaps from a system administrator point of view, managing the efficiency of a well-functioning system is laudable and socially useful. But if most classrooms are currently failing to produce the outcomes we want, then who cares if some districts produce identically failing outcomes whose costs are 20% cheaper?

Cost-minimization as an overall goal in education seems almost irrelevant in the current context, in large part, because quality is so difficult for both experts and nonexperts to observe. What we want is quality-controlled cost minimization. If we insist on cost-minimization without minding quality, insisting that schools procuring supplies always choose the lowest bidder, we could very well end up with not the most efficient but the lowest quality products in our schools. I am not at all sure I would want my school district to automatically go with cost-minimizing bids, say for school food, playground equipment, or scientific tools like microscopes. The quality dimension seems missing when Merrifield writes, "School districts can competitively contract out school management to gauge the minimum cost of existing approaches to schooling" (Merrifield, 2009, p. 281).

How can they conduct meaningful cost comparisons without undertaking very costly expert assessments of quality? It is doubtful that any company or individual has collected enough data points to estimate an efficiency frontier for producing computer operating systems, or mobile phone systems, or commercial airplanes. Those industries get along fine, in part thanks to their pricing power, which means that they operate on the strict interior of their production sets rather right along the efficient frontier.

Merrifield acknowledges how difficult it is to identify a single amount of money that would provide for an "adequate education." He is surely right that this is indeed a difficult task. What is almost certain, however, is that this amount is less than infinity. Rather than a vision of extreme scarcity, I would suggest that we might benefit by focusing instead on the abundant resources we already enjoy and then work then to better apply them. We have an abundance of resources like experienced adults with time on their hands who could be employed to deliver better education to our children—if only we could find better institutional arrangements to coordinate those abundant resources and the children who need better educational services. I would urge us to stay focused on imagining new institutional forms that can deliver much improved services rather than getting hung up on the less important problem of estimating cost envelopes when quality is very difficult to observe and precisely measure.

ARE MARKET-ORIENTED VALUES OF *X'* THE INSTITUTIONAL EXPERIMENTS WE SHOULD RUN FIRST?

Are real-world markets good at inducing the kind of competition that most of us would like to see in our nation's school system? Merrifield apparently thinks so. He writes, "General experience indicates that innovation and cost-cutting are among the first things to falter when the pressure to lets up" (Merrifield, 2009, p. 272). Thinking of CEO salaries (does anyone think that competitive pressure is what determines them?) or lavish spending on status goods for top managers that surely defy principles of cost-cutting,

one can easily point to systematic departures from cost minimization among Fortune 500 firms operating in ostensibly "competitive" environments. Whereas "competition" in economics means price taking, the word "competitive" in the real business world refers to a company that has accumulated enough pricing power and financial or political resources to insulate itself from price pressure. A firm that sells at a price near marginal cost is viewed as weak and likely to die—not competitive and efficient.

Merrifield acknowledges that winners in private markets do not always exemplify the traits of competition that we want to see in the school system: "History shows that industry leaders eventually become unable to adequately adjust to change" (Merrifield, 2009, p. 272). This, to me, then begs the question of why something we call "the market" should naturally be our first choice when looking for ideas about new institutional systems to deliver schooling services to our children. One can be 100% in favor of competition and yet deeply suspicious that laissez-faire approaches to the allocation of educational services will achieve any gains from competition.

Merrifield at times implies that we have a theoretical basis and, at times, implies we have an empirical record to support his claim that "markets" have a "mostly strong track record outside K–12 education" (Merrifield, 2009, p. 274) and should therefore be examined in K–12 education. Giving the claim the benefit of the doubt and ignoring the myriad counterexamples that so easily come to mind, one only has to recall that introductory micro principles textbooks routinely list conditions for gauging how optimistic we should be about the performance of market mechanisms, on the basis of theory rather than data.

Criteria favoring market institutions that are frequently cited in economics textbooks are homogeneity of the good, ease with which quality can be assessed, symmetry of information about quality of the good, absence of externalities, and eventual diminishing marginal returns technology (e.g., the good is not a so-called natural monopoly for which economies of scale would naturally favor a single producer). Nearly all of these criteria would seem to raise questions—not necessarily damningly negative a priori conclusions but rather genuine questions—about the effectiveness of decentralized markets for allocating educational services.

Even for a laissez-faire skeptic like me, reexamining this list of conditions that favor markets resensitizes and refocuses the sometimes nuanced but hugely important task of finding institutions that promote competition. It is important to distinguish laissez-faire as an institutional configuration from the actual degree of competitiveness and efficiency it produces. Laissez-faire can lead to an absence of competition. And there are many dimensions along which our institutions can encourage or discourage product differentiation and competition, implying that two institutional schemes, which should be thought of as vectors, say X_1 and X_2, may not be easy to order in terms of which is more market oriented or more laissez-faire.

"MARKET FORCES" IS AMBIGUOUS

Merrifield writes that his essay "identifies the challenges and opportunities that can be addressed through policy experiments that harness market forces" (Merrifield, 2009, p. 274). But what are market forces? Is there a scalar valued index of "market-ness" that can order all policy ideas on a single-dimensional spectrum? I think not.

SIGNALS AMISS

Merrifield writes, "A likely high price for a significant innovation drives much business enterprise" (Merrifield, 2009, p. 274). These high prices, however, are in many cases the result of an institutional intervention—patent laws—that intentionally shut down price competition for a decade or more for particular products while perhaps stimulating a different kind of competition for acquiring new patents. Price signals in a more decentralized economic world without patent laws would surely send different signals. Which institutional setting—patent law or no patent law—is more of a "market system?" The question of price signals motivating innovators to make big bets necessary to transform our educational system implies that price controls and protected monopolies could, at least following the logical principle of patent law, be part of the solution.

Merrifield refers to the virtues of market price signals: "Market price signals drive continuous improvement, cost reduction, and product evolution . . . by identifying the reward for topping the existing cost and quality standards reflected in competitive prices" (Merrifield, 2009, p. 274). But price competition is rare in many important industries and, when it takes place among firms with market power (nearly always the case), competition can lead to innovation suppression (e.g., shelving patents), hiding of valuable information, and negative externalities like pollution.

In the context of educational services, all those potential problems are real possibilities. Especially worrying among these potential problems is intellectual pollution. One can imagine a form of pollution occurring as firms who set educational curriculum create as a by-product children socialized with a muted, numbed, or missing appreciation of nonmaterialist definitions of happiness in the canon of Western thought—and beyond. If market mechanisms (whatever is meant by this vague term) produce educational curricula geared to growing the economy as measured by gross domestic product, would skills, sensibilities, and training in aesthetic decoding based on world views aiming for more than increased quantities of consumption of goods and services traded in private markets financed by labor income have any room in the curriculum? For example, the most famous philosophers of ancient Greece disparaged acquisitiveness and

identified high-quality social relations, and the leisure time needed to produce them, as indispensable inputs for achieving happiness. Do we really want institutions that let such ideas die?

Just like some market-oriented economists tend to blame poor people for being poor—that is, they were too lazy, they made the wrong decisions, they made friends with the wrong kids on the school playground—it sounds, at times, as if Merrifield wants to blame bad performing schools on educators themselves rather than on parents or, more culpable still, I think, those who designed the institutions that govern our school system. Merrifield quotes Albert Shanker, former President of the American Federation of Teachers, saying essentially that teachers are to blame. I would much rather blame administrators, managers, principals, and others with antifreedom, antithinking bureaucratic impulses. Rather, we need to channel competition into a subset of the multidimensional set of factors that determine educational quality and focus on the ones that respond positively to competition.

The personal politics of many leading economists who developed models of market economies are instructive in decoding what the word "market" means when describing different institutional configurations for organizing a nation's school system. When we read the Fundamental Welfare Theorem of Economics stating that "Under conditions A, B, and C, competitive equilibria are Pareto efficient," many evangelical free marketeers in the economics profession seem only to see the last phrase, "competitive equilibria are Pareto efficient" and ignore the hypothesis of the statement.

Others among us read the statement of the theorem and are drawn to how stringent the conditions are. Nobel Prize winning economists like Hurwicz, Ken Arrow, or Joe Stiglitz, for example, advocate real-world institutions that channel competition in particular directions and impose strict regulation in others. These economists emphasize that the gap between "Under conditions A, B, and C" and the real world is far too great to expect the conclusion of the theorem to hold without further institutional modification.

With regard to risk-assessment technologies, the recent financial market crisis would seem to imply that inferior innovations can displace incumbents whose technology is superior. Thus, innovation does not always imply progress or monotonically increasing consumer welfare. Did AIG's financial market innovations (new insurance products and new financial derivatives that obscured risk rather than efficiently reallocating it) constitute a superior technology? For comparison, recall that Merrifield writes, "Possible increased profit through innovation always drives some market entry and replacement of marginal incumbent producers" (Merrifield, 2009, p. 279).

Despite these reservations, I say, yes, let us try some of the projects envisaged by Merrifield. Seeing that markets are highly heterogeneous and malleable—and imperfect—gives us a larger set of potential policy tools, not a smaller one. If there is inertia, for example, then inertia can in theory be

productively harnessed to help lock in at an institutional configuration that is performing well. There would seem to be many good-enough routes to envisioning and then achieving meaningful improvements.

One specific that Merrifield writes on which I think is a terrific idea is coupon credits: "coupons that families can use later to finance higher education and or tuition at chartered schools that offer premium services. The coupon amount is debited from the per pupil public funding paid to the school. Credits, as opposed to cash rebates, avoid some potential for fraud" (Merrifield, 2009, p. 285).

INEQUALITY

A few countries have, quite admirably in my view, created educational systems in which nearly all children can expect to receive a rather good education. In the United States, however, despite being the richest country in the world, families' wealth and income dramatically condition their babies' expected quality of education. This inequality of educational opportunity is a defining characteristic of life in different U.S. neighborhoods, in different school systems, in families with different incomes. These differences are powerfully correlated with access to education, health, and virtually all basic services that comprise the material quality of life. Because inequality is a defining feature of life in the United States, educational reforms will naturally be evaluated, in part, on the basis of how they affect inequality.

Merrifield accepts increased inequality as a by-product of allowing for more diversity and encouraging innovation in educational services. He argues that high-services schools with rich customers will innovate and produce technological spillovers that eventually will benefit all. Merrifield writes:

> So, through well-implemented good ideas, as well as bankruptcy for purveyors of the rest, entrepreneurial initiative gradually increases the quality and diversity of well-known services and drives down tuition levels (the copayment amount) until it is difficult for a newcomer to make a normal rate of return on investment by entering the schooling market. (Merrifield, 2009, p. 279)

As a counterexample to this claim that markets ineluctably make consumers better off, think about U.S. agricultural production since the beginning of the 20th century. If one shops at major grocery stores in the United States for tomatoes, for example, certainly agribusiness has managed to reduce unit costs and produce a large quantity of tomatoes for low cost, providing a benefit along the cost dimension. But the quality of these tomatoes in terms of flavor (as anyone who shops for produce in Italy—or India or Mexico—quickly

discovers) is lower than that of tomatoes sold elsewhere using much older technology and, almost without doubt, lower than the quality of tomatoes produced in the United States 100 years ago. Competition and technological innovation that make fat profits do not necessarily leave most consumers with more choice, with higher quality, or generally better off.

Merrifield is right that deregulation of what schools are allowed to charge would provide positive data and welfare improving outcomes for many. But I think perhaps he overstates the case in saying that this deregulation would hurt no one:

> As schools of choice, chartered schools that seek a copayment must offer superior services to be competitive with schools that cost less. Opening the school system to especially costly specialized services helps some families, hurts no one, and yields much valuable data. It increases school funding without raising taxes. (Merrifield, 2009, p. 283)

Some of those who do not receive those superior services will, no doubt, be put at a competitive disadvantage. Perhaps the American education system is already so unequal that allowing for temporarily more inequality in order to raise overall performance levels is a worthwhile tradeoff—perhaps. I think this is what Merrifield is arguing for. I only wish that his essay acknowledged the oftentimes dramatically experienced costs of inequality in the lives of many Americans and the possibility of increasing inequality as a negative side-effect of otherwise desirable policy ideas.

DIVERSITY IN THE SUPPLY OF EDUCATIONAL SERVICES

In advocating the need for more diversity in educational supply (a richer variety of teaching styles, combinations of services, and objectives sought after as primary performance metrics), my intuition tells me that Merrifield is absolutely right. An organic educational system, like a rich environment with great biodiversity, should not be a monoculture, avoiding the monoculture's vulnerability to episodes in which a single pathological element in the mix can lead to extinction. As a consumer of educational services, it might be nice to have a longer menu of choices with very different kinds of educations offered. (On the other hand, long menus might also be confusing and gut-wrenching to parents, for example, having to choose between many expensive high-services options versus low-cost low-services options for one's child.) It is important to acknowledge, however, that diversity of educational supply is almost necessarily in conflict with another goal frequently put forward, which is standardization and the gains achieved by having common metrics of performance, common language, and common content.

TIME HORIZONS AND TRANSITION COSTS

I have written before (Berg, 2009) on the importance of going beyond comparisons of two equilibria to consider costs along the transition path to the new equilibrium. In the case of school choice, I think it should be an ethical requirement that we seriously attempt to account for transition costs, especially for those who will be worse off, for example, by taking part in experimental educational supply systems that prove to be inferior even to the status quo. In his current essay, I appreciate very much that Merrifield mentions the transition cost issue.

A closely related issue is the time horizon over which we can expect predicted gains to be fully achieved. Merrifield writes, "For useful insights, the latter approach will require a long time horizon" (Merrifield, 2009, p. 280). It would be nice to speculate in more detail about how long. I am guessing that we are talking about something on the order of a decade or more, especially in light of transition times required in privatization experiments in Eastern Europe, for example. To accelerate the time horizon and minimize transition costs, introducing competition experiments within the public school system may have a lot to speak in its favor.

Merrifield puts forward a very interesting idea regarding multiyear bidding: "Those terms set the bar for the bidding on schools in later years, as well as for the renewal of contracts. Staggering (some multiyear contracts awarded or renewed each year) accelerates the critical 'top-this' process" (Merrifield, 2009, p. 281). I think Merrifield is right and I would be happy to see multiyear bidding institutions be put in place. This further conclusion, however, is almost surely too simple: "Competitive pressures will gradually reveal and improve the cost, specialization, and quality possibilities" (Merrifield, 2009, p. 281).

Regarding the time dimension in school choice analysis, Merrifield very usefully reminds us that our notions of satisfactory or adequate school performance are, and perhaps should be, dynamical concepts—they change with the times. For example, we expect more math skills than we did a century ago. These disparate points under the theme of the time dimension in school choice should be kept in mind, especially as we undertake before/after event studies or cross sectional analyses of existing test score data.

EXPERIMENTAL CONTROLS

From an experimental point of view, using an entire district as a control is a problem. For this purpose, randomized trials are much better. With randomized trials, there is no room for measured gains from school reform to be attributed to preexisting differences in treatment versus control populations used in the study. Perhaps Merrifield could modify this claim or elaborate on the benefits of his proposed control when he writes, "The districts

without voucher programs serve as the ideal control group for the effects on the district that offers vouchers" (Merrifield, 2009, p. 281).

CONCLUSION

Merrifield's emphasis on expanding parents' range of choice about where to send their kids to school for the purpose of generating valuable new data, and the potential for school systems to learn from each other, is right on target: "That K–12 school systems might greatly benefit from one or the other is a major premise of this essay's discussion of needed insightful school choice experiments" (Merrifield, 2009, p. 275).

At times, it seems that Merrifield thinks of cost minimization and innovation as synonyms. It is true that innovations in technology can lead to cost reductions. But innovations that render older techniques obsolete can actually lead to higher prices. Think of medical technology, for example. Today's technology, presumably in the eyes of most contemporary observers, produces higher quality health care outcomes than the technology of centuries past. But technological innovation has not led to broad price reductions in the provision of medical services.

Merrifield also usefully lists (and has previously documented in fine detail) the failings of past attempts to introduce more competition to the supply of educational services: "The typical key restrictions in charter and voucher laws include limits on the number of new independent schools, caps on total enrollment, open admission requirements, curriculum content rules, and price control through voucher and charter school copayment bans" (Merrifield, 2009, p. 276). Merrifield is right about sometimes cynical policies that serve political imperatives but do not address real educational needs: "Indeed, the political process often funds ineffective programs, and bad economic policy is sometimes good politics" (Merrifield, 2009, p. 277). But the same is true in private business. Sometimes low-quality production is good business even when it is bad for consumers. Think, for example, about recent items in the news about toxic imports of pet food and housing construction that sickened many consumers. Or the famous quote from an Archer Daniels Midland executive caught on FBI cameras in the act of collusive price-fixing: the executive laughed as alleged competitors agreed on a lower limit on prices that "competitor is our friend and the customer is our enemy."

One can view this is an evolutionary step along a path in which badly behaving companies go bankrupt and new companies improve quality as an investment in their reputations. But in these cases, the bankruptcies have not in fact materialized and the evolutionary interpretation seems further from reality than the view I put forward previously: namely, that, just as referees are needed to make professional sports competition interesting, so too

we need strong regulation (e.g., random health inspections and detailed product labeling requirements) to achieve competition in the dimensions that make consumers better off (e.g., raising quality rather than dangerous quality-lowering cost cutting).

Thus, for me, there is much to disagree with in Merrifield's assertion that "priorities are easy to discern from a price system." The priority of families' health, and that of their pets, was not easy to discern in the prices of imported goods that contained difficult-to-detect poison. Price information is, of course, sometimes valuable. But in many cases, price information is woefully incomplete, because product quality is multidimensional and very difficult for consumers to directly observe.

Proponents of decentralized, deregulated mechanisms for allocating educational services should keep in mind severe asymmetries in information (e.g., school managers who are on the premises of educational facilities most of the day versus parents who do not see what happens in the classroom most of the time). They should keep in mind the difficulty for parents to quickly and cheaply assess product quality. And they should keep in mind the large gap between private versus social marginal benefit resulting from positive externalities that occur when parents increase the quality of the educational services they procure for their children.

I keep coming back to my earlier concern that Merrifield assumes that markets will provide a desirable menu of educational choice. This is a substantive question and should not automatically be assumed. Merrifield writes, "Do they trust parents to define the appropriate menu of schooling options through their choices, and if so, are the district leaders willing to subvert the district administration's self-interest in sustaining public school enrollments?" (Merrifield, 2009, p. 281). The question is not whether parents create the choice sets. It is whether the institutions of public education, private education, or a mixed platform, would provide the best choice sets according to various performance and consumer choice metrics.

Merrifield's bold premise is right on target and deserves wholehearted support from reformers of every political stripe when he writes: "I will discuss both approaches. True, the discussion of new institutions will appear radical or just irrelevant because of assumed political infeasibility. But it is the duty of social scientists to present them anyway" (Merrifield, 2009, p. 277). Amen to that.

Too often economists adopt a narrowly Popperian view of their scientific research program, aspiring only to whittle down the list of possibly true theories by occasionally falsifying one with data analysis. As Merrifield points out, scientists should also be in the business of creating new ideas. Institutional design for the provision of educational services is an area long overdue for creative synthesis of new ideas, and Merrifield is undoubtedly a leading voice advancing us down this path. Let the new experiments and data generation begin!

REPLY BY JOHN MERRIFIELD

Both as author and editor, I appreciate Nathan Berg's ability to detect and eloquently describe and critique every assumption I made and some that I did not realize that I made. That serves the journal's mission of advancing and better defining the frontier of knowledge relevant to the journal, which will include reform strategies of all kind, not just school choice, in future volumes. It also provides a treasure trove of research topics and testable hypotheses.

Berg also identifies some needs for clarification. The motivation for "Defining Continuous Improvement and Cost Minimization Possibilities through School Choice Experiments" was the search for a definition of "funding adequacy." Numerous multibillion dollar lawsuits struggled to specify what a school system must contain to provide an "adequate education" and what it will cost to create such a system? I described some market experiments that could help ferret out such information, which means ways to exploit pursuit of self-interest and highly dispersed knowledge and talent by facilitating greater rivalry among producers of schooling services. There was no attempt to assert that the provision of schooling should be driven by a laissez-faire pure free enterprise driven menu of schooling options. For reasons just abundantly provided by Berg, we are a long way from having an evidence base from which to confidently specify the optimal mix of political and market accountability in primary and secondary education, much less the underlying key details of which kinds of interventions are helpful. Movement in that direction is precisely why we need some experimentation, including possibilities that represent only modest departures from the high intervention status quo to more "radical" change in the market-politics mix. The former have the virtue of being more likely to be implemented, but with the drawback of revealing less information and the risk of misleading members of the body politic about the nature of systems with drastically less intervention.

In the context of the funding adequacy debate, the starting point is a definition of adequacy: what school systems should do. Then, we ask how we can know what the minimum cost of optimal/adequate practices are and whether we should allow families with means to pursue schooling for their children that greatly exceeds the adequacy standard that families with lesser means will have to settle for. Given the great deal of inequity in the current system, freedom to deploy earnings toward superior schooling does not necessarily create more inequity than we already have. Also, Berg's discussion of the disadvantage that would result seems to imply a finite amount of economic opportunity, that lower income families would be made worse off by higher income families' increased expenditure on schooling. Holding back upper income families might shrink the current gap between upper and lower, but not necessarily to the benefit of the lower income families. Indeed, policies that ultimately led to a wider gap between more and less advantaged families might be most beneficial to the less advantaged if the result

is an increased rate of improvement for all. Also, it is well worth noting that there is an enormous amount of income mobility and that the membership of more and less advantaged groups is constantly changing, so that inequities, to the extent they exist, do not benefit a particular set of people at the expense of another particular set of people. That said, I have to remind the reader that my point in the article was not assert the desirability of that as a policy but to assert its usefulness as a procedure to generate significant data.

My use of the term "existing approaches to schooling" does not mean classroom practices. I meant governance and funding policies that influence the nature of the school system, all of the schools, public and private, combined. There is a large and debilitating tendency to equate school and school system policies. The latter, given the diversity of children and educators, is about creating the conditions that will lead to mix of schooling services that will engage children in useful and important learning.

A key element of my experiment suggestions was an expanded role for consumer choice, to test the efficacy of increased bottom-up subjective accountability to families and decreased top-down objective accountability to government officials.[1] I did not predict risk-free improved results. High producer turnover in most industries speaks to the high propensity for error in satisfying consumer demand, and to the extent that greater reliance on free enterprise has advantages over political control, it may be because the former corrects errors more readily.

To address Berg's query about my definition of market mechanisms, I mean increased contestability of the provision of services, with the degree to which the political process defines what will be provided varying among the proposed experiments. Though Berg's discussion of textbook descriptions of the conditions of competitive markets creates a "straw man" that was not an issue in my delineation of useful experiments, it still served the useful purpose of identifying research questions that might attend the results of the experiments. Though there is some danger in Berg's frequent discussion of what "can" happen, mostly because it is very unlikely to, but also because it may already exist. For example, Berg's cautions about *possible* outcomes of price competition are a pretty good description of problems we already have, and they are rare in competitive industries. And speaking of price competition, certainly some high prices are sustained by market entry barriers like patents, but even in those cases high prices are important pieces of information and strong incentives to economize and develop substitutes.

Berg posits a false interpretation of an Albert Shanker quote, that teachers are to blame for the woeful outcomes of the public school system. Shanker meant that the current system provides teachers little incentive to change, much less provide them a basis to discern the right changes. However, they are responsible for supporting organizations—their unions—that resist change, including some that might be in the interests of many teachers but not of their union.

I have take issue with Berg's tomato example. He's taken the delineation with potential market malfunction a bit too far when he asserts that cheaper but

flavorless tomatoes are a sign that markets work badly. It may be the result of government intervention. It is a surprisingly regulated enterprise. More likely, the lack of tasty tomatoes is the result of conscious choice by consumers. Not enough prefer the improved the flavor to pay the higher price. Or put another way, unless government regulations are in the way, entrepreneurs would supply the tastier version if enough consumers were willing to buy them.

Finally, Berg notes the inadvisability of entire districts as controls (benchmarks) against which to gauge treatment effects of policies in place in other, otherwise similar districts. He prefers randomized trials. Both approaches are potentially useful and potentially misleading depending on the circumstances. Quasi-experimental design is the formal name of the control group approach, and depending on the issue and the available data, may be the only way to assess the significance of systemwide differences. Randomized trials are appropriate for smaller scale issues where the necessary shortage of the slots in the program to be evaluated does not significantly color the outcome of the experiment. Unfortunately, for large scale school choice programs, there either is not a group of randomly denied program applicants or there is a shortage, which we know from economic theory and abundant experience can significantly effect the behavior of school operators and the potential for new schools to form. Of course, poorly chosen control groups—districts, sometimes—would certainly invalidate the findings of a study. Quasi-experimental design needs to be applied with great care. Randomized trials have been used too often, sometimes assuming away the critical issues, and other times creating misleading findings about competition that does not exist when suppliers face persistent excess demand.

Nathan Berg provided a thorough accounting of the potential issues that should examined by experiments that examine the effects of increased contestability of service provision, but he went a bit beyond the theme of experimentation to ferret out vital information and made the imperfections of free enterprise-based approaches of varying intensity seem more significant as likely issues than they typically are or are likely to be for a school system.

NOTE

1. See http://www.texaspolicy.com/pdf/2001-veritas-2-3-school.pdf for a good, nontechnical discussion of the differences.

REFERENCES

Berg, N., (2009). Illusive competition in school reform: Comment on Merrifield's "Imagined evidence and false imperatives." *Journal of School Choice*, *3*(2), 290–306.

Merrifield, J. (2009). Defining continuous improvement and cost minimization possibilities through school choice experiments. *Journal of School Choice*, 3(3), 271–289.