

The Potential Impact of Vouchers

Patrick J. McEwan

*Department of Economics
Wellesley College*

This article provides an updated review of recent empirical research on the potential impact of private school vouchers. It addresses 3 questions: (a) do students that use vouchers to attend a private school obtain better outcomes than would be obtained in a public school? (b) Do vouchers encourage student sorting and how does sorting affect student outcomes? (c) Does the offer of vouchers promote competition, improving outcomes of students who remain in public schools? Conclusions are that African American students who are offered vouchers experience small achievement gains. The results are highly sensitive to analytical assumptions and are not evident for other racial or ethnic groups. The evidence further indicates that large-scale voucher plans encourage sorting that could lower the achievement of public school students. There is no compelling evidence that such losses are outweighed by competitive gains in public schools. The conclusions on sorting and competition are most applicable to unrestricted choice plans in which flat-rate vouchers are offered to a large number of students with few eligibility restrictions. Some emerging evidence indicates that alternate approaches to policy design could yield more promising results.

Vouchers are tuition coupons that parents can redeem at schools of choice. They are rarely used in the United States, with publicly funded

I am grateful to Phil Levine and Anthony Rolle for their comments. All conclusions are my sole responsibility.

Requests for reprints should be sent to Patrick J. McEwan, Wellesley College, Department of Economics, 106 Central Street, Wellesley, MA 02481. E-mail: pmcewan@wellesley.edu

programs in Cleveland, Milwaukee, and Florida.¹ Nevertheless, it seems likely that vouchers will eventually be implemented on a wider scale. In June 2002, the Supreme Court held that the inclusion of religious schools in the Cleveland voucher program was not an unconstitutional establishment of religion under the First Amendment. Because existing private schools are mostly sectarian, the court's decision in *Zelman v. Simmons-Harris* removed a legal barrier to the implementation of vouchers. Since then, Colorado's governor has signed a voucher bill, and plans are afoot in other states to push for voucher legislation (Richard, 2003; Walsh, 2002).

This article reviews the extant research on vouchers and asks whether it can provide a useful guide to voucher policy. An earlier review identified three questions that should frame an assessment of vouchers (McEwan, 2000b). First, do students that use vouchers to attend a private school obtain better outcomes than would be obtained in a public school? Second, do vouchers encourage student sorting—perhaps cream-skimming—by ability, income, race, or other attributes, and how does sorting affect student outcomes? Third, does the offer of vouchers promote competition, perhaps leading to improved outcomes among students who remain in public schools?

The first question is relevant to assessing the impact of small-scale voucher programs that restrict eligibility to a small number of students, perhaps from poor families or low-achieving schools. The second and third questions are relevant to the assessment of larger-scale plans that extend eligibility to many students (thus facilitating widespread sorting and competition). Based on research available at the time, McEwan (2000b) concluded that evidence supported a small-scale offering of vouchers to poor children, especially African Americans in urban areas. At the time, the evidence was not sufficiently compelling to justify strong support for—or opposition to—large-scale voucher programs. Other reviews shared this agnostic conclusion (Gill, Timpane, Ross, & Brewer, 2001; Ladd, 2002; Levin, 1998; Neal, 1998, 2002).

It is a propitious moment to reconsider the questions. First, there has been tremendous growth in empirical research, much of it devoted to pre-

¹In 2000, statewide voucher initiatives were defeated at the polls in California and Michigan (Walsh, 2000). Outside the United States, in countries from Chile to New Zealand, there are intriguing experiences with vouchers and large-scale school choice, but these experiences seem to have rarely filtered down to the U.S. debate. For an overview of international experiences with vouchers and school choice in Chile, New Zealand, Sweden, and elsewhere, see Plank and Sykes (2003).

viously neglected topics (such as peer effects).² Second, we have access to additional years of pioneering randomized voucher experiments in several cities (Howell & Peterson, 2002; Krueger & Zhu, 2002).³ Though lauded for their quality, the experimental findings have been dogged by controversy (even among their authors). There is, therefore, much to be gained by sifting through competing claims.

The article concludes that a case may exist for small-scale vouchers that are targeted at poor African Americans in urban areas (in contrast, there is no evidence that vouchers will improve the achievement of other racial or ethnic groups). However, the evidence is much less supportive of large-scale vouchers, particularly if they provide a flat per-pupil voucher to all students.

A Framework for Assessing the Impact of Vouchers

Three Questions

A voucher plan makes some families eligible to receive tuition coupons for attending a private school. In practice, eligibility has been extended to students in poor families (e.g., Cleveland and Milwaukee), to students in low-achieving schools (e.g., Florida), and to all students (e.g., Chile's national voucher plan).⁴ The offer of a voucher, all else equal, reduces the price of attending a private school for eligible families and encourages some proportion to demand private rather than public schooling.⁵ Whether families actually attend private schools in the short run is further determined by the local supply (e.g., admission slots in local private

²Much of the new research is by economists. For example, only 9 articles in the National Bureau of Economic Research Working Paper series were concerned with education in 1998 (as judged to their Journal of Economic Literature code). By 2002, this had risen to 48. Many of the new articles analyzed topics relevant to the voucher debate, including private school effects, sorting, peer effects, and competition.

³McEwan (2000b) only reviewed the first-year results from recent randomized experiments.

⁴As Figlio and Page (2003) illustrated with Florida data, these decisions can lead to large differences in the size and characteristics of eligible populations. Even more complicated eligibility schemes have been proposed, including a recent bill in Colorado that would extend eligibility to poor, low-achieving students who happen to reside in districts with a given number of low-achieving schools (Richard, 2003).

⁵Some plans may restrict the ability of private schools to charge tuition beyond the voucher (so-called add-on payments), whereas other plans would allow add-ons.

schools) and the willingness of private schools to accept vouchers.⁶ In the long run, the schooling market will further adjust to the voucher offer.⁷ First, growing demand on the part of voucher recipients may lead new private schools to enter the market and existing private schools to alter their scale, inducing a further reshuffling of students across schools. Second, public schools may face incentives to alter their behavior, especially if public school revenues are penalized by exiting students.

The *net* impact of vouchers on student outcomes is the result of three separate effects. First, do students that *use* vouchers to attend a private school obtain better outcomes than would be obtained in a public school? Second, vouchers encourage students to attend different schools, thus altering the distribution of student attributes—such as ability, race, or income—across public and private schools.⁸ Thus, one must ask two related questions: Do vouchers encourage student sorting and—if so—how does sorting affect student outcomes? Third, one must consider the possible responses of public schools to the introduction of vouchers. It is unlikely that they will be indifferent to the departure of many students, especially if their revenues are penalized. Thus, one must ask whether the offer of vouchers promotes competition, thereby leading to improved outcomes among students who remain in public schools.⁹

It is important to keep two issues in mind. First, it is necessary to answer *all* the questions to derive meaningful inferences about the net impact of vouchers. The point may seem obvious, but it is occasionally lost in the din.

⁶In some plans, specific categories of private schools might be excluded from participation (e.g., Catholic schools; although this seems less of a concern in the wake of *Zelman v. Simmons-Harris*). In other plans, private schools might be empowered to reject voucher students altogether, or to reject specific students that do not meet their admission criteria. Nonetheless, individual plans may regulate admissions of participating schools, perhaps by mandating lottery admissions when schools receive too many applicants.

⁷The work of Nechyba (2000) makes it clear that housing markets will adjust as families potentially alter their residential choices. This article does not consider that issue. See Neal (2002) for a discussion in the context of vouchers.

⁸The exact nature of this sorting cannot be predicted and, indeed, can likely be affected by the exact features of the voucher plan (Epple & Romano, 2002). The common presumption of many is that sorting will resemble cream-skimming, in which the most able or highest income are the first to exit public schools. This is, however, an empirical question.

⁹The answer, to some extent, depends on whether public schools have much scope to alter student outcomes within existing resource constraints. The usual presumption of voucher proponents is that public schools operate inefficiently and therefore have much room for competitive efficiency gains. The evidence for this conclusion is drawn from the raft of production function studies summarized by Hanushek (1997), who showed few consistent partial correlations between school resources and student achievement. This article does not resolve that debate but instead focuses on the literature that attempts to directly measure the competitive effects of private school competition.

For example, voucher supporters highlight the virtues of competition on public schools, whereas opponents emphasize the deleterious consequences of cream-skimming and declining peer quality in public schools. It is possible that both effects exist and work in opposite directions (for a theoretical illustration of this point, see Hsieh & Urquiola, 2002).

Second, it is unlikely that all voucher plans will produce similar effects (despite the tendency here and elsewhere to refer to a generic voucher scheme). Most obviously, the size of the eligible population will greatly condition the likely outcomes. If a small proportion of students are voucher-eligible, then effects are limited to students who use vouchers to attend private rather than public schools. If a large proportion of students are voucher-eligible, then widespread student sorting and competition may ensue. Other potentially relevant design features include the size of the voucher, the ability of private schools to charge add-on payments beyond the voucher amount, and the relation between the voucher and student attributes (e.g., larger vouchers for less privileged students). Some of these issues are addressed—to the extent that evidence permits—later in the article.

Methodological Issues

In answering the three questions, researchers must establish a causal relation between various “treatments” (e.g., attending a private school, peer-group characteristics, private school competition) and student outcomes (e.g., academic achievement).¹⁰ Consider the problem of identifying the causal effect of attending a private school rather than a public school on achievement. Ideally, the researcher would measure multiple students’ achievements after attending a private school, and the same students’ achievement *had they instead attended a public school*. Of course, it is impossible to estimate a particular student’s achievement in two states of the world. But it is the difference between the observed and the counterfactual states that would describe a causal effect.

Instead, researchers attempt to identify reasonable estimates of the counterfactual. These rely on comparing two groups of students: a treated group that attends private school and a control group that does not. Whether the control group provides a reasonable counterfactual is the crux of the researcher’s dilemma. The best method of ensuring a good

¹⁰For a detailed overview of evaluation design, see Shadish, Cook, and Campbell (2002). For a more intuitive summary, with specific references to education, see McEwan and McEwan (2003).

counterfactual is randomized assignment to treatment and control groups.¹¹ In fact, most voucher research does not employ randomized, or even deliberate, assignment to treatment and control groups. Rather, it relies on treatment and comparison groups that are formed by selection.

In nonexperimental studies of private school effects, for example, students choose whether or not to attend a private school, and schools choose whether or not to accept them. The students who remain in public schools—the counterfactual in this case—are potentially quite different from private school students in ways that are both observed (e.g., parental education) and unobserved (e.g., motivation). The danger is that these individual differences will lead to differences in student outcomes across treatment and comparison groups that are mistakenly attributed to the treatment. Statistical techniques like regression analysis can be used to control for observed differences between treatment and comparison groups that affect achievement—thus rendering the comparison group a more appropriate counterfactual. However, there is a lurking threat that unobserved differences remain. If these unobserved differences are correlated with student outcomes, then it difficult to isolate the causal effect of the treatment.

In recent years, researchers have applied a wide range of econometric methods to address these biases. The methods usually involve a search for variation in treatment status that is unrelated to student outcomes, or a so-called “natural” experiment.¹² Doing so requires the identification of an instrumental variable that is correlated with treatment status (i.e., the probability of attending a private school), but uncorrelated with student outcomes.¹³ If these conditions are fulfilled, then something akin to a randomized experiment can be recovered from the data.

Private School Effects

Nonexperimental Research

The literature on private school effects falls into two categories: a large one that includes nonexperimental studies of secondary school outcomes and a smaller one that includes experimental studies and findings from el-

¹¹In fact, the voucher experiments conducted thus far have not randomly assigned students to private and public schools. They have, more practically, randomly assigned students to receive or not receive a private school scholarship. This is discussed in a later section.

¹²For an overview, see Angrist and Krueger (1999).

¹³More correctly, the instrument must be uncorrelated with unexplained variation in student outcomes (i.e., the error term in the regression that explains student outcomes).

ementary and secondary students in several cities. In reviewing evidence from the first category, McEwan (2000b) concluded that attending private secondary schools—mainly Catholic—leads to small gains, if any, in the academic achievement of secondary students (see pp. 114–117 and the citations therein). In contrast, there are consistent links between Catholic school attendance and attainment, including the probability of graduating from high school or attending college.¹⁴ These effects are generally larger for African Americans in urban areas.

As the previous section emphasized, the causal interpretation of such results hinges on the instrumental variables that are used to correct for selection bias. The literature employs two kinds of instruments: (a) measures of the Catholic religious affiliation of the parents and (b) measures of the local density of Catholic schools or populations. In each case, the authors presume that the instruments are correlated with the probability of attending a private secondary school, but uncorrelated with (unexplained) student outcomes. Several authors have expressed doubts about whether the second condition is fulfilled and, hence, whether the “corrected” results truly have a causal interpretation (e.g., Figlio & Stone, 1999; Murnane, Newstead, & Olsen, 1985).

In a recent article, Altonji, Elder, and Taber (2002) explored the controversy in detail. Using two national datasets¹⁵ and multiple instrumental variables, the authors examined the causal effect of Catholic school attendance on academic achievement and attainment. Several tests of instrument quality led the authors to conclude that none of the common instrumental variables are appropriate, given the strong likelihood that they are correlated with unexplained student outcomes.¹⁶ They pessimistically concluded that randomized experiments provide the best hope of convincingly identifying the private school effect.

The vast nonexperimental literature has devoted little attention to other student outcomes such as risky behavior or civic participation. Using NELS data, Figlio and Ludwig (2000) found that Catholic school attendance reduces teen involvement in sexual activity, arrests, and the use of cocaine, although there are no effects on the use of alcohol, tobacco, or mar-

¹⁴See, for example, Evans and Schwab (1995), who used the High School and Beyond Survey; Neal (1997), who used the National Longitudinal Survey of Youth; and Grogger and Neal (2000), who used the National Education Longitudinal Study (NELS) data. Figlio and Stone (1999), also using the NELS data, found somewhat weaker results on attainment using different instrumental variables.

¹⁵NELS and the National Longitudinal Study.

¹⁶For example, they find consistent correlations between the instruments and student outcomes—conditional on a wide range of other student characteristics—in a sample of public school eighth graders (who almost never attend Catholic high school).

ijuana. Contrary to other findings, the effects appear to be concentrated among students who live in two-parent families in the suburbs.¹⁷ Another article, using the National Longitudinal Study of Adolescent Health, found no evidence that Catholic schools affect a wide range of social behaviors (Mocan, Scafidi, & Tekin, 2002). Finally, Dee (2003) found that individuals who attended Catholic schools were more likely to vote as adults.

Experimental and Quasi-Experimental Research

The experimental evidence is sparse, and it mainly explores the impact of vouchers on short-term outcomes like academic achievement. Howell and Peterson (2002) reported the findings from pioneering randomized voucher experiments conducted in New York City; Dayton, Ohio; and Washington, D.C. The most recent findings encompass 3 years of results from New York and Washington and 2 years from Dayton. Private-school vouchers, funded by private sources, were offered to public school students from low-income families. The vouchers were capped at \$1,200 to \$1,700 per year, depending on the city, and the families were expected to supplement the voucher amount to cover the full tuition. A treatment group of voucher recipients was chosen randomly from among the applicants, as was a control group of nonrecipients.¹⁸

Not all students in the treatment group actually used a voucher. In New York City, for example, about three quarters of students used a voucher in at least 1 year (Krueger & Zhu, 2002), with lower rates in other cities. In analyzing the results, one can compare the outcomes of the entire treatment group—including students who declined a voucher—to the entire control group. Doing so measures the effect of *offering* a voucher (i.e., the “intent to treat”). One can also apply more sophisticated methods to estimate the effect of actually *attending* a private school (i.e., the “treatment on the

¹⁷Figlio and Ludwig (2000) used a somewhat more plausible set of instruments, distinguishing this article from previous work on achievement and attainment. They posited that variation across geographic areas in the availability of public transportation—a proxy for local transportation costs facing parents—will be associated with Catholic school attendance. They further assumed that it is uncorrelated with unexplained student outcomes.

¹⁸For further details of the experimental design, see Howell and Peterson (2002). For especially detailed accounts of the design and conduct of the New York experiment, see Hill, Rubin, and Thomas (2000) and Mayer, Peterson, Myers, Tuttle, and Howell (2002).

¹⁹In a regression with achievement as the dependent variable and private school attendance as the main independent variable, one treats the latter as endogenous in a two-stage

treated”).¹⁹ In practice, the latter are scaled-up estimates of the intent-to-treat estimates. Neither is more correct than the other, but they are interpreted differently.

Howell and Peterson (2002) reported the effect of attending a private school. After 3 years, there was a 9.1% gain in composite test scores for African Americans in New York City. There were no statistically significant gains after 3 years for African Americans in Washington, D.C. And, after 2 years in the Dayton experiment, there were positive effects for African Americans—but they were only statistically significant at 10% and disappeared after controlling for student background variables.²⁰ In general, the positive results for African Americans were driven by the upper-elementary grades, although there is no obvious explanation for this. In sharp contrast to these results, there are no statistically significant effects for other groups (mainly Hispanics in New York City, Whites in Dayton, and a mixture of ethnicities in Washington).

Krueger and Zhu (2002) reanalyzed the New York City data and found less optimistic results. In estimating the effect of being offered a voucher, they found that prior results were sensitive to the definition of the sample (Howell & Peterson, 2002; Mayer et al., 2002). The original analyses excluded students who did not take a baseline pretest (including, for example, the entire cohort of kindergartners). Using a sample that was 44% larger than the original, Krueger and Zhu reported third-year estimates for African Americans that were smaller in magnitude and, depending on the model specification, sometimes statistically insignificant. They also used alternate definitions of students’ racial definitions and found that results were further attenuated. This critique, in general, seems less important given the inherent subjectivity of measuring race. At best, Krueger and Zhu showed that treatment effects vary across subgroups of African Americans when race is defined differently.

In a response, Peterson and Howell (2004) sharply disputed the results of the secondary analysis. However, their arguments do not acknowledge basic concepts in experimental design, particularly with regard to

least-squares framework. It is instrumented with a variable indicating whether a voucher was offered. The instrument should be highly correlated with actual private school attendance, whereas uncorrelated (by design) with unexplained achievement.

²⁰In a typical experiment, one might anticipate that estimated treatment effects would be more precisely estimated after controlling for student background.

²¹For example, they argue that “adding students for whom no test scores are available at baseline raises the risk of introducing bias” (Peterson & Howell, 2004, p. 7). In general, it is true that controlling for baseline test scores diminishes bias in estimates of treatment effects, but only in experimental samples that are already compromised by attrition from treatment or control groups (for a discussion, see Shadish et al., 2002). In the case of the New York experi-

the use of an expanded sample.²¹ A separate response by Mathematica Policy Research—collaborators in the New York evaluation—generally accepted the new results (Myers & Mayer, 2003). It concluded that the results are indeed sensitive to important analytical decisions about the sample and the definition of race. A conservative reading of the accumulated findings is that offering vouchers leads to small improvements in African American test scores, but these effects are sensitive to the exact definition of race. As before, there are no effects for other racial and ethnic subgroups.

Lastly, there are findings from quasi-experimental evaluations of vouchers. Two studies of the Milwaukee voucher program took advantage of the fact that voucher recipients were supposedly selected at random if individual schools received too many applicants (Greene, Peterson, & Du, 1998; Rouse, 1998). The putative lotteries created a number of experiments at the level of each school, allowing a treatment group composed of each school's voucher recipients to be compared with a control group of the randomly rejected applicants. Using various subsets of data and multiple statistical methods, Rouse concluded that there were small effects in mathematics and none in reading.²²

The early controversies over the Milwaukee findings highlighted the importance of sound research design (and paved the way for later randomized experiments). But the results themselves appear to have had little enduring impact on opinions about vouchers—an ironic coda to the vitriolic debate of the late 1990s. That may be because the early evaluations were limited to students attending a small number of secular private schools. Since then, the Milwaukee program has been expanded to include religious schools (and *Zelman v. Simmons-Harris* has apparently removed some obstacles to doing so in Cleveland and elsewhere). In this context, recent randomized experiments seem to hold greater promise for informing policy decisions.

Finally, a team of economists conducted a quasi-experimental evaluation of a Colombian program that awarded secondary school vouchers to more than 125,000 low-income students in public schools (Angrist, Bettinger, Bloom, King, & Kremer, 2002). Individual cities almost always used lotteries to allocate vouchers when applicants exceeded the number of available scholarships (similar in spirit to the Milwaukee evaluations). ment, eliminating observations without baseline scores introduces a large amount of attrition. Krueger and Zhu (2002) were correct in stating that "because of random assignment ... estimates are unbiased even without conditioning on baseline information" (p. 3).

²²This finding is at odds with an earlier evaluation that instead used a sample of Milwaukee public school students as the comparison group (Witte, 1998). For a more complete discussion of the Milwaukee findings, see McEwan (2000b, pp. 120–121).

In a comparison of lottery winners and losers, the authors found that a voucher offer produced test score gains of about 0.2 standard deviations.

External Validity

Even if the previous research has a causal interpretation, one should still ask whether the results possess external validity—that is, whether they can reliably inform voucher policy in other contexts. There is no recipe for determining external validity, but some caveats can be mentioned. First, the preceding results are best generalized to similar populations of students (i.e., poor, minority students who have expressed an interest in receiving a voucher). The New York City experiment, for example, found no effects for Hispanic students (one of the few findings that everyone agrees on). Logically, one might be skeptical about the worth of offering vouchers in heavily Hispanic communities, at least until there is a better understanding of the context-specific reasons for their effects in New York City. This point was lost in the debate leading up to the defeat of a statewide voucher referendum in California in 2000, a state in which 43% of public elementary and secondary students in 2000 were Hispanic and 9% were African American (National Center for Education Statistics, 2003). To further complicate matters, the Hispanic population of New York City is mainly Puerto Rican and Dominican, whereas that of California is Mexican and Chicano. There is no way of gauging whether the “Hispanic” effect—or lack thereof—would generalize across ethnic or regional subgroups of the Hispanic population.

Second, the previous findings are best generalized to contexts in which voucher students would attend Catholic schools (for, in existing research, students overwhelmingly attend Catholic schools). Many proposed voucher plans, however, are larger in scale and encourage the creation of new private schools. The current evidence cannot tell us whether newly created schools would duplicate the effects (or lack thereof) of Catholic schools.²³ Some empirical evidence argues in favor of caution when making assumptions about the effects of new private schools. In Chile’s national voucher plan, many nonreligious private schools were created in addition to large numbers of existing Catholic schools that opted to accept vouchers. The new private schools appear to produce lower achievement, all else equal, than Catholic schools (McEwan, 2001; McEwan & Carnoy, 2000). In the United States, legislation has spurred the growth of privately

²³On this point, see McEwan (2000b) and Neal (2002).

managed, publicly funded charter schools. The incipient evidence suggests that charter schools in California, Michigan, and Texas produce achievement that is similar to that of public schools (Bettinger, 1999; Hanushek, Kain, & Rivkin, 2002; Zimmer et al., 2003).²⁴

Third, current research does not specify exactly *why* private schools might produce better outcomes than public schools. One possibility is that private schools use more resources than public schools or allocate the same resources differently. Another possibility is that private schools enroll students with different characteristics, on average, than public schools (e.g., ability) and that private students benefit from exposure to “better” peers. (The next section cites the mounting evidence on the existence of peer effects.) If the private school effect is mainly a peer effect, then it would likely change in the wake of a large-scale application of vouchers. A widespread use of vouchers would lead to extensive student sorting across public and private schools—probably altering the distribution of peer-group characteristics.²⁵

There is some evidence on the role of peer groups in explaining private school effects. Some nonexperimental studies controlled for a limited range of peer variables—such as mean classroom or school socioeconomic status—when estimating a private school effect (e.g., Neal, 1997); other studies controlled for a wider array of variables (e.g., Figlio & Stone, 1999). The latter tended to find smaller private school effects. One implication is that part of the private school effect is “explained” by different endowments of peer-group characteristics across private and public schools. Somers, McEwan, and Willms (2004) analyzed data from 10 Latin American countries. The average private school effect across countries was 0.3 standard deviations, when controlling for individual characteristics such as socioeconomic status. The average effect declined to 0.04 standard deviations on controlling for the mean socioeconomic status of peer groups.²⁶ This evidence does not imply that typical estimates of private school effects—lacking controls for peer-group variables—are incorrect. But, even if they are nominally unbiased, they have a very specific interpretation.

²⁴All of the studies have access to longitudinal data on student performance. In particular, the recent studies on California (Zimmer et al., 2003) and Texas (Hanushek et al., 2002) used exceptionally rich panel data on individual students. Using fixed effects models, therefore, they were able to control for student characteristics that affect achievement—both observed and unobserved—that are constant across time or have a constant time-trend. Though not randomized experiments, these studies allow much more complete controls for student characteristics than previous studies.

²⁵For similar points, see Neal (2002) and McEwan (2000b).

²⁶For a similar illustration using data from Chile, see McEwan (2001).

For that reason, they are difficult to generalize to voucher systems that would encourage widespread student sorting and reshuffling peer-group characteristics across schools.²⁷

Sorting and Peer Effects

The Effects of Vouchers on Sorting

McEwan (2000b) reviewed the evidence on sorting in small-scale voucher programs in several U.S. cities, finding mixed evidence on whether those who take advantage of vouchers are substantively different from other eligible families. In some cases, voucher-users appeared to have parents who were more educated and involved in their child's education, implying that vouchers might exacerbate stratification. Yet, in other cases, voucher-users were more likely to be low-achieving and belong to minority groups. The evidence was further clouded by the fact that most of these programs restricted eligibility to lower income families. Thus, even if participating families were relatively more privileged than other eligible families, they were certainly less privileged than most families in public schools. Thus, cream-skimming might occur in some instances, but the skimming is done from a decidedly small bottle.

In fact, there is even greater interest in how large-scale voucher plans would alter sorting. The evidence on this point is limited because there are few such policies. Figlio and Stone (2001) examined the *current* enrollment patterns in private and public schools, using NELS data. They found, perhaps unsurprisingly, that cream-skimming is evident when private schools charge tuition. Private school students are more likely to be White and have a high family income, high socioeconomic status, and high ability. Figlio and Stone viewed this as *prima facie* evidence that a private school effect could be determined, in part, by peer characteristics (as the last section suggested). Still, this evidence does not tell us about the patterns of sorting when students face *diminished* constraints on their ability to choose private schools. One such example is Chile's large-scale voucher plan, implemented in 1980. Hsieh and Urquiola (2002) suggested that the first-order consequence of the reform was to induce cream-skimming on a large scale. In municipalities with large increases in private school market share, public schools displayed large declines in

²⁷In the language of economists, it is difficult to generalize the results of partial-equilibrium voucher experiments to the general-equilibrium outcomes of large-scale voucher plans.

socioeconomic status and test scores (relative to all schools in the municipality).²⁸

There is also recent evidence from public school choice programs in which students are no longer constrained to attend neighborhood public schools. In Chicago, roughly half of public students opt out of the neighborhood high school to which they are assigned. Using exceptionally rich panel data on students, one study found that higher ability students were much more likely to exercise choice and that they tended to choose schools with higher ability peers (Cullen, Jacob, & Levitt, in press). In New Zealand's system of public choice, Fiske and Ladd (2000) found that "choosing" families were most likely to opt for higher socioeconomic status schools and that additional choice led minorities to become increasingly concentrated in low socioeconomic status schools (also see Ladd & Fiske, 2001). Dee and Fu (2003) considered whether the introduction of charter schools in Arizona—a form of public school choice—led to cream-skimming. Using panel data on schools' enrollments, they found that the introduction of charter schools reduced the share of White enrollments in public schools by 2%.

Several authors have constructed computational models that simulate the sorting behavior of parents in response to the introduction of vouchers. In one such model, Epple and Romano (1998) found that the introduction of a flat per-pupil voucher—available to all students—leads to increasing income and ability stratification across schools. Their model's results are driven by the assumption that school quality is influenced by the mean ability of peer groups (i.e., that peer effects exist). Because families care about school quality, they are assumed to also care about school peer groups. Epple, Figlio, and Romano (2004) found that the model's key implications are consistent with U.S. data. The model is also broadly consistent with sorting patterns in Chile, where similar flat per-pupil vouchers were offered.

The preceding results are subject to a caveat, however, given that there are many possible voucher schemes that could be implemented. In an extension of their model, Epple and Romano (2002) showed that voucher-induced stratification diminishes, for example, if larger vouchers are given to lower ability students, thereby reducing private schools' incentives to "cream-skim." The additional results are suggestive that policy design matters a great deal in determining patterns of student sorting. Although a fair amount of evidence suggests that cream-skimming would occur in *un-*

²⁸The main point of their article was that the massive sorting induced by vouchers severely complicated efforts to assess whether competition improved student outcomes (a point that is taken up again in the next section).

restricted choice plans, it would be unwise to generalize those results to *any* voucher plan.

The Effects of Sorting on Student Outcomes

Sorting is particularly important if student outcomes are influenced by characteristics of their peer groups. Thus, a vast nonexperimental literature tests for peer effects using regression analysis to estimate partial correlations between individual students' outcomes and the characteristics of their school-based peer groups (e.g., mean income, parental education, or ability). That literature has produced a mixed bag of positive and negative correlations, and few definite conclusions have emerged about the role of peers in determining individual achievement.²⁹ In part, the lack of consensus stems from the fact that peer groups are formed by selection. For example, parents choose residences—and, by corollary, schools—based on a range of school attributes, including the characteristics of other students. Imagine that highly motivated families choose schools because the other parents are highly educated. In a regression analysis that does not control for motivation and other individual variables, it is hard to disentangle the causal effects of peer attributes (mean parental education) and unobserved individual attributes (motivation). For a broad overview of these methodological challenges, see Moffitt (2001) and Evans, Oates, and Schwab (1992).

One solution would be the identification of instrumental variables that are correlated with peer-group variables but uncorrelated with unexplained student outcomes. Two recent studies have done so (Gaviria & Raphael, 2001; Robertson & Symons, 2003), but they do not provide a strong case for the quality of their instruments (Moffitt, 2001). Other studies have made use of extensive longitudinal data on student achievement available in Texas. Hanushek, Kain, Markman, and Rivkin (2003) controlled for observed and unobserved student characteristics (such as motivation) by including fixed effects. They found that the achievement level of peers has a positive effect on individual achievement. Using a different empirical approach with the same data, Hoxby (2000) compared the achievement of adjacent cohorts that had slightly different peer characteristics (which she argued are random fluctuations). Like Hanushek et al., she found that student achievement is influenced by peer achievement.

²⁹The starting point in this literature is the Coleman Report (Coleman et al., 1966). A more recent study, using data from several countries, is Zimmer and Toma (2000). For complete citations, see McEwan (2003).

McEwan (2003) used Chilean data to compare the achievement of twins that were assigned to classrooms with different peer characteristics. Doing so provided better controls for the unobserved background characteristics of students and families (because twins share many of these characteristics). The results suggested that the mean of mothers' education in a classroom is strongly associated with individual achievement. And finally, there is a burgeoning literature that estimates peer effects in colleges (Sacerdote, 2001; Williams & Zimmerman, 2003; Zimmerman, 2003). It relies on the apparently random assignment of roommate pairs in college housing, thus avoiding the selection problems that are endemic to this literature. However, the findings are difficult to generalize to elementary and secondary education.

The Effects of Competition

As public schools lose students and revenues under a voucher system, they may face competitive pressures to raise the achievement of remaining students and forestall their exit. But, as with sorting, there are few concrete opportunities to measure the impact of competition in large-scale voucher systems. Instead, a growing literature has sought to estimate the effects of private school competition on public school outcomes in current schooling systems.

All such studies identify a measure of private school competition, usually the proportion of private school students in a local schooling market. There is little clarity on what level of aggregation constitutes a "local" market. In practice, authors have used the proportion of private school students in districts, countries, larger metropolitan areas, and even states. Jepsen (2002) showed that results are sensitive to these analytical assumptions. Regardless of the exact measure, authors use regression analysis to estimate the effect of competition measures on public school outcomes, controlling for a range of student background variables.

The authors of these studies face two empirical challenges in identifying the causal effects of competition (Dee, 1998; McEwan, 2000b). First, private school enrollments and public school outcomes are determined simultaneously (that is, each variable exerts causal influence on the other). Although more private school enrollments may improve public school outcomes via competition, it is possible that low public school outcomes cause additional private schools to gain market share. The lat-

ter tends to bias downward the estimated effects of competition. Second, the level of private school enrollments in a given community may be confounded with unobserved features of the community that also affect achievement. For example, private schools may locate in certain communities because they are wealthier or have “better” students. If these variables are not held constant in a regression analysis, then the estimated effect of competition is biased, perhaps upward.

To solve these problems, authors have attempted to identify instrumental variables that are correlated with local private school enrollments but uncorrelated with unexplained public school outcomes. Most authors used variants of the Catholic population density in a particular community, following roughly the same logic as previous studies of private school effects. However, the dubious performance of similar instruments in that research (e.g., Altonji et al., 2002) might lead one to interpret these results with caution. In reviewing an early subset of competition studies, McEwan (2000b) found weak effects. Jepsen (2002) arrived at a similar conclusion, reviewing a broader array of evidence. In the most complete review of the competition literature, Belfield and Levin (2002) found that an increase of private school market share by 1 standard deviation produces, on average, less than a 0.1 standard deviation increase in local public school outcomes. In sum, the evidence on current schooling systems does not provide strong support that private school competition improves public school outcomes.

Yet, one might ask whether it is even possible to estimate the effects of competition in the previous empirical framework (Hsieh & Urquiola, 2002). For the sake of argument, let us imagine that authors succeed in identifying a “perfect” instrument for the private school market share (i.e., a variable that is strongly correlated with private school market share and uncorrelated with unexplained student outcomes). That would be akin to a randomized experiment in which a number of communities were randomly selected to receive a higher private school market share (the treatment group) and other communities were randomly denied the treatment (the control group). The essential premise of existing research is that one can compare the posttreatment outcomes of public school students in each group and estimate the causal effect of competition. Consider, however, that the treatment leads to student sorting in treated communities—sorting that probably occurs on variables that are imperfectly observed by the researcher (e.g., motivation). If the sorting resembles cream-skimming, for example, then the achievement of “treated” public school students may fall simply because motivated students have moved to private schools (just as the op-

posite would be true if the least motivated students moved to private schools).³⁰ Regardless, the causal effect of competition on public school outcomes cannot be easily separated from sorting effects.³¹

Hsieh and Urquiola (2002) argued that the best researchers can do is to estimate the *net* effect of competition and sorting on public school outcomes. Using longitudinal data from Chile's voucher system, they assessed whether the mean academic achievement of *all* students in municipalities changed in response to increasing private school market share over time. They found little evidence for either increases or decreases in achievement, consistent with prior analyses of the same data (McEwan, 2000a). One plausible explanation is that competition and sorting effects worked in opposite directions and produced little net impact.

Evidence on competition has slowly emerged from voucher programs in the United States. Greene (2001) evaluated the Florida A+ accountability program, in which students in low-achieving public schools were offered vouchers. He argued that the threat of vouchers led to large gains in achievement of those schools, implying some role for competition. But other authors found similar gains for low-achieving schools in North Carolina, for example, that were not threatened with vouchers (Ladd, 2002; Ladd & Glennie, 2001). Given this, Ladd argued that achievement gains are more likely caused by "the increased scrutiny, shame and additional assistance associated with being labeled a low-performing or 'failing' school" (p. 15). Another explanation, forwarded by Hanushek and Raymond (2002), is simply that schools experienced transitorily low achievement in the year they were selected because of unlucky circumstances. Then, in subsequent years, their achievement reverted upward to the mean, a common effect that is easily confused with competition unless empirical adjustments are made.³²

Finally, Hoxby (2003) examined whether the Milwaukee voucher program spurred competitive improvements in public schools. The program

³⁰The decline could reflect a compositional effect (remaining public students are different in unobservable ways) or a peer effect (remaining public school peer groups are different in unobservable ways).

³¹As Hsieh and Urquiola (2002) observed, "when competition results in greater sorting, there is simply no instrument that would allow us to isolate the effect of choice on the public sector's productivity" (p. 15).

³²See Hanushek and Raymond (2002) for a discussion. Kupermintz (2001) analyzed the possibility of mean reversion in the Florida data. Kane and Staiger (2001) reported evidence of measurement error and mean reversion in North Carolina test score data. More recently, Chay, McEwan, and Urquiola (2003) found that mean-reverting measurement error may account for a large proportion of test-score gains among a group of low-achieving schools in Chile that were voucher-eligible.

was started in 1990, but a very small proportion of Milwaukee students actually participated in those early years (see Rouse, 1998, for an evaluation of program effects on voucher-users). The program was expanded in subsequent years—particularly after the 1998–1999 school year—and Hoxby argued that public schools experienced a much larger threat of losing students.

To evaluate the effects of competition, Hoxby (2003) identified a treatment group composed of schools that faced the greatest threat from competition. In these schools, at least two thirds of students were eligible to receive vouchers because they were also eligible for free or reduced-price lunch. She compared the achievement growth of these schools to a group of “less-treated” schools (in which fewer students were voucher-eligible) and “untreated” schools (in which no students were eligible). The untreated schools were necessarily drawn from Wisconsin schools outside of Milwaukee. Overall, Hoxby found that “treated” schools experienced more rapid growth in fourth-grade achievement than other groups, and she ascribed this to the causal effect of competition. She did not control for the changing characteristics of fourth graders over time, however, leaving open the possibility that student sorting affected student achievement. Ladd (2002) argued persuasively that the composition of treated schools was especially prone to change. Only the relatively poorer students were eligible to use vouchers, and they were likely to be relatively lower achieving than other students in the same school. Their departure would be expected to raise the mean achievement of treated schools, independently of any effects from competition.

Conclusions

Research that can inform the voucher debate has grown dramatically in the past 5 years. The question at hand is whether the research base can adequately predict the potential impact of voucher plans. McEwan (2000b) concluded that the evidence could warrant the implementation of small-scale voucher programs that are targeted at low-income, African American students in urban areas. That earlier conclusion stills holds, albeit with some caveats. The offer of vouchers may lead to small test-score improvements for some African Americans—judging from recent experimental evidence—but this depends on the city, the definition of race, and other analytical assumptions. Whatever the findings, it is certainly worth continuing these experiments for additional years. They provide the best available opportunity to assess whether the offer of vouchers produces ef-

fects on long-term outcomes, such as high school graduation, college entrance, and wages. Currently, there is only nonexperimental evidence on such outcomes, even though most would argue that they are ultimately more important than test scores.

McEwan (2000b) further concluded that early research was not helpful in forecasting the impact of a large-scale voucher plan. That conclusion is now only partly true. The mounting evidence on sorting suggests that *unrestricted* choice—as in large-scale open enrollment plans or Chile’s voucher system—can lead to cream-skimming. The best recent evidence on peer effects suggests that cream-skimming could lower the achievement of remaining public school students. This is not necessarily a concern if the achievement declines from sorting are outweighed by gains from competition. However, most of the literature on private school competition does not suggest that gains would be large. More alarmingly, it is not at all clear that research succeeds in identifying the causal effect of competition. As an alternative, Hsieh and Urquiola (2002) estimated the net effect of sorting and competition in Chile’s voucher system and found that it could be simply zero or very small.

The preceding conclusions are most relevant to unrestricted choice plans where flat-rate vouchers are offered to a large number of students with few eligibility restrictions. It is quite plausible that a voucher system could be designed that ameliorates some of the negative effects of sorting and maximizes the benefits of competition.³³ For example, Epple and Romano (2002) explored the potential impact of several different voucher schemes (including larger vouchers for some students and restricted add-on payments). Their work suggests that the voucher debate could fruitfully move toward issues of policy design and leave behind a simplistic debate of yea versus nay.

References

- Altonji, J. G., Elder, T. E., & Taber, C. R. (2002). *An evaluation of instrumental variables strategies for estimating the effects of Catholic schools* (Working Paper 9358). Cambridge, MA: National Bureau of Economic Research.
- Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized field trial. *American Economic Review*, 92, 1535–1558.
- Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3A, pp. 1277–1366). Amsterdam: Elsevier.

³³This point is shared by other reviewers, including Ladd (2002) and Neal (2002).

- Belfield, C. R., & Levin, H. M. (2002). The effects of competition between schools on educational outcomes: A review for the United States. *Review of Educational Research*, 72, 279–341.
- Bettinger, E. (1999). *The effect of charter schools on charter students and public schools*. Unpublished manuscript, Massachusetts Institute of Technology.
- Chay, K., McEwan, P. J., & Urquiola, M. (2003). *The central role of noise in evaluating interventions that use test scores to rank schools* (Working Paper 10118). Cambridge, MA: National Bureau of Economic Research.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Cullen, J. B., Jacob, B. A., & Levitt, S. D. (in press). The impact of school choice on student outcomes: An analysis of the Chicago public schools. *Journal of Public Economics*.
- Dee, T. S. (1998). Competition and the quality of public schools. *Economics of Education Review*, 17, 419–427.
- Dee, T. S. (2003). *The effects of Catholic schooling on civic participation*. Unpublished manuscript, Swarthmore College.
- Dee, T. S., & Fu, H. (2003). *Do charter schools skim students or drain resources?* Unpublished manuscript, Swarthmore College.
- Epple, D., Figlio, D., & Romano, R. (2004). Competition between private and public schools: Testing stratification and pricing predictions. *Journal of Public Economics*, 88, 1215–1245.
- Epple, D., & Romano, R. E. (1998). Competition between public and private schools, vouchers, and peer-group effects. *American Economic Review*, 88, 33–62.
- Epple, D., & Romano, R. (2002). *Educational vouchers and cream skimming* (Working Paper 9354). Cambridge, MA: National Bureau of Economic Research.
- Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, 100, 966–991.
- Evans, W. N., & Schwab, R. M. (1995). Finishing high school and starting college: Do Catholic schools make a difference? *Quarterly Journal of Economics*, 110, 941–974.
- Figlio, D., & Ludwig, J. (2000). *Sex, drugs, and Catholic schools: Private schooling and non-market adolescent behaviors* (Working Paper 7990). Cambridge, MA: National Bureau of Economic Research.
- Figlio, D. N., & Page, M. E. (2003). Can school choice and school accountability successfully co-exist? In C. M. Hoxby (Ed.), *The economics of school choice* (pp. 49–66). Chicago: University of Chicago Press.
- Figlio, D. N., & Stone, J. A. (1999). Are private schools really better? In J. Robst (Ed.), *Research in labor economics* (Vol. 18, pp. 115–140). Stamford, CT: JAI.
- Figlio, D. N., & Stone, J. A. (2001). Can public policy affect private school cream skimming? *Journal of Urban Economics*, 49, 240–266.
- Fiske, E. B., & Ladd, H. F. (2000). *When school compete: A cautionary tale*. Washington, DC: Brookings Institute.
- Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics*, 83, 257–268.
- Gill, B. P., Timpane, P. M., Ross, K. E., & Brewer, D. J. (2001). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools*. Santa Monica, CA: RAND Education.
- Greene, J. P. (2001). *An evaluation of the Florida A-Plus accountability and school choice program*. New York: Center for Civic Innovation, Manhattan Institute.
- Greene, J. P., Peterson, P. E., & Du, J. (1998). School choice in Milwaukee: A randomized experiment. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from school choice* (pp. 335–356). Washington, DC: Brookings Institute.

- Grogger, J., & Neal, D. (2000). Further evidence on the effects of Catholic secondary schooling. In W. G. Gale & J. R. Pack (Eds.), *Brookings-Wharton papers on urban affairs* (pp. 151-193). Washington, DC: Brookings Institute.
- Hanushek, E. A. (1997). Assessing the effect of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19, 141-164.
- Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18, 527-544.
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). *The impact of charter schools on academic achievement*. Unpublished manuscript.
- Hanushek, E. A., & Raymond, M. E. (2002). Improving educational quality: How best to evaluate our schools? In Y. K. Kodrzycki (Ed.), *Education in the 21st century: Meeting the challenges of a changing world* (pp. 193-224). Boston: Federal Reserve Bank of Boston.
- Hill, J. L., Rubin, D. B., & Thomas, N. (2000). The design of the New York School Choice Scholarship Program Evaluation. In L. Bickman (Ed.), *Donald Campbell's legacy*. Thousand Oaks, CA: Sage.
- Howell, W. G., & Peterson, P. E. (2002). *The education gap: Vouchers and urban schools*. Washington, DC: Brookings Institute.
- Hoxby, C. M. (2000). *Peer effects in the classroom: Learning from gender and race variation* (Working Paper 7867). Cambridge, MA: National Bureau of Economic Research.
- Hoxby, C. M. (2003). School choice and school productivity: Could school choice be a tide that lifts all boats? In C. M. Hoxby (Ed.), *The economics of school choice* (pp. 287-341). Chicago: University of Chicago Press.
- Hsieh, C.-T., & Urquiola, M. (2002). *When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program*. Unpublished manuscript, Princeton University and Columbia University.
- Jepsen, C. (2002). The role of aggregation in estimating the effects of private school competition on student achievement. *Journal of Urban Economics*, 52, 477-500.
- Kane, T. J., & Staiger, D. O. (2001). *Improving school accountability measures* (Working Paper 8156). Cambridge, MA: National Bureau of Economic Research.
- Krueger, A. B., & Zhu, P. (2002). *Another look at the New York City school voucher experiment* (Working Paper 9418). Cambridge, MA: National Bureau of Economic Research.
- Ladd, H. F. (2002). School vouchers: A critical view. *Journal of Economic Perspectives*, 16, 3-24.
- Ladd, H. F., & Fiske, E. B. (2001). The uneven playing field of school choice: Evidence from New Zealand. *Journal of Policy Analysis and Management*, 20, 43-63.
- Ladd, H. F., & Glennie, E. J. (2001). Appendix C: A replication of Jay Greene's voucher effect study using North Carolina data. In M. Carnoy (Ed.), *School vouchers: Examining the evidence*. Washington, DC: Economic Policy Institute.
- Levin, H. M. (1998). Educational vouchers: Effectiveness, choice, and costs. *Journal of Policy Analysis and Management*, 17, 373-391.
- Mayer, D. P., Peterson, P. E., Myers, D. E., Tuttle, C. C., & Howell, W. G. (2002). *School choice in New York City after three years: An evaluation of the School Choice Scholarships Program* (Rep. No. 8404-045). Princeton, NJ: Mathematica Policy Research.
- McEwan, P. J. (2000a). *The impact of vouchers on school efficiency: Empirical evidence from Chile*. Unpublished doctoral dissertation, Stanford University.
- McEwan, P. J. (2000b). The potential impact of large-scale voucher programs. *Review of Educational Research*, 70, 103-149.
- McEwan, P. J. (2001). The effectiveness of public, Catholic, and non-religious private schools in Chile's voucher system. *Education Economics*, 9, 103-128.

- McEwan, P. J. (2003). Peer effects on student achievement: Evidence from Chile. *Economics of Education Review*, 22, 131–141.
- McEwan, P. J., & Carnoy, M. (2000). The effectiveness and efficiency of private schools in Chile's voucher system. *Educational Evaluation and Policy Analysis*, 22, 213–239.
- McEwan, E. K., & McEwan, P. J. (2003). *Making sense of research*. Thousand Oaks, CA: Corwin.
- Mocan, H. N., Scafidi, B., & Tekin, E. (2002). *Catholic schools and bad behavior* (Working Paper 9172). Cambridge, MA: National Bureau of Economic Research.
- Moffitt, R. A. (2001). Policy interventions, low-level equilibria and social interactions. In S. Durlauf & H. P. Young (Eds.), *Social dynamics*. Cambridge, MA: MIT Press.
- Murnane, R. J., Newstead, S., & Olsen, R. J. (1985). Comparing public and private schools: The puzzling role of selection bias. *Journal of Business and Economic Statistics*, 3, 23–35.
- Myers, D. E., & Mayer, D. P. (2003). *Comments on "Another look at the New York City voucher experiment."* Retrieved July 20, 2003, from the World Wide Web: <http://www.mathematica-mpr.com/PDFs/anotherlook.pdf>
- National Center for Education Statistics. (2003). *Digest of education statistics 2002*. Washington, DC: Author.
- Neal, D. (1997). The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15, 98–123.
- Neal, D. (1998). What have we learned about the benefits of private schooling? *Federal Reserve Bank of New York Policy Review*, 4, 79–86.
- Neal, D. (2002). How vouchers could change the market for education. *Journal of Economic Perspectives*, 16, 25–44.
- Nechyba, T. J. (2000). Mobility, targeting, and private-school vouchers. *American Economic Review*, 90, 130–146.
- Peterson, P. E., & Howell, W. G. (2004). Efficiency, bias, and classification schemes: A response to Alan B. Kruger and Pei Zhu. *American Behavioral Scientist*, 47, 699–717.
- Plank, D. N., & Sykes, G. (Eds.). (2003). *Choosing choice: School choice in international perspective*. New York: Teachers College Press.
- Richard, A. (2003, April 9). Gov. Owens pledges to sign Colorado voucher bill. *Education Week*. Retrieved July 20, 2003, from the World Wide Web: <http://www.edweek.org/ew/ewstory.cfm?slug=30colo.h22>
- Robertson, D., & Symons, J. (2003). Do peer groups matter? Peer group versus schooling effects on academic attainment. *Economica*, 70, 31–53.
- Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. *Quarterly Journal of Economics*, 113, 553–602.
- Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, 116, 681–704.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Somers, M.-A., McEwan, P. J., & Willms, D. (2004). How effective are private schools in Latin America? *Comparative Education Review*, 48, 48–69.
- Walsh, M. (2000, November 15). Voucher initiatives defeated in Calif., Mich. *Education Week*. Retrieved July 20, 2003, from the World Wide Web: <http://www.edweek.org/ew/ewstory.cfm?slug=11vouch.h20>
- Walsh, M. (2002, October 16). Voucher advocates plan a multistate legal battle. *Education Week*. Retrieved July 20, 2003, from the World Wide Web: <http://www.edweek.org/ew/ewstory.cfm?slug=07institute.h22>
- Williams, G. C., & Zimmerman, D. J. (2003). *Peer effects in higher education* (Working Paper 9501). Cambridge, MA: National Bureau of Economic Research.

P. J. McEwan

- Witte, J. F. (1998). The Milwaukee voucher experiment. *Educational Evaluation and Policy Analysis*, 20, 229–251.
- Zimmer, R., Buddin, R., Chau, D., Gill, B., Guarino, C., Hamilton, L., et al. (2003). *Charter school operations and performance: Evidence from California*. Santa Monica, CA: RAND Education.
- Zimmer, R. W., & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, 19, 75–92.
- Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85, 9–23.