

14.48, Economics of Education  
Prof. Frank Levy  
**Lecture 15**

**Questions from last lecture:**

**Q:** How do you reconcile researches with policy agendas with good economic research when evaluating the studies we look at?

**A:** Well, that's what we're trying to do today in lecture

**Q:** How common it is for different sets of researchers to reach different conclusions with the same data?

**A:** When a line of research is new, there are often various and conflicting hypothesis, but over time, opinions usually converge as more research is done. In the case of school vouchers, you would ideally want to run many experiments in a variety of cities so you can verify whether they are effective under in a variety of places.

**Q:** In the Milwaukee experiment, is there a problem of selection bias?

**A:** There was an eligibility requirement: A family's income had to be less than 1.75 times the poverty line. Therefore, we can't say anything about how the voucher program might affect students from wealthier backgrounds.

**Q:** The average cost per student in public schools is much larger than the average cost per student in private schools. Why aren't school districts shifting towards privatization on their own to take advantage of these lower costs?

**A:** To scale privatization up nationally, you need to know that the cost structure can be scaled up too. A major portion of any school's costs is teacher salaries. Private (especially parochial) schools are generally paid less than public school teachers. There is no guarantee that there would be a large enough pool of teachers willing to work for lower wages in order to make the lower cost structure of private schools feasible on a larger scale.

**Q:** What do you do when you have multiple studies with different results?

**A:** We'll discuss this soon.

N= # of students in a public school system

M= # of students who apply for vouchers

X= # students awarded vouchers

Y= # of students who actually use vouchers

Z= # of students who use vouchers for more than 1 year

What is your treatment group? X. The treatment in this experiment is being awarded a voucher. The government cannot control whether or not a student uses the voucher. They can only award the vouchers.

What is your control group? M-X, the students who apply for vouchers but do not receive one. Why isn't our control group N-X? Because then we'd have a selection bias issue.

### Milwaukee experiment complications:

- 1) *Students applied to a specific school, not to the voucher program as a whole.* Oversubscribed schools help lotteries to allocate spots. Why does this do to our experiment design? We are trying to eliminate any systematic differences between lottery winners and losers that may affect achievement. We want to isolate the effect of voucher schools that might affect achievement. However, if the private schools are of different qualities, the probability of getting into the better schools will be lower than the lower performing schools because more students will apply to the better schools. The particular school you apply to affects your chance of getting a voucher, which affects the school, you go to which affects student achievement. Therefore, we have to control for which lottery a student applied to. There were effectively 3 private schools included in the experiment, one attended almost exclusively by Hispanic students and two attended by almost exclusively African American students so they researchers could extrapolate which schools a specific student applied as long as they knew his or her race.
- 2) *Attrition from the control group.* Ideally, the control group would continue to attend Milwaukee public schools. However, many lottery losers enrolled in suburban public schools or other private schools w/o a voucher. The more highly motivated students and families from the control group were seeking out other opportunities even if they lost the lottery resulting in their attrition from the control group because, once these students leave the Milwaukee public school system, we don't have data on them. The control group is whittled down to lottery losers with less motivated parents. Therefore, the analysis may be overestimating the effects of the voucher program.
- 3) *Lack of base comparison data.* When you look at a test score, you want to control for individual background. In this study, we'd want to use an earlier test score for comparison in addition for family fixed effects. The problem with using earlier test scores is that the data is usually unavailable.

### Parameters of 1996 Milwaukee experiment:

- 1) Income limited to less than \$21,000
- 2) Mean reported income of applicants was \$12,000
- 3) Mean reported income of families with children enrolled in Milwaukee public schools was \$43,000
- 4) Experiment couldn't involve more than 1% (later 1.5%) of Milwaukee public school students. This resulted in a small sample size, very little change in the composition of students in the Milwaukee public school students and little incentive for competition between Milwaukee public schools and private school

	<u>1990</u>	<u>1991</u>	<u>1992</u>	<u>1993</u>
<b>M=Number of applicants</b>	583	558	558	559
<b>X=Number of vouchers awarded</b>	376	452	321	395

3 studies use data from the Milwaukee voucher experiment

1) *Witte et al:*

These researchers thought the best approach was to take as the control group a set of public school students matched to the voucher students matched with various characteristics.

Results: Students awarded vouchers showed no significant improvements in test scores.

Problems: It's very difficult to control for/match parental involvement and ambition.

We don't want a random sample of Milwaukee public school students because the system wide average income is much higher than the average income of voucher applicants and recipients. Therefore, the treatment students should be matched with a control student based on income, pre-program achievement test scores and other controls. The matched control group isn't ideal but it is certainly better than nothing.

2) *Green et al:*

This study compares losers and winners in the same lottery.

Problems: They have a very small control group. Additionally, MPS records don't tell you what schools students applied to. As discussed earlier in the lecture, the researchers use ethnicity to infer what school students applied to.

Results: Vouchers are most effective after a student has been in the program for 3 years or more.

Math: Students attending voucher schools for 3 years scored 5 percentile points (.24 SD) above the control group. After 4 years, voucher students score 10.7 percentile points (.51 SD) above the control group.

Reading: Student attending voucher schools for 3 years gain 2-3 percentile points on reading achievement tests. Students attending voucher schools for 4 years gain 5.8 percentile points. These results are significant only when you take the 3rd

What's wrong with this approach?

The treatment group is those who used the vouchers for 3-4 years and NOT all those students awarded vouchers. Not every student who started in the voucher program continued using a voucher for 4 years. In fact, the program had a 50% attrition rate, and those who dropped out of the program are arguably systematically different from those who stayed. Those who dropped out were likely not doing well in their private school whereas those who did stay were doing well. The treatment group used by these researchers was actually just a subset of the true treatment group.

3) *Rouse:*

$$(TestScore)_{i,t} = \alpha_0 + \sum_{k=1}^4 \beta_k z_k + \delta_i (win_{t-1}) + (FixedEffects)_i + (EthnicControlsForLotteries)_i + \varepsilon_i$$

The data Rouse uses does not say whether students actually used the vouchers.

Results: On average, students awarded vouchers had better math scores than students not awarded vouchers, but there was no effect on reading scores. Winning the lottery added 1.5-2.3 percentile points (.32-.48SD) to math scores per year after winning a lottery.

When considering these results (and results from other policy analysis studies) we need to consider first whether the results are statistically significant. If they are, then we need to consider whether the effects are big enough to warrant government action. Without all the time, money and resources in the world, governments have to pick and chose which reforms to enact.