

Evaluation Tradecraft

Evaluation Tradecraft presents short articles about the art of evaluation in housing and urban research. Through this department of *Cityscape*, the Office of Policy Development and Research presents developments in the art of evaluation that might not be described in detail in published evaluations. Researchers often describe what they did and what their results were, but they might not give readers a step-by-step guide for implementing their methods. This department pulls back the curtain and shows readers exactly how program evaluation is done. If you have an idea for an article of about 3,000 words on a particular evaluation method or an interesting development in the art of evaluation, please send a one-paragraph abstract to marina.l.myhre@hud.gov.

Sensitivity of Treatment on Treated Effects in the Housing Vouchers Welfare Experiment to Alternative Measures of Compliance

Daniel Gubits
Abt Associates Inc.

Mark Shroder
U.S. Department of Housing and Urban Development

The views expressed in this article are those of the authors and do not necessarily represent the official positions or policies of the U.S. Department of Housing and Urban Development or the U.S. government.

Abstract

In the social experimentation literature, the treatment-control outcome difference is the “intention to treat” and the adjustment of that difference to reflect actual participation in treatment as the “treatment on treated” (TOT) effect of the intervention. Previous contributions to this literature have been silent on the sensitivity of TOT to alternative definitions of treatment.

In this article, we apply alternative methods of estimating treatment-on-the-treated to data from the Effects of Housing Vouchers on Welfare Families experiment. The final

Abstract (continued)

report on that experiment employs an original method of calculation of TOT and finds that early negative impacts on earnings fade out after 1.5 to 2 years. We test for sensitivity of these results to alternative concepts of participation: participation at time of measurement, exposure to treatment over time, and definition of the intervention as housing assistance per se rather than as vouchers.

We find that the published TOT results are qualitatively robust to the definition of treatment. We believe this finding is likely to apply more generally in large, well-controlled experiments.

Introduction

Controlled experiments to obtain unbiased estimates of the impacts of social interventions began in the 1960s in the United States and have since spread worldwide (Greenberg and Shroder, 2004). Because these are field (not laboratory) experiments, and because the subjects are human beings with rights, members of treatment groups do not always get the treatment that the researchers intended them to receive, and members of control groups may obtain access to the treatments that researchers intended to deny them or to equivalent treatments in the community.

The value of experimental findings may depend on the extent to which policymakers can accurately adjust for this noncompliance. The difference between treatment and control is an unbiased and consistent estimate of the impact of the difference in the regimes those two groups experience. The policymaker, however, is usually not interested in this difference per se but in the impact on those actually exposed to the treatment intervention for two reasons. First, when compliance with the experimental protocol is not universal, the treatment-control difference generally will understate the impact of the intervention. Second, when the treatment is costly to implement, the noncompliant members of the treatment group do not add to the costs, but the noncompliant members of the control group do. Thus noncompliance may distort the benefits and costs of the treatment.

Compliance and noncompliance, in general, are not random but reflect systematic differences in the personalities and backgrounds of the subjects. Exclusion of data from noncompliant subjects would destroy the integrity of the experiment; thus, some other method must be applied.

A convenient and well-accepted adjustment to experimental findings is called “treatment on treated” (TOT). The first rigorous articulation of TOT occurs in Bloom (1984). The full defense using the notation of Bayesian probability theory is in Imbens and Angrist (1994) and Imbens and Rubin (1997), who consider TOT (called by them the “local average treatment effect [LATE]”) the only valid measure of intervention impact that an experiment can supply.

The TOT/LATE concept has been highly influential. Citations to “Treatment on Treated” in Google Scholar as of June 2016 number 896; to “Local Average Treatment Effect,” 5,070. The concept has faced some controversy, centering on whether it is inherently meaningful (see, for example, Imbens

[2010] and the response by Deaton [2010]), but, to our knowledge, the literature on social experiments does not address the sensitivity of TOT to the definition of participation. This concern does not seem trivial. If the intervention is a training program, is it meaningful to include as participants those individuals who showed up on the first day and were never seen again? Must they actually graduate to qualify?

In the simplest case, let there be just one treatment and one control group, with no access whatever to the treatment by control group. A regression for the outcome of interest will have a binary variable for treatment on the right-hand side of the equation, and the estimated coefficient for this dummy is called the *intention to treat* (ITT), because the researchers intended to expose all members of the treatment group to the treatment. If, however, only a fraction p of the treatment group actually is exposed to it, and the impact on the fraction of noncompliers $(1-p)$ can be assumed to equal zero, then

$$\text{TOT} = \text{ITT}/p. \tag{1}$$

Thus, if, other things equal, a treatment raises earnings by an average of \$1,000 but only 20 percent of the treatment group actually received it, the value of the treatment to those actually exposed was, on average, \$5,000.

If some members of the control group *did* manage to obtain the treatment (or its equivalent) in the community, call the fraction of the group that these “crossovers” represent c . The natural assumption is that the impact of the treatment on the crossover controls will be the same as the impact on the compliant members of the treatment group. If so,

$$\text{TOT} = \text{ITT}/(p - c). \tag{2}$$

To continue the previous example, if 10 percent of the controls received the treatment despite not being assigned to it, the value of the treatment must average \$10,000.

To our knowledge, the literature on social experiments does not address the sensitivity of TOT to the definition of participation. We previously noted a concern about the minimum length of exposure to treatment, but the concern is broader. Suppose some other training program was available, and that some control group members enrolled in it. How closely must an alternative training program approximate the treatment curriculum for control group members enrolled in the alternative to be considered crossovers?

Some treatments continue during the course of the experiment. Are the effects of participation supposed to be static or dynamic? Suppose a mother receives a food supplement for her family in the second and fourth quarters of the year, but not in the third. Do we expect the feeding program to affect her behavior only in the second and fourth quarters or in the third quarter as well?¹

¹ We had some expectation that biometricians would have preceded us in this area, and perhaps would have offered a theoretical solution, because inconsistent compliance by patients in clinical trials could bias the evaluation of new drugs. Our hasty review of a vast literature indicates that the compliance problem is well known but that theorists tend not to accept the practical adjustments that practitioners make.

We investigate these issues using data from a large randomized experiment on the effects of portable housing assistance on the behavior of families with children receiving welfare.² A housing voucher will subsidize the rent of a low-income family in a privately owned decent and affordable unit, if the family can find such a unit in the community. The United States spends about \$20 billion a year on housing vouchers for nearly 2.2 million low-income households (HUD, 2016a), of whom about 46 percent are families with children (HUD, 2016b).

We focus on the findings involving earnings impacts. The impact of assistance on labor supply is theoretically ambiguous (see Shroder [2002] for a literature review). On the one hand, the subsidy formula taxes income, including earnings, at 30 cents on the dollar by increasing the family's required contribution to rent and, in principle, permits the recipient to lease a decent unit with no earnings at all. On the other hand, the voucher cuts the price of housing only and may stimulate demand for other goods and reduce the opportunity cost of job search. One important finding of the experiment, not further tested here, was that it essentially eliminated homelessness among the families that actually used the voucher, and this added stability might be expected to improve functioning in the job market.

Exhibit 1 extracts the critical information from the final evaluation of the experiment by Mills et al. (2006).³ During a 3-year period, the members of the control group had average earnings of \$485 a month (\$17,458 divided by 36 months). Adjusting for random differences between groups, the average treatment group member earned \$5 a month less. All the negative impact essentially occurred in the first 18 months after voucher eligibility (about \$17 a month), with positive but statistically insignificant impacts thereafter. Using one definition of compliance (discussed later in this article), the published analysis concludes that the TOT effect of using the voucher was about -\$17 per month overall, with TOT of about -\$33 per month in the first 18 months, subsequently fading to insignificance.

Exhibit 1

Impact of Housing Vouchers on Biannual and Total Earnings

Half-Year	Control Mean (\$)	ITT Impact (\$)	TOT Impact (\$)
1	2,651	- 124**	- 306**
2	2,837	- 100	- 174
3	2,889	- 76	- 195
4	3,007	16	- 20
5	3,029	30	- 32
6	3,046	72	103
Total	17,458	- 182	- 624
N			8,664

ITT = intention to treat. TOT = treatment on treated.

* p < .10. ** p < .05.

Source: Extracted from Mills et al. (2006: 100, exhibit 4.9); see our discussion at footnote 3

² The eligibility for the program technically extended to eligible nonrecipients of the Temporary Assistance for Needy Families (TANF) program. About 20 percent of the sample was not receiving TANF at baseline.

³ We have suppressed standard deviations and “all sites but Los Angeles” data from Exhibit 4.9 of Mills et al. (2006). Including Los Angeles adds about 1,000 subjects but subtracts 6 months of data. Non-Los Angeles ITT estimates are significantly negative at the .1 level in the second and third half-years, insignificantly negative in the fourth and fifth, and insignificantly positive in the sixth and seventh. In general, Los Angeles impacts by themselves are more positive than in other sites, but the differences are not statistically significant.

Although one would prefer having no negative earnings impacts whatever, these impacts are much smaller than those found for cash assistance in, for example, the Negative Income Tax experiments. In conjunction with negative impacts on homelessness and doubling up, they suggest that vouchers are an effective housing program but are not an effective antipoverty program.

These findings, however, provoke all the questions noted earlier. Some members of both treatment and control groups lived in public housing or Section 8 project-based assisted units before randomization, and some moved into such units thereafter. Public housing and Section 8 project units have the same 30 percent of income rent rule noted previously, but they are not portable—the tenant cannot take the assistance with her if she moves out. In calculating TOT, should one treat these project-based families as receiving an equivalent treatment? Some voucher holders left the program after first leasing up. Should impacts be adjusted to reflect the absence of subsequent participation?

Social experiments on the scale of the voucher demonstration cost millions of dollars to administer and are often simply too costly for agencies to fund, despite the major attractions of rigor and precision. We test whether any substantive inability to interpret their findings undermines these attractions.

In the remainder of this article, we describe the experimental data, define 10 alternative methods for estimating TOT in the context of this experiment, present the results of implementing these alternatives, and interpret the findings. We conclude that alternative adjustments to construct TOT do *not* substantively affect the findings from the experiment, and we think this conclusion will be applicable to the findings of most large, well-controlled randomized experiments.

Data

The data we use are from the Effects of Housing Vouchers on Welfare Families study (Mills et al., 2006), sponsored by the U.S. Department of Housing and Urban Development (HUD). In this study, families who either received or were eligible to receive Temporary Assistance for Needy Families (TANF) cash assistance were randomly assigned to either a treatment of an immediate offer of a housing choice (Section 8) voucher or a control group. Control group families were placed on housing authority waiting lists to receive a housing voucher. The study took place in six cities: Atlanta, Georgia; Augusta, Georgia; Fresno, California; Houston, Texas; Los Angeles, California; and Spokane, Washington. Data collected for the study included a baseline information form, unemployment insurance (UI) wage records, and HUD administrative records; they also included address history tracking, TANF and Food Stamp program administrative records, a followup survey, and qualitative indepth interviews. For this article, we make use of the baseline information, UI-derived earnings data, and housing subsidy receipt data derived from HUD records.

Not all treatment group members actually chose to accept the treatment offer. Roughly one-third of the treatment group families never leased an apartment using a housing voucher. In addition, about two-fifths of control group families eventually leased an apartment using a voucher. Mills et al. (2006) modeled the effects of using a voucher as cumulative over time and used an original method for estimating TOT impacts we call the “Orr method.”⁴ Mills et al. (2006) processed HUD

⁴ The “Orr method” is named for Larry Orr, its originator, who laid out the algebra and statistical properties of the model in appendix B of Mills et al. (2006).

administrative records to produce a series of period-by-period dummy variables that indicate whether a family had ever leased up with a housing voucher. The definition of treatment received as “ever leased up” implies that, if a family had ever leased up in period t , then it would also have “ever leased up” in all subsequent periods.

For this study, we used three other sets of dummy variables that describe the receipt of housing subsidies. In addition to (i) ever leased up with a housing choice voucher (HCV), we used (ii) leased up with a housing voucher in half-year t , (iii) ever received any housing subsidy (HCV, public housing, or project-based Section 8), and (iv) received any housing subsidy in half-year t .⁵

Dummy variable sets ii and iv take into account whether the housing voucher or housing subsidy has been relinquished. Therefore, in these sets, it is possible for the dummy variable for period $t + 1$ to equal zero if the dummy variable for period t equals 1.

We chose these alternative dummies out of consideration of the static labor supply model, which predicts a drop in labor supply due to the receipt of a voucher or housing subsidy (income effect) and due to the subsidy formula’s tax on earnings (substitution effect). If the effect of the voucher on earnings is solely through a labor supply disincentive effect, then the voucher’s effect should apply only in periods when the voucher is actually being received. Furthermore, other housing subsidies, such as public housing or project-based Section 8, with identical formulas to the HCV should have identical effects on earnings.⁶

Estimation

In this article, we present 10 sets of results for earnings impacts.⁷ We use two different estimation methods with each of the four sets of dummy variables that describe housing assistance receipt. Additional complexities with one of the estimation methods lead to 10 sets of results, rather than 8 sets. Exhibit 2 summarizes the overall estimation strategy.

The period-by-period Bloom (1984) method simply applies the Bloom adjustment ($1/(p - c)$) to the ITT in each period, using the appropriate definition of treatment to calculate treatment group participation (p) and control group participation (c).

The Orr method assumes that the time path of effects for treatment group participants and control group crossovers who initially receive the treatment in period 2 or after is *identical* to the time path of effects for non-crossover-like participants who initially receive the treatment in period 1. The Orr assumption parallels the assumption in the Bloom adjustment that the impact on crossovers is the same as the impact on compliant members of the treatment group.

⁵ Use of a voucher or housing subsidy for a period is defined as use of a voucher or housing subsidy on the first day of the period.

⁶ The authors thank Scott Susin for making the point that the high receipt of other housing subsidies in the Effects of Housing Vouchers on Welfare Families experiment would bias TOTs based solely on housing vouchers downward. Susin also made the point that the decrease of multiple program participation in the wake of welfare reform serves to increase the potential “bite” of the housing voucher marginal tax, suggesting a renewed focus on the prediction of the static labor supply model.

⁷ We have also computed 10 TOT estimates associated with the “all sites but Los Angeles” ITT estimates, but we do not present them here. The additional estimates are available from the authors. The conclusions we reach are not affected by this choice.

Exhibit 2

Static and Dynamic Estimation of TOT

Definition of Participation or Crossover in Period t	Estimation Method	
	Period-by-Period Bloom (static) Adjustment	Orr (dynamic) Adjustment
1.) Ever leased up (since RA) with a housing voucher	(1)	(5) (Method used in Mills et al., 2006)
2.) Uses a housing voucher in period t (takes account of voucher relinquishment)	(2)	(6) Calculates periods leased up as periods leased up since RA (7) Calculates periods leased up as periods leased up in current spell
3.) Ever received (since RA) a housing subsidy, including housing voucher, public housing, or project-based Section 8	(3)	(8)
4.) Receives housing subsidy (HCV, public housing, project-based Section 8) in period t (takes account of voucher relinquishment)	(4)	(9) Calculates periods of subsidy receipt as periods of receipt since RA (10) Calculates periods of subsidy receipt as periods of receipt in current spell

HCV = housing choice voucher. RA = random assignment. TOT = treatment on treated.

For example, if there were two periods, and subjects could be first exposed to treatment in either the first or the second period, under the Orr method,

$$ITT_1 = (p_{1,1} - c_{1,1})TOT_1,$$

$$ITT_2 = (p_{2,1} - c_{2,1})TOT_1 + (p_{2,2} - c_{2,2})TOT_2,$$

where $(p_{k,j} - c_{k,j})$ is the difference in the fraction of treatment and control groups who in period k of the experiment have been participating for j periods, and TOT_j is the treatment on treated effect for those who have been participating for exactly j periods. From standard linear algebra, it follows that the TOT vector is the product of the ITT vector and the inverse of the $(p_{k,j} - c_{k,j})$ matrix.⁸

An additional wrinkle here is that the adjustment requires information in each period on how many periods each family has been leased up (or received a subsidy). If cumulative effects of housing assistance since random assignment (RA) do not dissipate because of breaks in receipt, then counting the periods of receipt is appropriate. On the other hand, if cumulative effects of housing assistance since RA do dissipate because of breaks in receipt, then the salient length of receipt is the length of receipt in the current spell of housing assistance receipt.

⁸ The static Bloom adjustment can produce a sign reversal from ITT to TOT only if $(p - c)$ is negative (that is, controls get treatment more frequently than experimentals, which should not occur in well-controlled experiments). The Orr method, however, can produce a sign reversal simply because the TOT vector is the solution to a system of simultaneous equations (that is, a linear combination of the whole ITT vector). Further, the Bloom TOT estimate will always have the same t -statistic as the ITT estimate, because the standard error of the Bloom TOT is just the standard error of the ITT multiplied by the same scalar that transformed the ITT into the TOT. The Orr TOT in general will not have the same t -statistic as the corresponding ITT, however, because the standard error of the Orr TOT is a linear combination of the standard errors of the entire ITT vector.

The following varieties of TOT estimation are summarized in exhibit 2.

Bloom Estimates

(1) and (3):

Assumption in Method

The effect of having received a housing subsidy is not cumulative but may change over time since RA.

Assumption in Participation Definition

The effect occurs at initial lease up or receipt and impacts all outcomes thereafter.

(2) and (4):

Assumption in Method

The effect of having received a subsidy is not cumulative but may change over time since RA.

Assumption in Participation Definition

The effect occurs at lease up or receipt but disappears with relinquishment of the subsidy.

Orr Estimates

(5) and (8):

Assumption in Method

The effect of having received a subsidy is cumulative, and the time path of effects for treatment group participants and control group crossovers who initially receive the treatment in period 2 or after is *identical to* the time path of effects for non-crossover-like participants who initially receive the treatment in period 1.

Assumption in Participation Definition

The effect occurs at initial lease up or receipt and impacts all outcomes thereafter.

(6) and (9):

Assumption in Method

The effect of having received a subsidy is cumulative, and the effect for treatment group participants and control group crossovers who initially receive the subsidy in period 2 or after and who have held a subsidy for j periods is *identical to* the effect for non-crossover-like participants who initially receive the treatment in period 1, never relinquish the subsidy, and have held the subsidy for j periods.

Assumption in Participation Definition

The effect occurs at initial lease up or receipt but disappears during any subsequent periods of nonreceipt.

(7) and (10):

Assumption in Method

The effect of having received a subsidy is cumulative, and the effect for treatment group participants and control group crossovers whose current subsidy spell started in period 2 or after and whose current spell is j periods is *identical* to the effect for non-crossover-like participants who initially receive the treatment in period 1, never relinquish the subsidy, and have held the subsidy for j periods.

Assumption in Participation Definition

The effect occurs at initial lease up or receipt but disappears during any subsequent periods of nonreceipt.

Results and Discussion

Exhibit 3 presents the results of our calculations, and the final column summarizes the results.⁹ Our view of the results is that they show that the method of calculation simply does not matter very much. The maximum range of estimates for any particular half-year is in the second half-year, where the high and the low TOT estimates are, respectively, -\$40 and -\$61 per month but are not statistically different from zero. The policy implication would be precisely the same, using either the highest or the lowest estimate of the range.

During 3-year period, the 10 differences in estimation method yield cumulative TOT of between -\$458 and -\$770 (that is, between -\$13 and -\$21 per month). A policymaker would react to the highest number in substantially the same manner as to the lowest.

Why should the significant differences in TOT estimation methods result in such insignificant differences in their end products? We think the answer is in the source of the data: a large, well-controlled experiment. The size of the sample (7,622 or 8,664, depending on whether one excludes Los Angeles) greatly reduces the standard error of the ITT, and the variance-covariance matrix of the TOT in all versions is a linear transformation of the variance-covariance matrix of the ITT.¹⁰ We say the experiment was well controlled because, at all times, a substantial difference existed in participation in treatment, however defined, between experimentals and controls.

Exhibit 4 illustrates the latter point. It presents the half-year scalars constituting the four different Bloom adjustments. A substantial, relatively consistent gap always exists in the rate of participation between treatment and control groups. As a consequence, the different methods of estimation all produce reasonably similar results.

⁹ Column (5) of exhibit 3 uses the same TOT estimation method as Mills et al. (2006). A quick comparison of the point estimates in exhibit 1 with column (5) reveals a discrepancy between what is ostensibly the same set of results. This discrepancy comes from two sources: (1) the HUD administrative data used in this article were abstracted at a later time than the data used in Mills et al. (2006) and may contain corrections of records and (2) data cleaning decisions, which are consistent across the sets of subsidy receipt dummies in this article, may differ from those in Mills et al. (2006).

¹⁰ See Mills et al. (2006: B-9).

Exhibit 3

TOT Estimates for All Sites

	Control Mean ^a (\$)	ITT Impact (\$)	Period-by-Period Bloom (static)										Range of 10 TOTs										
			(1)		(2)		(3)		(4)		(5)			(6)		(7)		(8)		(9)		(10)	
			Voucher Ever	Voucher Period	Subsidy Ever	Subsidy Period	Subsidy Ever	Subsidy Period	HCV Ever	HCV Period	HCV Ever	HCV Period		HCV Ever	HCV Period	Subsidy Ever	Subsidy Period	Subsidy Ever	Subsidy Period	Subsidy Ever	Subsidy Period	Subsidy Ever	Subsidy Period
Half-year 1	2,651	-124**	-429**	-429**	-510**	-510**	-429**	-429**	-429**	-429**	-429**	-429**	-510**	-510**	-510**	-510**	-510**	-510**	-510**	-510**	-510**	-429 to -510	
	3,434	(59)	(205)	(205)	(243)	(243)	(205)	(205)	(205)	(205)	(205)	(205)	(243)	(243)	(243)	(243)	(243)	(243)	(243)	(243)	(243)	-242 to -368	
Half-year 2	2,837	-100	-279	-281	-368	-364	-242	-242	-242	-242	-242	-242	-364	-364	-350	-348	-345	-345	-345	-345	-345	-181 to -270	
	3,705	(67)	(187)	(188)	(246)	(243)	(202)	(202)	(202)	(202)	(202)	(202)	(243)	(243)	(257)	(256)	(255)	(255)	(255)	(255)	(255)	-181 to -270	
Half-year 3	2,889	-76	-204	-209	-270	-269	-184	-184	-184	-184	-184	-184	-269	-269	-251	-253	-250	-250	-250	-250	-250	-181 to -270	
	3,868	(73)	(195)	(200)	(258)	(257)	(215)	(215)	(215)	(215)	(215)	(215)	(257)	(257)	(273)	(272)	(273)	(273)	(273)	(273)	(273)	+45 to +121	
Half-year 4	3,007	16	45	49	60	62	95	95	95	95	95	95	62	62	87	99	87	87	87	87	87	+45 to +121	
	4,091	(80)	(223)	(243)	(297)	(307)	(241)	(241)	(241)	(241)	(241)	(241)	(307)	(307)	(307)	(311)	(317)	(317)	(317)	(317)	(317)	+37 to +135	
Half-year 5	3,029	30	90	105	121	135	37	37	37	37	37	37	135	135	59	71	52	52	52	52	52	+37 to +135	
	4,225	(83)	(249)	(292)	(334)	(372)	(251)	(251)	(251)	(251)	(251)	(251)	(372)	(372)	(322)	(332)	(352)	(352)	(352)	(352)	(352)	+156 to +393	
Half-year 6	3,046	72	240	307	319	393	156	156	156	156	156	156	393	393	196	233	211	211	211	211	211	+156 to +393	
	4,268	(86)	(287)	(366)	(380)	(468)	(267)	(267)	(267)	(267)	(267)	(267)	(468)	(468)	(343)	(364)	(397)	(397)	(397)	(397)	(397)	-458 to -770	
Total, all half-years	17,458	-182	-538	-458	-649	-553	-567	-567	-567	-567	-567	-567	-553	-553	-770	-707	-754	-754	-754	-754	-754	-458 to -770	
	20,359	(365)					(1,085)	(1,085)	(1,085)	(1,085)	(1,085)	(1,085)											

HCV = housing choice voucher. ITT = intention to treat. RA = random assignment. TOT = treatment on treated.

N = 8,664. ** p < .05.

^a The standard deviation of the control group mean for each half-year appears below it in this column.

Exhibit 4

$(\rho - c)$ for Four Bloom Adjustments

Half-Year	1 Voucher Ever	2 Voucher Current	3 Subsidy Ever	4 Subsidy Current	Range of Adjustments
1	0.288	0.288	0.242	0.242	.242–.288
2	0.359	0.358	0.273	0.276	.273–.359
3	0.373	0.364	0.282	0.284	.282–.373
4	0.361	0.330	0.270	0.261	.261–.361
5	0.333	0.284	0.248	0.222	.222–.333
6	0.300	0.235	0.226	0.183	.183–.300

The $(p_{k,j} - c_{k,j})$ matrix in the Orr adjustments is not reproduced to save space. All entries to the northeast of the diagonal are zeros, the entries representing the first two periods after RA are large positive proper fractions, and many of the remaining entries are small negative proper fractions, because the controls begin to gain access to subsidy at a faster rate than the experimentals. These small negative values result in small effective weights on the corresponding values of the ITT covariance matrix. This result merely confirms the good judgment of the researchers (and HUD) in the selection of sites and the oversight of the demonstration: control group members at these sites had relatively poor access to housing subsidy throughout the period of the demonstration, and data collection did not extend beyond the period where catchup began to threaten the value of the evidence.

Thus, in this experiment we find that TOT estimates are relatively insensitive to a wide range of differences in assumptions. We believe that this conclusion ought to be generally applicable to most large, well-controlled experiments. As long as the control group will not be exposed to the treatment or to a close substitute for the treatment, reasonable alternative estimates of TOT should not exhibit substantively significant deviations in a large sample.

Acknowledgments

The authors thank Lydia Taghavi and Paul Dornan for facilitating their use of experimental and administrative data and Phil Gleason and Steve Bell for helpful comments.

Authors

Daniel Gubits is a principal associate at Abt Associates Inc.

Mark Shroder is Associate Deputy Assistant Secretary for Research, Evaluation, and Monitoring in the Office of Policy Development and Research at the U.S. Department of Housing and Urban Development.

References

- Bloom, Howard S. 1984. "Accounting for No-Shows in Experimental Evaluation Designs," *Evaluation Review* 8 (April): 225–246.
- Deaton, Angus. 2010. "Instruments, Randomization, and Learning About Development," *Journal of Economic Literature* 48: 424–455.
- Greenberg, David, and Mark Shroder. 2004. *The Digest of Social Experiments*, 3rd ed. Washington, DC: Urban Institute Press.
- Imbens, Guido, and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62 (2): 467–475.
- Imbens, Guido, and Donald Rubin. 1997. "Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance," *Annals of Statistics* 25 (1): 305–327.
- Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature* 48: 399–423.
- Mills, Gregory, Daniel Gubits, Larry Orr, David Long, Judie Feins, Bulbul Kaul, Michelle Wood, Amy Jones & Associates, Cloudburst Consulting, and the QED Group. 2006. *Effects of Housing Vouchers on Welfare Families*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.
- Shroder, Mark. 2002. "Does Housing Assistance Perversely Affect Self-Sufficiency? A Review Essay," *Journal of Housing Economics* 11: 381–417.
- U.S. Department of Housing and Urban Development. 2016a. "Budget Authority by Program, Comparative Summary, Fiscal Years 2015–2017." http://portal.hud.gov/hudportal/documents/huddoc?id=1.1-Budget_Authority.pdf.
- . 2016b. "Picture of Subsidized Households." huduser.gov/portal/datasets/picture/yearly-data.html.