Social Experiments in Housing

Mark Shroder
U.S. Department of Housing and Urban Development
Office of Policy Development and Research

Abstract

In the past quarter century, HUD has funded eight evaluations of social programs in the housing field using random assignment of households or individuals to treatment and control groups. The existence and results of these rigorous studies often are not well known. The object of this paper is to describe the completed demonstrations, to outline the scope and objectives of those still under way, to discuss unfinished business—important issues that experiments might address—and to summarize the power and limits of the experimental method.

The paper summarizes completed demonstrations along the following dimensions: timeframe, treatments tested, outcomes of interest, sample size, target population, number of treatment groups, number and location of sites, major findings, design issues, replicability, generalizability, information sources, and whether there is public access to the data. Reports on the ongoing experiments do not discuss findings, design issues, replicability, or generalizability.

Both the power and the limits of the experimental methodology in its application to housing issues are explored. Experimental methods allow identification of the response of individuals to program changes, holding markets constant. They do not, however, capture all the impacts on individual behavior when the program changes significantly affect the operations of the markets themselves.

Undergraduates used to be taught that social science hypotheses could not be tested experimentally. This seemingly reasonable proposition turns out to be false. Social science hypotheses have been extensively tested both in the laboratory and in the field. Field tests usually attempt to measure the effects of particular social programs, which in turn embody, not always explicitly, hypotheses about human behavior. In this paper I report on tests of the efficacy of housing-related programs that meet a restrictive definition of *social experiments.*

Social experiments are field studies of social programs in which individuals, households, or (in rare instances) firms or organizations are randomly assigned to two or more alternative treatments. The primary research objective of social experiments is the measurement

of impacts from the alternative treatments on market behavior and corresponding government fiscal outcomes. A social experiment has at least the following features:

- **Random assignment.** The creation of at least two groups of human subjects who differ from one another by chance alone.
- **Policy intervention.** A set of actions ensuring that different incentives, opportunities, or constraints confront the members of each of the randomly assigned groups in their daily lives.
- **Followup data collection.** The measurement of market and fiscal outcomes for the members of each group.
- **Evaluation.** The application of statistical inference and informed professional judgment about the degree to which the policy interventions had impacts—that is, caused differences in outcomes between the groups.

Random assignment of subjects, when effectively implemented, serves to eliminate selection bias that would usually be found when a program is evaluated. Individuals who participate voluntarily in programs generally differ both in observed and in unobserved ways from individuals who do not.[1]

Greenberg and Shroder (1997) summarize 143 completed social experiments and 74 that were still ongoing in late 1996. This paper summarizes 5 completed and 3 ongoing experiments funded by the U.S. Department of Housing and Urban Development over the past 25 years. The existence and results of these rigorous studies often are not well known. The object of this paper is to describe the completed demonstrations and to outline the scope and objectives of those still under way.

These demonstrations were designed to answer three fundamental questions:

- Can housing subsidies help very low-income renters find adequate housing in the private market, and if so, how should such subsidies best be structured?
- Is financial counseling a cost-effective intervention for homeowners who have defaulted on FHA-insured mortgages, and can prepurchase counseling reduce default risk?
- Can the Federal Government effectively respond to one of the most serious imbalances in the U.S. economy—the concentration of poverty in the central city and the concentration of employment and educational opportunity in the outlying suburbs?

This paper does not report original research on these subjects. My goal simply is to make information about these studies a little more accessible. Half of the studies are almost unknown, even in housing circles, and the other four are not in any danger of overexposure. The paper provides capsule synopses of completed demonstrations along the following dimensions: timeframe, treatments tested, outcomes of interest, sample size, target population, number of treatment groups, number and location of sites, major findings, design issues, replicability, generalizability, information sources, and whether there is public access to the data. Unfinished business remains in each of these issue areas, and some conceptually feasible experiments are discussed.

Both the power and the limits of the experimental methodology in its application to housing issues are explored in the final section. Experimental methods allow identification of the response of individuals to program changes, holding markets constant. They do not, however, capture all the impacts on individual behavior when program changes

significantly affect the operations of the markets themselves. However, social experiments in housing have increased the stock of useful knowledge and substantial increments could be obtained from this methodology in the future.

Rent Subsidy Experiments

In the three experiments described below, an integral part of the experiment is the provision of subsidy to a low-income person or household to be used as partial payment of the rent required by the owner of a private unit selected by the tenant. All three were specifically authorized by acts of Congress. The first two have an important place in the history of housing policy. They were shaped by—and in turn helped to shape—the continuing controversy over the form in which housing assistance should be provided.

Housing Allowance Demand Experiment

Synopsis. This demonstration, conducted from 1973–77, tested the effects of two forms of housing subsidy on a large sample of low-income renter households. Subjects were followed for 2 years.

Timeframe. Data collected April 1973 to February 1976. Final report issued June 1980.

Treatments Tested. The principal treatments tested were payments to households based on a "housing gap" and payments based on a percentage of the rent.[2]

For the housing gap treatment, a panel of experts at each site estimated the cost of housing meeting certain standards in modest neighborhoods in that city. This number was C^*. Payment (P) was based on the formula $P=dC^*–bY$, where Y was disposable income less $300 per year for each working member of the family, and d and b were experimental parameters (higher d and lower b imply a greater generosity). Treatments also varied in housing requirements: a minimum rent requirement, set at .7 or .9 of C^*, or a minimum standards requirement, under which occupied units would be inspected for conformity with standards for health and safety. Households living in units that did not meet the standard specified for the treatment to which they were assigned could not receive payments.

Exhibit 1

Housing Gap Payments

Treatment	*d*	*b*	Housing Requirement
1	1	.15	Minimum Standards
2	1.2	.25	Minimum Standards
3	1	.25	Minimum Standards
4	0.8	.25	Minimum Standards
5	1	.35	Minimum Standards
6	1.2	.25	Minimum rent (.7)
7	1	.25	Minimum rent (.7)
8	0.8	.25	Minimum rent (.7)
9	1.2	.25	Minimum rent (.9)
10	1	.25	Minimum rent (.9)
11	0.8	.25	Minimum rent (.9)
12	1	.25	No requirements

In the percentage-of-rent treatments, payment (P) was determined by the formula $P=aR$, where R is rent and a is a program parameter. There were no housing requirements.

Exhibit 2

Percentage-of-Rent Treatments

Treatment	*a*
13	.6
14	.5
15	.4
16	.3
17	.2

Controls were paid $10 a month for filling out a monthly form and $25 for periodic interviews. Experimentals were not paid for interviews. The payment system for experimentals lasted 3 years. Information was collected for 3 years.

Target Population. Renter households residing in the counties of the experiment and meeting the following tests: (1) Disposable income (minus $300 per year per worker) less than one-quarter of the *C** figure for households of that size in that city; (2) assets of under $5,000 (under $10,000 if 62 or older); (3) two or more related persons of any age or with household head who was handicapped, disabled, 62 or older, or displaced by an urban renewal project; (4) resident in unsubsidized housing—public housing tenants were only eligible if they moved.

Outcomes of Interest. (1) Enrollment. (2) Rate of participation (actual receipt of payment). (3) Effects on housing expenditures, quality, and residential segregation.

Sample Size. The 3,600 households (1,800 in each city) were invited to enroll either as experimentals or as controls. The numbers actually enrolling (not necessarily receiving payments) were recorded for the initial enrollment offer and 2 years later. The difference between the second and third columns is attrition, which in this case was of independent interest.

Number of Treatment Groups. 19 (with two control groups).

Number and Location of Sites. Pittsburgh, Pennsylvania, and Phoenix, Arizona.

Research Components.

- Process analysis: Obtained reasons for refusal to enroll, choice not to participate, and condition of initial housing units.[3]
- Impact analysis: By comparison of means, regression, and discrete-response techniques.
- Benefit-cost analysis: A relative cost-effectiveness study was conducted, comparing costs of housing allowances and public subsidies with housing construction.

Major Findings

- Process analysis: Many families refused to enroll, a finding that is important in estimating the costs of a national program. Refusal to enroll did not seem to be related to the variables in the experiment, but to a lack of interest in receiving public assistance. Many of those who did enroll did not participate (receive a payment), and this

Exhibit 3

Sample Size

Treatment	Initial	2 Years
1	212	181
2	91	63
3	133	77
4	128	82
5	137	75
6	85	58
7	132	89
8	124	79
9	88	60
10	145	88
11	137	78
12	145	103
13	66	49
14	235	190
15	265	179
16	258	176
17	176	111
Controls	950	603

A different way of presenting the second-year numbers is as follows:

Exhibit 4

Second-Year Numbers

Treatment Type	Phoenix	Pittsburgh	Total
Housing Gap (minimum standards)	174	204	378
Housing Gap (minimum rent)	207	245	452
Housing Gap (no requirements)	40	63	103
Percentage of Rent	298	407	705
Controls	282	321	603
Total	**1,001**	**1,240**	**2,241**

non-participation was substantially affected by the stringency of the housing requirements, the household's race, and the relative availability of housing meeting the minimum requirements (which varied between the sites). Participation is stated as the percentage of those enrolling (all of whom were eligible for immediate payments on income grounds) who received one or more payments.

Higher payments increased participation. At an average monthly payment level of $43, one-fourth of all renters who had to meet housing requirements participated; at twice that level, twice as many participated.

- Impact analysis: To date no satisfactory explanation for the divergence (in sites) shown in Exhibit 6 has been found.
- Impact analysis: The impact on housing expenditures in the Housing Gap treatments differs according to whether the household initially occupied a unit satisfying the

Exhibit 5

Percentage of Those Enrolling That Received One or More Payments

Treatment/Site	Enrollment	Participation
No housing requirements		
Percent of rent		
Pittsburgh	82%	100%
Phoenix	87	100
Housing gap		
Pittsburgh	78	100
Phoenix	90	100
Minimum standards*		
Pittsburgh	75	40
Phoenix	84	54
Minimum rent (.7)		
Pittsburgh	74	81
Phoenix	82	74
Minimum rent (.9)		
Pittsburgh	73	58
Phoenix	81	54

*This is the treatment that best approximates the current Section 8 tenant-based assistance program, although a number of features differ.

Exhibit 6

Estimated Experimental Effect on Housing Expenditures and Services Among Households Meeting Requirements 2 Years After Enrollment

Site/Treatment	Change in Expenditures	As Share of Payment	Change in Services
Pittsburgh			
No restriction	2.6%	5.7%	3.4%
Percentage of rent	8.0	14.0	3.0
Minimum rent (.7)	–3.6	–7.8	0
Minimum rent (.9)	8.5*	23.3	0.9
Minimum standards	4.3	8.6	3.1
Phoenix			
No restriction	16.0*	19.0	12.6*
Percentage of rent	8.0	23.7	–1.0
Minimum rent (.7)	15.7	25.5	11.0*
Minimum rent (.9)	28.4*	41.3	18.9*
Minimum standards	16.2*	27.4	10.2*

Services are measured with a hedonic index based on characteristics of the housing unit. An asterisk (*) indicates a difference that is statistically different from zero.

Exhibit 7

Estimated Experimental Effects

Site/Treatment	Change in Expenditures	As Share of Payment	Change in Services
Pittsburgh			
Satisfactory			
Minimum rent (.7)	2.4%	(5.7%)	0.5%
Minimum rent (.9)	4.6	(13.7)	–0.7
Minimum standards	1.1	(2.3)	0.8
Unsatisfactory			
Minimum rent (.7)	8.7	(15.4)	–0.9
Minimum rent (.9)	15.8*	(38.8)	3.1
Minimum standards	7.5*	(14.2)	5.6
Phoenix			
Satisfactory			
Minimum rent (.7)	–1.2	(2.7)	2.5
Minimum rent (.9)	7.4	(15.4)	4.2
Minimum standards	–0.7	(2.1)	8.2*
Unsatisfactory			
Minimum rent (.7)	42.0*	(41.7)	20.2*
Minimum rent (.9)	42.6*	(50.0)	26.0*
Minimum standards	23.6*	(32.8)	10.5*

An asterisk denotes statistically significant difference from zero.

minimum standards at enrollment. Of those whose units did not, most that participated satisfied the requirements by moving. The numbers of those who moved and the distances they moved were such that the impact on residential segregation would have been negligible.

- Impact analysis: Estimates of income elasticity range from .29 to .34 in Pittsburgh and from .26 to .44 in Phoenix.[4] Estimates of price elasticity range from –.11 to –.18 in Pittsburgh, –.23 to –.24 in Phoenix. Differences come in part from econometric specification, in part from differences in the definition of income.

Design Issues.

- Payments guaranteed over 3 years may not induce the same behavioral changes as a permanent program. For instance, a family that would need to move to receive payments might also realize that a second move would be necessary at the close of the experiment, because they could not afford the unit the experiment had subsidized. Most households change their housing units infrequently, and the effect of the experiment will occur with some lag. However, the timing of the lag is not known with certainty and requires modeling assumptions.
- The measure of housing services, a hedonic quality index, was developed for this experiment and is not completely satisfactory.
- The least generous plans were assigned only to very low-income members of the sample, because otherwise many enrolled persons would have been eligible only for

very small payments or none at all, limiting their benefits from participating in the experiment. However, this tends to confound the treatment effect with the characteristics of households assigned to the treatment.

- "... the price elasticity per se is unlikely to be of much use in designing a housing-allowance program. A percent-of-rent formula offers such attractive opportunities for mutually beneficial fraud on the part of landlords and renters that (it) is hard to imagine it ever being implemented." (Harvey Rosen, in Bradbury and Downs, 1981.)

Replicability. The processes used for administration of the program were fully documented.

Generalizability. No one has ever claimed that Phoenix and Pittsburgh, between them, represent the United States. With appropriate qualification, HUD and other analysts drew the following conclusions:

1. The single most important finding of this experiment is the extremely low-income elasticity of housing demand among low-income people. One implication of this finding is that housing allowances would not result in large inflation of rents. Another implication is that in the objectives of a housing-allowance program, there is a tradeoff between assisting large numbers of people and improving the quality of the existing housing stock.

2. Results of this experiment and the other components of the EHAP were used by the Urban Institute to project total costs of housing-allowance programs using microsimulation techniques.

Funding Sources Other Than HUD. None.

Evaluators. Abt Associates. Key personnel: Stephen Kennedy and James Wallace.

Information Sources. Kennedy (1980); HUD-PD&R (1979, 1980); Struyk and Bendick, Jr. (1981); Bradbury and Downs (1981). Abt Associates also can make available a large assortment of unpublished reports.

Public-Use Access to Data. No public-use file exists.[5]

Freestanding Housing Voucher Demonstration

Synopsis. This demonstration, conducted from 1985 to 1988, tested the effects of housing vouchers with an experimental payment formula on a large sample of low-income families. Subjects were followed for 1 year.

Timeframe. April 1985–September 1988. Final report, May 1990.

Treatments Tested. 1. Certificate program (controls). This was the preexisting program. It pays a monthly stipend to the landlord on behalf of a tenant living in privately owned, existing housing. The amount of the payment is the difference between the rent (plus certain scheduled utility allowances, if they are not included in the rent) and the tenant's contribution, which is essentially 30 percent of adjusted income. Tenants must live in a unit meeting HUD's housing quality criteria, and the rent must be less than or equal to the local fair market rent (FMR) (set by HUD) and judged "reasonable" by the public housing agency (PHA). From the time of enrollment into the program, tenants have 2 to 4 months to find acceptable housing under the program.

2. Housing voucher program (experimentals). This treatment differs from the control program in the payment formula; the housing unit must still meet HUD quality criteria. The housing assistance payment is equal to $P - .3Y$, where P is the local rental payment standard, initially set equal to the FMR,[6] and Y is income. Thus, the PHA no longer sets a ceiling on gross rent. The tenant has, on the one hand, an incentive to obtain housing at a lower cost than FMR, if it can be found, and, on the other hand, the option to secure housing that costs more than the PHA would allow under the certificate program.

Target Population. Lower income families certified as eligible for Section 8 who live in large urban areas.

Outcomes of Interest.

- Success rate (percentage of those enrolled that find acceptable units and become recipients).
- Rent payments.
- Rent burdens.
- Program payments.
- Administrative costs.

Sample Size. 12,390, evenly divided between certificates and vouchers. Many of the analyses, however, use subsamples of about 4,500.

Number of Treatment Groups. Two (with one control group).

Number and Location of Sites. Sixteen—Atlanta, Georgia; Boston, Massachusetts; Buffalo, New York; Cleveland, Ohio; Dayton, Ohio; Minneapolis, Minnesota; Montgomery County, Maryland; New Haven, Connecticut, New York City, New York; Oakland, California; Omaha, Nebraska; Pittsburgh, Pennsylvania; St. Petersburg, Florida; San Antonio, Texas; San Diego, California; and Seattle, Washington.

Research Components.

- Process analysis: Not conducted.
- Impact analysis: Comparison of means and OLS.
- Benefit-cost analysis: Not conducted.

Regressions on housing quality appear to show that roughly half of the higher rent payments under the voucher plan go to improved housing quality, with the other half going to higher landlord income. A major finding is that both certificate and voucher holders tended to look for housing in a "Section 8 submarket" in which landlords were much more constrained by the monopsony of the PHA in the certificate program than by the tenant's self-interest in keeping the rent down in the voucher program.[7]

Design Issues. The most obvious problem is the absence of a process analysis. The voucher program changes the budget constraint of the PHA as well as the subjects, and the absence of a process analysis means that we do not know how PHAs responded or whether their responses affect experimental results.

Replicability. High. HUD continues to administer both programs.

Exhibit 8

Major Findings

	Voucher	Certificate
Success rate, overall	64.6%	61.0%*
Success rate when *P*=FMR[a]	64.4%	59.5%*
Total rent paid by recipients	$463	$437*
Rent burden as percentage of income		
At initial payment	34%	31%
At annual recertification	35	31*
Among recipients who did not move	28	31**
Among recipients who moved	39	31
Monthly assistance payments, overall average	$310	$293**
Initial assistance payment	307	287**
Payment at recertification	304	298
Administrative cost per slot		
Initial eligibility	$579	$598
Annual ongoing	257	261

* Difference significant at .05 percent level.
** Difference significant at .01 percent level.
[a] PHAs had some discretion about changing *P*. See note 7.

Generalizability. Designed for generalizability to the population of large urban PHAs. Two special caveats: (1) the sample is drawn from applicants for the current Section 8 program, and (2) more important, many experimentals were renting from landlords with substantial Section 8 experience. If the entire program changed over to vouchers, landlord rentsetting behavior might change as well.

Funding Sources other than HUD. None.

Evaluators. Abt Associates. Key personnel: Mireille L. Leger and Stephen D. Kennedy.

Information Sources. Leger and Kennedy (1990).

Public-Use Access to Data. No public-use file exists.[8]

HUD/VA Supported Housing Program

Synopsis. This demonstration, conducted from 1991 to 1998, tests the effects of the combination of housing vouchers and intensive case management on homeless veterans with serious psychiatric or substance abuse disorders. Subjects will be followed for 3 years.

Treatments Tested.

- Controls received standard VA (Department of Veterans Affairs) treatment (i.e., referral to outpatient clinic and treatment there on an appointment basis).

- The first experimental group received intensive case management (i.e., a caseworker regularly visited the subject and tried to help him solve problems of daily life, housing among them).
- The second experimental group received intensive case management and housing vouchers.

Target Population. Homeless veterans with serious psychiatric or substance abuse disorders.

Outcomes of Interest. Housing (number of days in housing—that is, where participants lived for the previous 90 days), mental health status, employment, health service utilization, income, substance abuse, and satisfaction with services.

Number of Treatment Groups. Three, including one control.

Funding Sources Other Than HUD. The VA paid for medical treatment and case management, and for the evaluation.

Evaluators. Robert Rosenheck, VA Medical Center, West Haven, CT, and Linda Frisman, Connecticut Department of Mental Health.

Unfinished Business. With hindsight, one can see extraordinary missed opportunities in the Housing Allowance Demand Experiment. The variation in the generosity of the basic benefit and in the tax rate (parameters *d* and *b*, respectively) would have permitted estimation of not only the effects of the subsidy on housing demand but also on the labor supply and cash-transfer dependence of the households. These are contentious issues that arise continually in policy discussion, and it should have been possible to answer them with the data from this experiment.[9] This experiment cost $31.6 million in mid-1970s dollars and could have significantly contributed to subsequent policy.

The Welfare to Work Voucher demonstration funded in FY 1999 has the potential to fill this gap. HUD has announced its intention to use random assignment of vouchers to families that are receiving, are eligible for, or have received Temporary Assistance to Needy Families in the past 2 years to determine whether housing assistance can help such families obtain and/or retain employment. The tentatively selected sites are Augusta and Atlanta, Georgia; Fresno City and County, California; Houston, Texas; Los Angles City and County, California; Spokane, Washington; and Springfield and metropolitan Boston, Massachusetts.

Homeownership Counseling Experiments

The Clinton administration and a variety of private-sector actors have entered into a national partnership to raise the rate of homeownership in the United States. This will entail outreach to households that do not presently own homes. Some of these households might previously have been refused credit because underwriters considered them too risky. Others might have shunned the market because, rightly or wrongly, they feared the risks inherent in property ownership. These are not new issues. For several decades financial counseling has been a leading strategy in reducing the risks associated with lending to a specific household. The issue is whether this is an efficient strategy.

The three experiments reported here are not well known.[10] All were initiated during the late 1970s when the Federal Housing Administration (FHA) faced a rising delinquency rate on home loans it had insured. It guaranteed these loans under one of three programs:

- Section 203, the standard FHA loan program, guarantees the mortgage up to 97 percent of the property value. The 203 premium is priced on an "acceptable risk" basis and is not supposed to result in an ex ante subsidy from the government.
- Section 221(d)(2), for loans in blighted urban areas, required a downpayment of 3 percent of the property value, but the mortgage itself could be insured for up to 100 percent of the property value. The premium in this program was subsidized ex ante.
- Section 235, for low- to moderate-income homebuyers, required a downpayment of $200. Both the premium and the interest rate on the loan were subsidized; the interest rate to the buyer was 1 percent.

Losses from default in the 1970s had cost all three programs far more than their proponents envisioned. Congress had responded with programs to counsel homeowners, but the effectiveness of these programs was not clear.

Counseling for Delinquent Mortgagors

Synopsis. This demonstration, conducted in 1974, tested the effects of counseling on a large sample of defaulting or delinquent Section 235 mortgagors. Subjects were followed for 6 to 12 months.

Timeframe. The demonstration was conducted from July 1974 to January 1975. Analysis was conducted of loan status as of July 1975.

Treatments Tested. Controls received no counseling, and were subject to the usual procedures for Section 235 mortgagors who fell behind on their payments. Experimentals received an offer of financial counseling; the fact that a household was being counseled might or might not affect lender forbearance.

Target Population. Section 235 mortgagors in delinquency or default.

Outcomes of Interest. Mortgage repayment.

Sample Size. 705 controls, 605 experimentals.

Number of Treatment Groups. Two groups (one control).

Number and Location of Sites. Five—Atlanta, Georgia; Columbia, South Carolina; Detroit, Michigan; Los Angeles, California; and Seattle, Washington.

Research Components. Limited process analysis. Impact analysis by comparison of unadjusted means. Benefit-cost analysis conducted.

Major Findings.

- Process analysis: There were 10 agencies in the 5 sites. They differed in emphasis on home versus office counseling, debt collection services (offered to build support from creditors), and interest in referral to other social services.
- Process analysis: Only one-fourth of those referred were actually counseled. Almost one-fourth of those referred could not be contacted, and more than half (334) declined to make counseling appointments, although 184 of the latter went through a precounseling interview. Those who declined counseling, as a group, had a lower foreclosure rate than the controls.

- Impact analysis: As of July 1975, financial counseling had a significant positive impact on loan repayment by experimentals.
- In-home counseling appeared substantially more effective than in-office counseling.
- Cost-effectiveness: Counseling has net benefits of $174.82 to $440.32 (1974 dollars) per case, depending on the benefit-cost model chosen. The higher figure reflects the net *social* benefits; costs to the taxpayer of continued Section 235 subsidies (which is a transfer from one party to another, not a social cost) and to the government of foregone tax revenues from debt forgiveness as a consequence of foreclosure (which also is not a social cost) constitute most of the difference among the models.

Exhibit 9

Loan Status of Those Referred for Financial Counseling, by Outcome

	Referred (percent)	Not Referred (percent)	Difference (percentage points)
Good outcomes			
Current	41.3	38.8	+2.5
Fewer months in default	19.7	14.6	+5.1
Total improvement	61.0	53.4	+7.6
Bad outcomes			
Foreclosed	14.0	21.7	–7.7*
More months in default	12.1	9.3	+2.8
Total decline	–7.7*	+2.8	-2.9*

* Statistically significant difference from zero at 95 percent confidence level.

Design Issues. Assignment in Los Angeles was nonrandom. Los Angeles data should not have been included in the analysis, but this probably does not much affect the impact analysis, as one of the three agencies in Los Angeles was conspicuously ineffective.

Replicability. Probably replicable, with caveats. The text is unduly terse and does not fully characterize the criteria for eligibility of the target population or define counseling (which presumably is defined by HUD regulations located elsewhere).

Generalizability. The study was intended to generalize to other urban areas with high default rates. Note, however, that these are defaults on a subsidized loan program that is no longer offered.

Funding Sources Other Than HUD. None.

Evaluators. Les Rubin and Eugene Johnson.

Information Sources. U.S. Department of Housing and Urban Development, Office of Policy Development and Research (1977).

Public-Use Access to Data. None.

Detroit Default Counseling Demonstration

Synopsis. This 1976–1977 demonstration tested the effects of counseling on a large sample of delinquent or defaulted FHA-insured mortgagors in the Detroit area, although the findings are based on a small sample of the total. Subjects in the analysis sample were followed for 3 to 4 years.

Timeframe. The demonstration was conducted from October 1, 1976, to October 30, 1977. Outcomes as of August 1, 1980, were analyzed.

Treatments Tested. Controls were not offered counseling. Experimentals were offered counseling with an emphasis on money management; they were randomly assigned to one of four counseling agencies.

Target Population. Mortgagors in delinquency or default in any one of the three FHA programs described in the introduction to this section.

Outcomes of Interest. Mortgage repayment.

Sample Size. About 4,400 experimentals were referred for counseling and 4,600 controls were not. For the impact analysis, the evaluators took a random sample of 195 experimentals and 185 controls.

Number of Treatment Groups. Two groups (one control)

Number and Location of Sites. One—Detroit, Michigan.

Research Components.

- Process analysis: Conducted.
- Impact analysis: Comparison of means.
- Cost-effectiveness analysis: All the necessary components of such an analysis were assembled.

Major Findings.

- Process analysis: The evaluators examined the philosophy and functioning of the four agencies. They "seemed to share the feeling that default was a result of many problems but that most defaults could be cured if the homeowner was willing to follow a carefully prepared budget. All the counseling agencies offered guidance to the homeowner in debt management…" Most counselors conducted interviews in their own offices, not in the subjects' homes.
- Process analysis: Just 787 (about 19.7 percent of experimentals who were contacted) actually received counseling; most of the remainder refused it. One agency had a rate of actual counseling that was three times higher than those of the other three agencies.
- Impact analysis: See exhibit 10.

Design Issue. The intake procedures resulted in a very large proportion of experimentals being referred more than once for counseling. The sample of a sample technique (380 households analyzed out of 9,000 randomly assigned) wastes huge amounts of potential information.

Exhibit 10

Mortgage Loan Status as of August 1, 1980—Analytic Sample

	Experimentals (195)	Controls (185)
Current	56.4%	62.7%
In default	16.4	17.8
Foreclosed	27.2	19.5

Replicability. The treatment agencies adhered to defined HUD standards for counseling, and the criteria for eligibility are clearly defined.

Generalizability. Poor. Detroit was ground zero for an explosion of defaults on FHA-insured mortgages in the 1970s. For example, in 1972 the national rate of default was 2.75 percent (an unacceptably high rate), while in Detroit, in June 1971, the default rate was 10.5 percent. Underlying this deterioration was the decline of the city itself, with huge declines in auto industry and related factory employment, and gross administrative deficiencies at the Detroit FHA office. High levels of default could be readily traced to inadequate property inspections, issuance of insurance for buyers who failed to meet underwriting standards, and mortgages well in excess of real market value of the collateral. At the time of the demonstration, 3.7 percent of all dwelling units in Detroit were abandoned, and HUD owned two-thirds of them. Counseling might well be ineffective in this situation without any necessary implications for more normal conditions.

Funding Sources Other Than HUD. None.

Evaluators. Sol Jacobson.

Information Sources. Morgan Management Systems, Inc. (1980).

Public-Use Access to Data. None.

Prepurchase Homeownership Counseling Demonstration

Synopsis. This 1978–79 demonstration tested the effects of prepurchase counseling on a large sample of low- and moderate-income households that were interested in buying a first home. Subjects were followed for no more than 1 year.

Timeframe. Enrollment occurred between June 1978 and early 1979. Counseling operations wound up in June 1979. Data collection continued to September 1979.

Treatments Tested. The controls only received a written Homebuyers Information Package (HIP). Experimentals were assigned to one of three treatment groups, which made the following services available:

- Group I (low intensity): Three group counseling sessions, for up to 7 hours total, with 6 to 10 households per session, and the HIP.
- Group II (medium intensity): Three group counseling sessions, with some individual counseling, for up to 10 hours total, and the HIP.
- Group III (high intensity): Individual advocacy counseling, including housing inspections, attendance at closings, and other individualized services, for up to 14 hours total, and the HIP.

Target Population. First-time prospective homebuyers whose income profiles matched Section 235 income limits (low- and moderate-income).

Outcomes of Interest. Participation, home purchases. The demonstration was also intended to measure the ultimate effect on mortgage default.

Sample Size. 949 experimentals and 292 controls were enrolled in the treatment groups.

Number of Treatment Groups. Four groups (including controls)

Number and Location of Sites. Three—Atlanta, Georgia; Philadelphia, Pennsylvania; and Phoenix, Arizona. There were two counseling agencies at each site.

Research Components. Process analysis conducted. Impact analysis conducted through regression and comparison of means. Benefit-cost analysis not conducted.

Major Findings.

- Process analysis: An intensive and expensive outreach campaign was conducted at the three sites. It had poor results. The evaluators estimated that if there were a national prepurchase counseling program with a comparable outreach effort, then for every 10,000 households that would consider buying a home, 1,960 would learn that a counseling program existed, 1,390 would inquire about it, 740 would enroll, and 477 actually would be counseled. In other words, less than 5 percent of the target population would be reached.
- Process analysis: About two-thirds of households that were assigned to experimental treatments used them.
- Process analysis: Although no limit was placed on the number or length of contacts for individual counseling, the average amount actually received was 2.2 hours.
- Impact analysis: Since the takeup was far below the anticipated level, it was not worthwhile to continue data collection to investigate the impact of prepurchase counseling on the default rate. Default is a relatively rare event; a productive analysis requires large samples.
- Impact analysis: Group counseling (Treatments I and II) reduced the home purchase rate. These groups had a combined home purchase rate of 21.6 percent, while Individual counseling (Treatment III) had a 23.7 percent rate, and the controls had a 29.1 percent rate. The difference between the group rate and the control rate is significant from zero at the 95-percent confidence criterion.
- Impact analysis: Among those who bought homes, "housing expenditures… show no statistically significant variation by Demonstration treatment group."

Design Issues. The major question is the effectiveness of the outreach, which is extensively documented in the reference cited below. "(T)he outreach message for the Demonstration was not perceived as relevant by more than half the eligible homebuyers who had heard of the program." A second question is the provision of the HIP to the controls. The HIP probably contained much of the information that the subjects could have expected to obtain from the counseling, and the process analysis showed that many subjects read it and liked it; this in turn suggests some redundancy in offering counseling as well to an audience that is motivated to read the information for itself.

Replicability. All outreach, eligibility for inclusion, and treatment elements are documented.

Generalizability. See remarks on outreach under Design Issues.

Funding Sources Other Than HUD. None.

Evaluator. Judith Feins.

Information Sources. Feins, Bain, Jr., and White, Jr. (1979).

Public-Use Access to Data. None.

Unfinished Business. The record on homeownership counseling is mixed and difficult to apply to the present market. We have evidence that default/delinquency counseling is modestly efficacious and cost-effective for Section 235 mortgagors at several sites but ineffective for FHA mortgagors in a catastrophic default situation. Section 235 insurance, however, is for good reason no longer offered, and one hopes that the calamity observed in the Detroit market will not be repeated. If Congress wanted to know whether to expand, modify, or eliminate the current counseling programs for delinquent or defaulting FHA Section 203 mortgagors whose characteristics, especially in income, would typically differ from those of Section 235 mortgagors and whose unsubsidized mortgages would typically have less value to the borrower, the experiments performed to date would not give direct and unambiguous evidence.

We also have one experiment demonstrating that a broad marketing strategy for prepurchase counseling will be difficult to implement—takeup rates will be disappointing. The evidence suggests that counseling will reduce the rate of home purchase, which is either a bad outcome (raising the homeownership rate is a major national housing policy objective) or a good outcome (avoiding default by unprepared buyers is highly desirable). In the absence of samples large enough to detect differences in default rates, one cannot distinguish between the two alternatives.

HUD or one of the government-sponsored enterprises (Fannie Mae or Freddie Mac) might consider whether an alternative, more targeted experiment met its particular objectives.[11] For example, let the target population be a pool of mortgage seekers with marginal credit characteristics. They might randomly be assigned to one of four treatments: Group I would have their loans approved. Group II would have their loans approved only if they completed a short, intensive counseling program. Group III would have their loans rejected. Group IV would have their loans accepted but only at substantially higher interest rates. The sponsor would then evaluate, after some length of time, which treatment was the most profitable.[12]

Spatial Mismatch Experiments

Since at least the close of the Second World War, the middle class have been moving to the suburbs and the poor have been moving to the cities. Retail, service, and support jobs have followed the middle class, even when their holders did not earn middle-class wages. The tax base of the cities began to erode, and public services and amenities in the suburbs became perceptibly better, while tax burdens became lower. Zoning and other land-use regulations helped to keep the cost of suburban housing high. Racial discrimination in housing, which was essentially legal up to 1968, and the inadequacy of public transportation helped to ration the accessibility of what affordable housing stock existed.

In short, many of the poor are in one place while the opportunities they need are in another. Two ongoing experiments will evaluate strategies to address the "spatial mismatch." These experiments are both reasonably well known.

Moving to Opportunity for Fair Housing Demonstration

Synopsis. This demonstration tests the effects of housing vouchers that only can be used in low-poverty areas on a large sample of very low-income families with children living in public or assisted housing projects in high-poverty areas. Subjects will be followed for 10 years.

Treatments Tested.

- Rental housing vouchers that may be used only in low-poverty (i.e., middle-class) neighborhoods, with search assistance to find appropriate units.
- Rental housing vouchers without locational restrictions, with no search assistance offered.
- A control group whose members do not receive vouchers but are allowed to remain, if they choose, in their current housing.

Target Population. Very low-income families that apply for the program and that have children; they also live in public housing or Section 8 project-based housing located in central-city neighborhoods with high concentrations of poverty.

Outcomes of Interest.

- Impacts of treatment on families' location choices, housing, and neighborhood conditions.
- Impacts of neighborhood conditions on employment, income, education, and social well-being of demonstration families.

Number of Treatment Groups. Three, including one control.

Sites. Five—Baltimore, Maryland; Boston, Massachusetts; Chicago, Illinois; Los Angeles, California; and New York City, New York.

Funding Sources Other Than HUD. Local matching funding was obtained for the counseling/search assistance programs.

Evaluators. The evaluation contract has not been awarded at this writing. Design, early implementation, and tracking tasks were performed by Abt Associates. Principal investigator is Judith Feins.

Information Sources. HUD-PD&R (1996); Goering, Kraft, Feins, McInnis, Holin, and Elhassan. (1999).

Bridges to Work

Synopsis. This demonstration tests the effects of job placement in the suburbs, transportation, and other supported services on a large sample of work-ready, low-income, inner-city residents. Subjects will be followed for 4 years.

Treatments Tested. Provision of placement in entry-level unsubsidized suburban jobs, transportation, childcare, and other supportive services.

Target Population. Work-ready, low-income, inner-city residents.

Outcomes of Interest. Earnings, employment, and dependence on government support.

Number of Treatment Groups. Two, including one control.

Sites. Five—Baltimore, Maryland; Chicago, Illinois; Denver, Colorado; Milwaukee, Wisconsin; and St. Louis, Missouri.

Funding Sources Other Than HUD. "Predemonstration" costs were partly funded by the U.S. Department of Transportation's Federal Transit Administration, The Ford Foundation, The Rockefeller Foundation, The John D. and Catherine T. MacArthur Foundation, and The Annie E. Casey Foundation. Demonstration costs are partly funded by the Ford, MacArthur, and Rockefeller Foundations as well as a variety of local sources at each of the demonstration sites.

Evaluators: Beth Palubinsky and Joseph Tierney.

Information Sources. A set of brief reports is available on the Web at www.huduser.org. Click on "Programs and Initiatives" and then "Program Evaluation Division."

Unfinished Business. The experiments under way test just two of four broad strategies that one might propose for addressing spatial mismatch—changing the residences of low-income households or raising their resources for solving the mismatch at their current residences. Two other broad strategies never have been tested experimentally, although national and local policy often center on them.

One strategy is to entice the middle class back into the city. We know from nonexperimental studies that police and education issues are substantively important to middle-class locational choice, and initiatives in these areas do not seem to lend themselves to experimental evaluation. One potentially testable policy instrument is a financial incentive to locate in the center city, for example, the substantial Federal income tax credit allowed for new homeowners in Washington, D.C. The feasibility of the experiment would depend on the success of the designers in targeting a group that was likely to respond to the incentive.

The other strategy is sometimes derisively called "gilding the ghetto"—major infrastructure investments to make the inner city irresistible to business and developers. Some version of this strategy has received significant funding for 40 years, under such names as Urban Renewal, Model Cities, Urban Development Action Grants, and Community Development Block Grants.[13] The strategy has met with success in some locations and failure in others. We do not really know what effect this strategy has had, because we do not know what the cities would have looked like in the absence of the funding.

Consider what an experimental evaluation of this strategy would involve. The unit of analysis here is the community, not the household. The outcomes of interest would be aggregate economic variables like property tax base and unemployment rate. To obtain statistical power, one would need many communities. To obtain any measurable impact, one would need a fair amount of money per community.

No administration would propose, and no Congress would approve, the use of a lottery to replace the present formula allocation of the Community Development Block Grant. One could imagine, however, a demonstration under which certain sites randomly received, say, a 10-percent increment to their grants, while comparable sites did not, and one could

imagine an evaluation following the money trail to determine whether there was any measurable impact from that increment.

Power and Limits of Social Experiments in Housing

Properly conducted social experiments have strong *internal* validity. If two groups really do differ only randomly before they are subjected to the treatment, then a significant difference in outcomes caused by something other than the treatment would be a rare event. Experimental data, accordingly, usually have greater explanatory power than nonexperimental data.

There are, in general, limits to this power, and there are specific limits to the power of social experiments in housing. Generally, random assignment might introduce other biases when an analyst attempts to generalize from the experimental sample to a larger population. The act of creating the experimental sample—for example, by requiring greater than normal outreach effort—might result in the subjects having characteristics different from those of the population that normally might be affected by the program. The fact that an experiment is being carried out might also change, positively or negatively, the features of the treatment. For these reasons social experiments may have weak *external* validity—it might be hazardous to generalize from the treatment effects on members of the experimental sample to some larger population.

The underlying facts in social experiments, unlike physical experiments, may change quickly. Carbon 14 is the same thing today that it was 200 years ago, but subsidized homeownership programs come and go. Results may therefore become dated, and extrapolation from past findings to current policy requires care and informed judgment.

Social experiments in housing face a further limitation. The social experiment may be treated as taking place in what economists call a partial equilibrium. The larger rental and labor markets in Baltimore, for example, are little affected by Moving to Opportunity or Bridges to Work, respectively. They might be significantly altered by larger scale programs. The definition of a marginal borrower today is not the same as the definition that lenders would use if a counseling program of known efficacy were required of all marginal borrowers, nor would the interest rates remain constant. In general the universal implementation of an experimental treatment might cause changes in market prices, quantities, and non-price exchange conditions. These “general equilibrium” consequences are difficult to measure experimentally, except at great cost.[14]

Even these generalizations about either the power or the limits of the methodology may need to be revised as we gain further experience. The potential scope of experiments is much wider than the actual experience to date. Of the 217 completed and ongoing social experiments documented in our book, 213 were tested on U.S. territory. Three took place in Canada, and one in the Netherlands. We know very little about what social experiments can or cannot tell outside of North America.

Moreover, the scope of the social experiment has been largely (but not entirely) limited to programs affecting persons who are in some way disadvantaged. Programs affecting the broad middle class seldom have been evaluated experimentally. One important example of a neglected experiment in the housing field is the income-tax deduction for mortgage interest by homeowners. This deduction is targeted for repeal in nearly all the proposals for replacing the current income tax system. Macro simulations of the impact of such a repeal on the housing market and on the U.S. economy generally vary from catastrophically negative to moderately positive. It has not apparently occurred to the makers of tax

policy to see what would happen to housing demand if a small set of taxpayers were presented with a radically different tax structure as a trial run.

Over time one would expect the geographic and issue restrictions to relax. The knowledge created by social experiments in housing to date has been valuable, and it is realistic to expect substantial further gains.

Author

Mark Shroder is an economist with the Office of Policy Development and Research, U.S. Department of Housing and Urban Development. He received his Ph.D. from the University of Wisconsin-Madison in 1991. He is the author, with David Greenberg, of The Digest of Social Experiments, *published by Urban Institute Press, and of a related article with Greenberg and Matt Onstott in the* Journal of Economic Perspectives. *He has published theoretical and empirical articles on income redistribution in Federal systems, rental assistance, rent/commuting tradeoffs for tenants of different races, and the Real Estate Settlement Procedures Act.*

This paper developed out of work with David Greenberg, University of Maryland-Baltimore County, and Matthew Onstott, U.S. Department of Education. Ideas and phrases in this paper may have been absorbed unintentionally from them. The major published joint products to date have been Greenberg and Shroder (1997) and Greenberg, Shroder, and Onstott (1999). The assistance of Lester Rubin and Martin Abravanel in uncovering and interpreting the homeowner counseling experiments is appreciated, as are the comments of Amy Bogdon, and of participants at the Mid-Year 1998 Meeting of the American Real Estate and Urban Economics Association meeting in Washington, D.C., and the 1999 meeting of the Southern Regional Science Association in Richmond, where earlier versions of this paper were presented. The author has sole responsibility for opinions expressed here, and this paper is not an expression of the policy of the U.S. Department of Housing and Urban Development.

Notes

1. The power and limits of the methodology are discussed in the last section. This paper is not intended as a general defense of social experiments. See the introduction in Greenberg and Shroder (1997) for a general treatment of the subject.

2. This was a "response surface" experiment, in which wide variation in two or more design parameters permits easy estimation of the elasticity of demand. This kind of experiment is an economist's dream but an administrator's nightmare; no new response-surface experiments have been initiated for the past 20 years.

3. Since the administrator and the evaluator were the same in this experiment, the process analysis was quite sketchy. For such a large experiment, the absence of information about what went on inside the "black box," and the failure to examine the way the experiment functioned in the concrete context of people's lives is disappointing, but typical of experiments in that period.

4. These estimates are much lower than similar values for middle- and upper-income households. The finding that low-income households have low elasticity of demand for housing relative to income is a major contribution of this experiment.

5. Some researchers had access to data that HUD intended to transform into a public-use file (see Bartik, Butler, and Liu, 1992).

6. During the course of the experiment the PHA had the option of setting the payment standard below the FMR. This allowed the agency both to serve more families and to receive more administrative fees.

7. Rent reasonableness enforcement by the PHA is now a feature of both the certificate and voucher programs.

8. Researchers who could demonstrate a reasonable analytic need might be able to obtain some data from HUD, provided the privacy of participants was fully protected.

9. The basic data of this demonstration have been destroyed. Researchers cannot go back and get them.

10. Quercia and Wachter (1996), in an article on measuring the performance of homeownership counseling, advocate random-assignment evaluation of counseling programs but do not reference the results of these actual experiments.

11. I am aware of two random assignment experiments conducted by for-profit firms. They were electricity pricing experiments conducted by power companies on small samples of commercial customers. A Fannie Mae or Freddie Mac experiment would probably require the collaboration of one or more private mortgage insurers.

12. The design suggested could be modified in many ways. One necessary element of experimental design would be determining who would pay for the counseling.

13. One could also include Empowerment Zones/Enterprise Communities, which will have nonexperimental evaluations.

14. To measure the effects of housing subsidies on the level of rents (among other things) in the private market, HUD set up saturation demonstrations in Green Bay and South Bend contemporaneously with the Housing Allowance Demand Experiment. See Lowry (1983). Subsidies had little apparent effect on rents; the relevant portions of the rental market had highly elastic supply. A sample size of two (markets) does not leave many degrees of freedom, but that is the point in the text. Demonstrations on a larger scale would quickly exhaust the support of the appropriators.

References

Bartik, Timothy J., J.S. Butler, and Jin-Tan Liu. 1992. “Maximum Score Estimates of the Determinants of Residential Mobility: Implications for the Value of Residential Attachment and Neighborhood Amenities,” *Journal of Urban Economics* 32:233–256.

Bradbury, Katharine L., and Anthony Downs, eds. 1981. *Do Housing Allowances Work?* Washington, DC: Brookings Institution.

Feins, Judith D., L. Dixon Bain, Jr., and Charles S. White, Jr. 1979. “Outreach and Participation in the Prepurchase Homeownership Counseling Demonstration: Interim Report.” Cambridge, MA: Abt Associates.

Goering, John, Joan Kraft, Judith Feins, Debra McInnis, Mary Joel Holin, and Huda Elhassan. 1999. *Moving to Opportunity for Fair Housing Demonstration Program:*

Current Status and Initial Findings. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Greenberg, David, and Mark Shroder, with the assistance of Matthew Onstott. 1997. *The Digest of Social Experiments*, 2nd ed. Washington, DC: Urban Institute Press.

Greenberg, David, Mark Shroder, and Matthew Onstott. 1999. "The Social Experiment Market," *Journal of Economic Perspectives* 13, 3 (Summer):157–172.

Kennedy, Stephen D. 1980. *Final Report of the Housing Allowance Demand Experiment.* Cambridge, MA: Abt Associates Inc.

Leger, Mireille L., and Stephen D. Kennedy. 1990. *Final Comprehensive Report of the Freestanding Housing Voucher Demonstration*, two volumes. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Lowry, Ira S., ed. 1983. *Experimenting with Housing Allowances: Final Report of the Housing Allowance Supply Experiment.* Cambridge, MA: Oelgenschlager, Gunn, and Hain.

Morgan Management Systems, Inc. 1980. "A Report on the Detroit Default Counseling Demonstration: An Assessment of Counseling as a Default Remedy," manuscript, Columbia, MD: Morgan Management Systems, Inc.

Quercia, Roberto G., and Susan M. Wachter. 1996. "Homeownership Counseling Performance: How Can It Be Measured?" *Housing Policy Debate* 7, 1:175–200.

Struyk, Raymond J., and Mark Bendick, Jr., eds. 1981. *Housing Vouchers for the Poor: Lessons from a National Experiment*. Washington, DC: Urban Institute.

U.S. Department of Housing and Urban Development, Office of Policy Development and Research. 1996. *Expanding Housing Choice for HUD-Assisted Families: Moving to Opportunity*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

——. 1980. *The Experimental Housing Allowance Program: Conclusions.* Washington, DC: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.

——. 1979. *Experimental Housing Allowance Program: A 1979 Report of Findings.* Washington, DC: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.

——. 1977. *Counseling for Delinquent Mortgagors II: A Staff Study*. Washington, DC: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.