



U.S. Department of Health and Human Services
Assistant Secretary for Planning and Evaluation
Office of Disability, Aging and Long-Term Care Policy

METHODOLOGICAL ISSUES IN THE EVALUATION OF THE NATIONAL LONG TERM CARE DEMONSTRATION

July 1986

Office of the Assistant Secretary for Planning and Evaluation

The Office of the Assistant Secretary for Planning and Evaluation (ASPE) is the principal advisor to the Secretary of the Department of Health and Human Services (HHS) on policy development issues, and is responsible for major activities in the areas of legislative and budget development, strategic planning, policy research and evaluation, and economic analysis.

ASPE develops or reviews issues from the viewpoint of the Secretary, providing a perspective that is broader in scope than the specific focus of the various operating agencies. ASPE also works closely with the HHS operating divisions. It assists these agencies in developing policies, and planning policy research, evaluation and data collection within broad HHS and administration initiatives. ASPE often serves a coordinating role for crosscutting policy and administrative activities.

ASPE plans and conducts evaluations and research--both in-house and through support of projects by external researchers--of current and proposed programs and topics of particular interest to the Secretary, the Administration and the Congress.

Office of Disability, Aging and Long-Term Care Policy

The Office of Disability, Aging and Long-Term Care Policy (DALTCP), within ASPE, is responsible for the development, coordination, analysis, research and evaluation of HHS policies and programs which support the independence, health and long-term care of persons with disabilities--children, working aging adults, and older persons. DALTCP is also responsible for policy coordination and research to promote the economic and social well-being of the elderly.

In particular, DALTCP addresses policies concerning: nursing home and community-based services, informal caregiving, the integration of acute and long-term care, Medicare post-acute services and home care, managed care for people with disabilities, long-term rehabilitation services, children's disability, and linkages between employment and health policies. These activities are carried out through policy planning, policy and program analysis, regulatory reviews, formulation of legislative proposals, policy research, evaluation and data planning.

The paper was written as part of contract #HHS-100-80-0157 between ASPE and Mathematica Policy Research, Inc., and contract #HHS-100-80-0133 between ASPE and Temple University. Additional funding was provided by the HHS Administration on Aging and HHS Health Care Financing Administration (now the Centers for Medicare and Medicaid Services). For additional information about this subject, you can visit the DALTCP home page at http://aspe.hhs.gov/_/office_specific/daltcp.cfm or contact the office at HHS/ASPE/DALTCP, Room 424E, H.H. Humphrey Building, 200 Independence Avenue, S.W., Washington, D.C. 20201. The e-mail address is: webmaster.DALTCP@hhs.gov. The Project Officer was Robert Clark.

METHODOLOGICAL ISSUES IN THE EVALUATION OF THE NATIONAL LONG TERM CARE DEMONSTRATION

Randall S. Brown

Mathematica Policy Research, Inc.

April 1986
Revised July 1986

Prepared for
Office of the Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services
Contracts #HHS-100-80-0157 and #HHS-100-80-0133

The opinions and views expressed in this report are those of the authors. They do not necessarily reflect the views of the Department of Health and Human Services, the contractor or any other funding organization.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
I. INTRODUCTION	1
II. THE BASIC DESIGN OF THE EVALUATION	3
A. Random Assignment.....	3
B. Data Sources Used in the Analysis.....	5
C. The Analysis Samples	8
D. Sample Sizes and Precision of Impact Estimates.....	10
III. STATISTICAL METHODOLOGY	13
A. The Regression Model.....	13
B. Testing Strategy.....	19
C. Interpretation of Estimates	22
D. Estimated Impacts for Subgroups.....	24
IV. SPECIFIC METHODOLOGICAL ISSUES EXAMINED	26
A. The Equivalence of Treatment and Control Groups at Randomization.....	26
B. The Comparability of Baseline Data for Treatment and Control Groups	28
C. The Effects of Sample Attrition	32
D. The Validity of Pooling Observations	38
E. Differences Between Early and Late Cohorts of Sample Members	41
F. Potential Problems with Regression Analysis	43
G. The Effects on Impact Estimates of Using Proxy Respondents	54
V. CONCLUSION	58
REFERENCES	61

LIST OF TABLES

TABLE II.1. Channeling Sites and Sample Sizes	4
TABLE II.2. Sample Sizes Used in the Evaluation	12
TABLE III.1. Mean Values of Explanatory Variables Used in Regression Model	17
TABLE IV.1. Percent of Full Sample Included in Analysis Samples	33
TABLE IV.2. Impact Estimates from Least Squares Regression and from Probit for Selected Binary Outcome Measures	46
TABLE IV.3. Impact Estimates from Least Squares Regression and from Tobit for Selected Truncated Outcome Measures	48
TABLE IV.4. Comparison of Least Squares, Tobit, and 2-Part Estimates of Channeling Impacts for Skewed Outcome Variables	52

ACKNOWLEDGMENTS

Many people contributed indirectly to this report, despite the fact that mine is the only name listed on the title page. As is true for every report in the channeling evaluation, the most important contributions were made by Peter Kemper, my co-principal investigator. His comments on drafts of this report and on the work on which it is based were insightful and invaluable. Peter Mossel was my co-author on the analyses of attrition bias and baseline data comparability summarized herein. Jennifer Schore and Nancy Holden also contributed to the attrition analysis, and to the analysis of the effects of outliers on impact estimates. Margaret Harrigan helped write the report that assessed the equivalence of the treatment and control groups. Shari Dunstan did the programming for the analysis of pooling and cohort effects and helped write the long memorandum documenting the results of the latter. Robert Applebaum helped to interpret the results from the analysis of the effects of using proxy respondents at followup. In addition, a number of research assistants ably programmed the necessary computations and several secretaries patiently performed the necessary word processing. Helpful comments on methodological issues were provided at various stages during the evaluation by Mary Harrahan and Robert Clark of the Department of Health and Human Services and by the internal and external-review panels set up by the Department.

I. INTRODUCTION

In September 1980, the National Long Term Care Demonstration-known as channeling-was initiated by the United States Department of Health and Human Services. It was to be a rigorous test of comprehensive case management of community care as a way to contain the rapidly increasing costs of long term care for the elderly while providing adequate care to those in need. The key goal was to enable elderly persons, whenever appropriate, to stay in their own homes rather than entering nursing homes.

Two models of channeling were tested, each implemented in 5 sites. Under the basic case management model, the channeling project assumed responsibility for helping clients gain access to needed services and for coordinating the services of multiple providers. This model provided a small amount of additional funding to fill in gaps in existing programs. But it relied primarily on what was already available in each community, thus testing the premise that the major difficulties in the current system were problems of information and coordination which could be largely solved by client-centered case management.

The financial control model differed from the basic model in several ways. The primary difference was that it established a funds pool to ensure that services could be allocated on the basis of need and appropriateness rather than on the eligibility requirements of specific categorical programs. The pooled funds could be used to purchase a broader range of community services than were covered by Medicare and Medicaid. Case managers were responsible for determining the amount, duration, and scope of services paid for out of the funds pool, subject to limits on the amount that could be spent on any one case.

The goal of the evaluation was to determine the impact of channeling on several key outcomes:

- Use of formal health and long term care services, particularly hospital and nursing home care and community services
- Public and private expenditures for health services and long term care
- Personal well-being of clients, including mortality, physical functioning, unmet service needs, and social/psychological well-being
- Caregiving by family and friends, including the amount of care provided, the amount of financial support provided, and caregiver stress, satisfaction, and well-being

Research on these topics have been conducted over the past two years, culminating in a series of final reports.

The credibility of the estimates obtained depends heavily on the methodology used to obtain them. Many previous studies of case management or community service programs have methodological flaws that raise serious doubts about the findings, including use of poorly matched comparison groups, restriction in the number and diversity of sites examined, small sample sizes, and inadequate data (see Kemper et al, 1986 for a review of previous studies). Thus, one of the initial purposes of the channeling demonstration was to provide a sound methodological basis for assessing the impacts of such programs.

The purpose of this report is to describe the methodology used throughout the channeling evaluation and to document the major analytical issues that posed potential threats to the credibility of the analysis. Because these topics are quite technical and affect all areas of the analysis, they received relatively little attention in the various final reports on specific channeling impacts. This document provides a more thorough explanation of the estimation procedures and test statistics employed, and summarizes our investigations of specific methodological issues.

Although the discussion of technical issues contained here is more comprehensive than that provided in the final reports on channeling impacts, it is not a detailed review of all of the analyses conducted. Such a document would be extremely long and would make the important issues less accessible to readers. Rather, the discussion here is intended to describe the methodological concerns that we had, how we examined them, and the conclusions we came to, with relatively little presentation of the statistical evidence. In the discussions of specific methodological issues we indicate how the interested reader can obtain more detailed documentation of the results of these investigations.

Issues concerning the overall design of the evaluation are not addressed here. Thus, there is no assessment of the generalizability of the evaluation results to settings other than the 10 sites in which the demonstration was implemented (see Kemper et al, 1986 and Carcagno et al, 1986 for discussion of this topic). Nor does this report attempt to provide a guide for future evaluations on which design features are essential and what pitfalls should be avoided. Although that would be useful, it requires consideration of the economic and political costs of incorporating such features, which deviates too far from the statistical topics on which this report is focused. We hope to address these issues in a separate forum.

The remainder of the report is organized into four chapters. Chapter II describes the basic design of the evaluation, including the data sources and sample sizes. Chapter III presents the statistical methodology used to estimate channeling impacts, and the test statistics and assessment strategy used to draw inferences about whether observed treatment/control differences were attributable to channeling or to chance. In Chapter IV, we describe briefly eight specific methodological problems that arose and were examined. Finally, Chapter V recaps the primary methodological findings and gives an overall assessment of the methodology used in the analysis.

II. THE BASIC DESIGN OF THE EVALUATION

The evaluation was designed to avoid the methodological shortcomings of many of the previous studies of long term care programs. The key features of the design were:

- Use of a control group that was randomly selected from eligible channeling applicants
- The collection of high quality data on a large number of observations
- Implementation of the demonstration in a number of different settings

The last point requires little elaboration. The demonstration was implemented in 10 sites, 5 implementing the basic model and 5 the financial control model. While the set of demonstration sites may not be representative of the nation as a whole and was heavily concentrated in the northeast, it did include both urban and rural areas, and there was considerable variation across sites in the availability of case management, community services, and nursing home beds. The sites and sample sizes drawn from each are given in Table II.1. The variation in sample sizes reflects the differences between sites in the size of the elderly population. (See Carcagno et al., 1986 for comparison of characteristics of the demonstration sites and a discussion of the process by which they were selected.)

The other key features of the design require somewhat more discussion. One of the most important aspects of the evaluation design is that the use of random assignment means that throughout the evaluation, impacts of channeling are defined as the difference between what actually happened to treatment group members and what would have happened to them in the absence of the demonstration. Since other forms of case management are available in the demonstration sites and since many individuals receive needed community-based services without any case management, the estimated channeling impacts are not the effects of channeling as compared to nothing, but rather channeling as compared to whatever case management and services were already available in the demonstration sites. It is equally clear then that the estimates cannot be interpreted more broadly as the effects of case management and community services in general. With this in mind, the random assignment procedures and the data sources used in the evaluation are described briefly below.

A. RANDOM ASSIGNMENT

The single most important feature of the evaluation was the use of an experimental design. Persons referred or otherwise applying to channeling were given a screening interview (usually by phone) to determine their eligibility for the program. Applicants had to be at least age 65, have at least a moderate level of disability in performing certain activities, and have needs for two or more services that were

expected to remain unmet for the next six months.¹ Eligible individuals were then randomly assigned by survey research staff in Princeton to either the treatment group, which was offered the opportunity to participate in channeling, or to the control group, which was denied that opportunity for the course of the demonstration.² Both groups were then given a baseline interview, and followups were attempted at six, twelve, and (for half the sample) 18 months thereafter to gather data on outcomes that channeling was intended to affect.

TABLE II.1. Channeling Sites and Sample Sizes			
Site	Treatment Group	Control Group	Full Sample
Basic Case Management Model			
Baltimore	471	271	688
Eastern Kentucky	246	242	488
Houston	401	273	674
Middlesex County	451	299	750
Southern Maine	264	260	524
Total	1,779	1,345	3,124
Financial Control Model			
Cleveland	388	191	579
Greater Lynn	309	308	617
Miami	450	297	747
Philadelphia	581	288	869
Rensselaer County	195	195	390
Total	1,923	1,279	3,202
All Sites	3,702	2,624	6,326
NOTE: These sample sizes are the numbers of research sample members with completed screening interviews that were available for analysis. A total of 6,341 individuals completed screens and were randomly assigned to the treatment or control groups, but 14 screen interviews were lost in the mail and one sample member's treatment status was misrecorded. These 15 cases were excluded from all analyses.			

Given the random assignment, the control group should closely resemble the treatment group in the aggregate on both observable and unobservable characteristics. Thus, it should provide as good an indication as possible of what would have happened during the followup period to treatment group members in that site in the absence of the demonstration. The impacts of channeling are estimated by comparing the post-randomization experience of the treatment and control groups, using the estimation procedures described in Chapter III. The random assignment ensures that estimated impacts will be unbiased.

It should be kept in mind that not all members of the treatment group actually participated in channeling (some refused, some died, other were terminated after being permanently institutionalized). Furthermore, not all treatment group members who did participate remained in the program for the duration of the analysis period. However, to ensure unbiasedness of estimates it was essential that the full treatment group (i.e., everyone offered the opportunity to participate) be compared to the full control group.

¹ See Carcagno et al (1986, Chapter VI) for a detailed description of the eligibility criteria.

² See Phillips et al (1986, pp. 39-46) for a detailed description of the randomization procedure used.

While channeling obviously has no impact on treatment group members who do not participate, deleting such observations would destroy the equivalence of the two groups being compared if, as one expects, treatment group members who drop out differ from those who participate.³ Without this equivalence, comparison of the two groups would no longer yield unbiased estimates of channeling impacts. Thus, treatment group members was ignored, and the treatment/control comparisons yield unbiased estimates of the effects of having the opportunity to participate in channeling.

B. DATA SOURCES USED IN THE ANALYSIS

The second most important feature of the evaluation design was the collection of comprehensive longitudinal data from a variety of sources on a large sample of individuals. These features increase the precision (i.e., decrease the variance) of the impact estimates-in two ways: (1) by gathering various data items from the best sources (e.g., nursing home data from Medicare, Medicaid, and provider records) the analysis variables contain less "noise" due to measurement error, and (2) the large sample sizes decrease the likelihood that observed treatment/control differences are due to chance rather than to the effects of channeling.

To conduct the evaluation, data were required on both the initial (preprogram) characteristics of the sample and on outcome variables which measure the post-randomization experience of the sample. Outcomes which channeling was expected to influence were grouped into 6 substantive areas:

- Hospital use
- Nursing home use
- Use of formal community care
- Receipt of informal care from family and friends
- Mortality
- Well-being of clients and their caregivers

In order to obtain the best data to address these issues, various sources were required, including both interviews with sample members and records from specific programs and agencies. These sources are described below.

1. Interview Data

Interview data⁴ sources include: (1) the screen interview, which was administered to all persons referred or applying to channeling to assess their eligibility for the program; (2) the baseline interview, administered to eligible sample members as

³ See Carcagno et al (1986, Chapter VIII) for some statistics on the proportion of treatment group members who were terminated from channeling, the reasons for termination, and the points at which termination occurred.

⁴ See Phillips et al (1986) for complete documentation of interview data collection procedures.

soon as possible after they were assigned to the treatment or control group (average length of time between screen and baseline was about one week for treatment group members and almost two weeks for controls); and (3) the followup interviews, administered 6, 12, and 18 months after randomization in order to obtain data on outcomes which channeling was hypothesized to influence.⁵

The Screen. The screen questionnaire, administered primarily by telephone by channeling intake workers, was designed primarily to assess eligibility for channeling and contained data on sample members' ability to perform various activities of daily living, their unmet needs for assistance of several types, and some sociodemographic characteristics. Applicants determined to be eligible for channeling were then randomly assigned to treatment or control status by research staff. Screen interviews were completed with 6,341 eligible sample members. Unfortunately, 14 screen interviews were lost in the mail and one case assigned to the control group was erroneously allowed into the channeling program, so that there are actually 6,326 individuals who could be included in the analysis. These observations thus constitute the full research sample, and we refer to it as such throughout this report.

The Baseline. The screen interview does not, however, contain the comprehensive data that were necessary for either the evaluation or the development of a care plan for channeling clients. A thorough, in-person baseline assessment of treatment group members was required in order for program case managers to develop an appropriate care plan for participants. A single instrument was developed that would serve both the purpose of care planning and research. It was considered important that channeling staff members collect the data necessary for developing an appropriate care plan; therefore, the baseline interview (but not the followup interviews) was administered by channeling staff for the treatment group and by research interviewers for the control group.⁶ Treatment group members who refused the baseline assessment interview could not participate in channeling, since no care plan could be developed for them. However, since these individuals could differ substantially from other treatment group members, nonresponding members of the treatment group were interviewed by research interviewers whenever possible. This enabled us to retain them in the analysis sample and thereby helped to preserve the equivalence of the treatment and control groups. Overall, 108 (3 percent) of the baseline interviews for the treatment group were administered by research interviewers. Sample members who failed to complete baselines were not followed up and were excluded from most of the channeling analyses.

The Followup Interviews. For sample members who completed the baseline, followup interviews at 6, 12, and 18 months after randomization were attempted by research interviewers to gather the data on sample members' outcomes that were

⁵ In addition to these surveys of sample members, there were also surveys of the primary caregivers of a subset of the sample members. Data from these surveys were used primarily in the evaluation of the effects of channeling on caregivers. Methodological issues related to this sample are examined in Christianson (1986).

⁶ How this difference affected the comparability of the baseline data for the two groups is summarized in Section B of Chapter IV in this report.

necessary to assess the impacts of channeling. Although a completed baseline was a condition for being contacted for a followup interview, a noncompleted 6-month interview did not make the sample member ineligible for a 12-month interview. Thus, some sample members who did not complete a 6-month interview did complete a 12-month interview.

The situation was different at the 18-month interview. First, to reduce the length of the data collection period and costs, only the first half of the sample members randomized were eligible for an 18-month interview.⁷ Second, an 18-month followup was attempted only if the sample member belonged to this first half of the sample (referred to as the 18-month cohort), and had a completed baseline, 6-month and 12-month followup interview.

2. Records Data

Records data used in the channeling evaluation included Medicare and Medicaid claims data, records data from providers of services (e.g., nursing homes) that sample members claimed in the interview to have used, financial control system data from the channeling projects (for channeling clients in financial control sites), and death records.

Medicare Claims Data. Medicare claims data were collected for all sample members who said that they were eligible for Medicare and for whom a valid Medicare identification number could be verified by HCFA. Nearly the entire sample (97 percent) was eligible for Medicare. Claims provided data on sample members' hospital use, some nursing home use, and use of other medical services and community-based services paid for by Medicare. See Wooldridge and Schore (1986) for a detailed discussion of Medicare data.

Medicaid Claims Data. Medicaid claims were collected for all sample members who said they were eligible for Medicaid at any interview and signed a consent form authorizing use of the data, if a valid Medicaid identification number could be verified by the state Medicaid program. Medicaid claims were a key source of data on nursing home outcomes and use of formal community services.

Provider Records Data. Data on the nursing home use of specific sample members were collected from nursing homes for sample members stating in an interview that they had spent time in that institution during the reference period or were living there at the time of the interview. Records data were also collected from area hospitals on those few sample members who were not on Medicare. For a random 20 percent subsample of the research sample, records were also collected from other types of service providers (e.g., home health agencies) that were named in followup interviews by sample members.⁸

⁷ See Phillips et al (1986) for a discussion of the 18-month cohort and interview.

⁸ See Phillips et al. (1986) for a detailed description of the provider records data.

Financial Control System Data. Because of the pooling of Medicare, Medicaid, and in some cases other government funds in the financial control model, data on use of formal community services by treatment group members in that model were obtained from the channeling project's records.

Death Records. Data on mortality were obtained from a search of state death records for all sample members who failed to complete their last scheduled interview. These data were supplemented by data on mortality obtained in the attempt to field the followup interviews and from client-tracking data (for treatment group members).

C. THE ANALYSIS SAMPLES

For 5 of the 6 categories of outcomes identified above, the sources and therefore the completeness of the necessary data differ. For each substantive area we defined "analysis samples," i.e., that subset of the full research sample for which the data necessary for analysis were available. The analysis samples for the substantive areas were:

- Mortality--full research sample
- Hospital outcomes--6, 12, and 18 month Medicare samples
- Nursing home outcomes--6, 12, and 18 month nursing home samples
- Well-being outcomes--6, 12, and 18 month followup samples
- Receipt of formal community based services and informal care--6, 12, and 18 month in-community samples

These samples and the relationship between them are described below.

Full Sample. This sample included all of the 6,326 initially randomized individuals, and was used to estimate the impacts of channeling on mortality, as measured by whether sample members were alive at 6 and 12 months after randomization and by the number of survival days as of the end of each period. The full 18-month sample, used to estimate impacts on mortality at 18 months, included the 3,165 members of the full sample who were in the 18-month cohort. A search of state death records was conducted for all sample members not known to be alive from the interviews, and these records data were supplemented with information on deaths obtained from attempts to field followup interviews and from channeling programs' client tracking system. Sample members identified as dead from either source were assumed to be alive; hence, there was no missing data on mortality. An analysis of the validity of this assumption, presented in Wooldridge and Schore (1986, Appendix F), makes use of Medicare claims data and updated status files to verify that the assumption is correct for virtually all sample members.

The Medicare Sample. The Medicare sample was employed to examine channeling's impacts on the use of hospital and other medical services, and on home health expenditures reimbursed by Medicare. The Medicare sample is the subset of the

6,326 initially randomized individuals (the full sample) who completed baseline interviews and who were either known to be Medicare entitled or known not to be Medicare entitled. This sample was used for analyzing outcomes in the first 12 months following randomization. To be consistent with the analyses of channeling's impacts on outcome measures obtained from followup interviews, the 18-month Medicare sample was restricted to those members of the Medicare sample who were also in the 18-month cohort.

The Nursing Home Samples. Because Medicare claims do not provide a complete history of nursing home use, the samples used for the nursing home analysis differed from those used for the hospital analysis. Most nursing home expenses are paid by Medicaid for Medicaid-covered individuals, and by private payment for those not covered by Medicaid. Therefore, the nursing home analysis employed a two-pronged data collection strategy, relying on Medicaid (and Medicare) records to provide complete nursing home information for sample members who were covered by Medicaid, and on provider (and Medicare) records for those who were not Medicaid-covered. However, in order to identify the relevant providers for this latter group, either a followup interview or caregiver interview had to have been completed.

These data requirements resulted in three nursing home samples, one for each six-month period. These are subsamples of the Medicare samples, and include individuals who either completed a followup interview, were Medicaid covered throughout the six-month period, or died in the period but had a caregiver who provided followup information. In addition, Medicare sample members who were dead throughout a six-month period, or who died during the period and were Medicaid-covered at the start of the period and at death were also included in the nursing home sample for that period.⁹

The Followup Samples. The followup samples were used to analyze outcomes obtained from the followup surveys administered at 6, 12, and 18 months after randomization. The two major categories of impact analyses which relied on these samples were those dealing with sample members' well being and functional ability and those dealing with case management services. The followup sample at 6 months included the subset of the screen sample with both a complete baseline and a complete 6-month followup interview. In like manner, the sample at 12 months was composed of screen sample members who completed a baseline and a 12-month followup (but not necessarily a 6-month interview). The 18-month sample included only those in the early cohort who completed a baseline and all three followup interviews.

The In-Community Samples. Estimation of channeling impacts on receipt of formal and informal care required data on these outcome variables from the followup surveys. The interview data on receipt of such services pertained to the reference week—the week during which the 6, 12, and 18 month interview was conducted, for sample members residing in the community at followup. Therefore, the 6, 12, and 18 month

⁹ We also estimated channeling impacts on the “survivor” samples, the subset of the nursing home sample consisting of sample members who were alive at the beginning of the period being analyzed.

in-community samples were composed of those sample members who completed the relevant interview and were living in the community during the reference week.¹⁰

These five different sample types form a hierarchy, with each being nested, or nearly so, within the one above it. See Brown et al. (1986) for a schematic representation of the relationship between them.

D. SAMPLE SIZES AND PRECISION OF IMPACT ESTIMATES

The number of observations in each of these analysis samples is given in Table II.2. Observations were roughly evenly split between the models, but included more treatments than controls. This imbalance arose because the rate of client intake was slower than expected and would not have yielded a sample size large enough to meet the desired precision standards if the sample were restricted to only the subset of applicants who were initially intended for the research sample.¹¹ The in-community sample at 12 months included only 40 to 50 percent of the full research sample for either model or experimental group, due in large part to the high mortality rate for the sample (nearly 30 percent by 12 months). Sample sizes at 18 months were much smaller, of course, since followup data were sought for only half of the original sample.

The samples were originally designed to meet a certain precision standard. The precision standard specified (see Kemper et al., 1982) was based on the accuracy of the estimates of channeling impacts on nursing home use, since that was the single most important, outcome measure and the source of most of the cost savings channeling was expected to generate. The sample was to be large enough so that if 50 percent of controls entered a nursing home, we would be able to identify channeling impacts of 6 percentage points or larger with 90 percent power and 90 or 95 percent

¹⁰ Restriction of the sample used for analysis to those living in the community during the reference week yields more meaningful estimates of service use, since sample members who were dead or in a hospital or nursing home at reference week were by definition receiving no formal or informal community care. However, this restriction of the sample would yield biased estimates of program impacts on formal and informal care if the program affected the mortality, hospital or nursing home use of sample members. Given the lack of channeling impacts on these other outcomes, we were able to use the in-community sample.

¹¹ The original sampling plan called for a research sample with an equal number of observations from each of the ten sites, with observations in each site equally split between treatment and control groups: about 300 treatments and 300 controls in each site, with about 240 of each group expected to be available for analysis (assumed 80 percent response rates). Our estimate was that in all but the smallest sites the supply of eligible applicants during the 12 month intake period over which the sample was to be drawn would exceed the number of observations required. Thus, in the smallest sites, applicants were assigned to treatment or control groups on an equal basis, but in medium sites three fifths were assigned to treatment states and two-fifths to control states with only 2/3 of the treatment group to be included in the research sample. In the largest sites, one-third of the applicants were assigned to control status and two-thirds to treatment status, with only half of the latter group to be included in the research. However, caseload buildup in several sites was slower than projected during the first 5 months of intake, so those treatment group members not originally intended for the research sample were in fact included in the analysis in order to achieve the desired total sample size. As a result, sample sizes differed across sites and the treatment group was larger than the control group in the medium and large sites. See Kemper et al (1982, pp. 37-39) for a more detailed explanation.

confidence (for a two-tailed or one-tailed test, respectively). This means that the sample was to be large enough that if channeling actually reduced the probability of nursing home admissions by 6 percentage points or more, there would be at least a 90 percent a priori probability that conventional statistical tests conducted on the sample would reject the hypothesis that channeling had no impact. With such samples, the probability of erroneously concluding that channeling had no impact when in truth it had affected nursing home use (type II errors) and the probability of erroneously concluding that channeling had reduced nursing home when in fact it had not (type I errors) would both be small.

The actual precision of our tests differed from this standard for 3 reasons. First, we actually conducted tests at the 95 rather than the 90 percent confidence level, using two-tailed tests, which meant that differences had to be as large as 6.6 percentage points in order to have 90 percent power of detecting them. Second, the actual sample sizes differed from the 1200 treatments /1200 controls that were required to produce the desired precision (see Table II.2), for the reasons cited above. If half of the control group had entered nursing homes, we would have been able to detect impacts of 6.9 percentage points or larger with 90 percent power and 95 percent confidence using the actual sample and two-tailed tests.

In fact, however, the control group use of nursing homes was much smaller than the assumed 50 percent, which had been used in the calculations because it resulted in the largest possible variance for a binary variable. This is the third and by far the most important reason why the precision of our estimates differed from what was originally planned. Given that only 13 percent of controls were admitted to nursing homes in the first 6 months, the variance was smaller than assumed. Thus, the sample was sufficiently large to detect impacts as small as 4.6 percentage points. Proportionately, however, 4.6 percentage points is over one-third of total actual use. Thus, unless channeling's impact on nursing home use was proportionately quite large, we cannot be highly confident that the treatment/control difference observed in our sample will be significantly different from zero statistically. Had control group use been equal to the assumed 50 percent, reductions due to channeling as small as 14 percent (6.9 percentage points) would have been detectable.

Despite this fact, the sample sizes are sufficiently large that it is very unlikely that channeling impacts large enough to make channeling a cost-effective program would go undetected by the statistical tests. Thus, the sample sizes used in the evaluation were large enough to ensure a low probability of either seriously overstating or understating channeling impacts.

TABLE II.2. Sample Sizes Used in the Evaluation									
	Basic Model			Financial Control Model			Full Sample		
	Treatments	Controls	Total	Treatments	Controls	Total	Treatments	Controls	Total
Number of Observations in Full Sample	1,779	1,345	3,124	1,923	1,279	3,202	3,702	2,624	6,326
6 Month Outcomes									
Medicare sample	1,608	1,104	2,712	1,795	1,047	2,842	3,403	2,151	5,554
Nursing home sample	1,281	903	2,184	1,548	861	2,409	2,829	1,764	4,593
Followup sample	1,181	834	2,015	1,405	757	2,162	2,586	1,591	4,177
In-community sample	974	692	1,666	1,198	625	1,823	2,172	1,317	3,489
12 Month Outcomes									
Medicare sample	1,608	1,104	2,712	1,795	1,047	2,842	3,403	2,151	5,554
Nursing home sample	1,359	935	2,294	1,577	881	2,458	2,936	1,836	4,752
Followup sample	1,052	701	1,753	1,212	658	1,870	2,264	1,359	3,623
In-community sample	838	552	1,390	974	521	1,495	1,812	1,073	2,885
18 Month Outcomes									
Number of Observations in 18-month Cohort	992	697	1,619	926	620	1,546	1,848	1,317	3,165
Medicare sample	823	592	1,415	871	501	1,372	1,694	1,093	2,787
Nursing home sample	644	475	1,119	730	399	1,129	1,374	874	2,248
Followup sample	404	281	685	471	249	720	875	530	1,405
In-community sample	310	218	528	359	195	554	669	413	1,082
NOTE: Sample sizes used in analyses were actually slightly smaller than these figures in some cases due to missing data on specific outcomes. Thus, these sample sizes differ slightly from those reported elsewhere. Some analyses based on the Medicare and nursing home samples were further restricted to sample members alive at the beginning of the analysis period. See Wooldridge and Schore (1986) for these sample sizes.									

III. STATISTICAL METHODOLOGY

Given the randomized design, unbiased estimates of channeling impacts could be obtained simply by comparing mean values of outcomes for the treatment and control groups. The approach actually used in the evaluation, however, was ordinary least squares regression. The regression models used in the evaluation, the types of statistical tests and significance levels employed, and the interpretation of estimates and tests are the topics of this section.

A. THE REGRESSION MODEL

Regression, or equivalently, analysis of covariance, offers three advantages over simple differences in means as a way of estimating program impacts. First, although the two experimental groups should have very similar average characteristics initially, there may in fact be differences between them, either by chance or because of different patterns of sample attrition for treatment and control groups. If these differences are fully reflected in the observed initial characteristics of the sample, the regression model can control for such differences between the groups. Second, the ratio of treatments to controls differs across sites, ranging from 1:1 to 2:1. If sample members differ across sites, the treatment/control differences in mean outcomes will reflect not only effects of channeling but the different distributions as well. Again, regression will control for these differences.¹² Finally, to the extent that outcomes are related to baseline or screen characteristics, regression can explain some of the variation between individuals, leading to more precise estimates of channeling impacts than are obtained from differences in means.¹³

The regression model used was

$$(1) \quad Y = a_0 + a_B T_B + a_F T_F + a_S S + a_X X + e,$$

where Y is the outcome variable that is hypothesized to be affected by channeling; T_B and T_F are binary variables equal to one for sample members in the basic (B) and financial control (F) sites; S is a set of binary site variables; X is a set of explanatory variables taken from the screen or baseline interviews; e is a disturbance term, and the a 's are coefficients to be estimated. Under this model the coefficients a_B and a_F measure the treatment/control differences in mean outcomes, controlling for any

¹² It can be shown that if a regression of outcomes on a treatment/control variable and a set of binary site variables (with no control variables) is estimated, the coefficient on the treatment variable will be a weighted average of the treatment/control differences in the individual sites, with the weight for any site dependent only on the proportion of all observations coming from the site and the ratio of treatments to controls in the sites.

¹³ Another procedure used by some analysts is multiple classification analysis (MCA), which is simply a regression model in which all of the control variables are categorical or qualitative (i.e., discrete). Because a few of our control variables are continuous, we use regression instead.

differences which exist between the two groups on baseline explanatory variables. Hence, a_B and a_F are our estimates of channeling impacts.¹⁴

The same regression model was used to estimate the impacts of channeling on all outcome measures examined in the evaluation. Although it may seem unlikely that the factors which affect well-being (for example) are exactly the same as the factors which affect nursing home use and other outcomes, there exists a strong justification for this approach. The outcome variables are highly interrelated and depend on each other as well as on many of the same exogenous variables. However, the interrelationship of the outcome variables is very complex, and trying to model it could lead to biased estimates, since some of the explanatory variables would be endogenous (i.e., correlated with the disturbance term e in the regression).¹⁵ Furthermore, we are interested first and foremost in the total effect of channeling on outcome variables, and are not particularly interested in how much of the impact on well-being (in our example) was indirectly due to channeling's effect on nursing home use and how much was due directly to the case management services provided by channeling. Therefore, we estimate the "reduced form" equation for the outcome variables. In the reduced form all explanatory variables must be exogenous (baseline/screen) variables and any exogenous variable that affects any of the interrelated outcome variables of interest is included. Thus, the explanatory variables in the reduced form include any baseline or screen variables that directly or indirectly affect the particular outcome variable being examined.¹⁶

The advantage of this approach is that we need not make arbitrary exclusions of explanatory variables from some outcome equations but not others. Including in the regression explanatory variables that do not really affect a given outcome variable, directly or indirectly, does not bias the estimates of channeling impact, and given the large number of observations available, has no discernible effect on the standard errors of the estimates. On the other hand, excluding from the regression equation explanatory variables that do affect the outcome of interest can lead to biased estimates. Thus, the reduced form approach is more likely to yield unbiased estimates of channeling impacts than would specifications in which the set of control variables assumed to affect a particular outcome variable is arbitrarily restricted. This approach has the added benefit of providing consistency across the many analyses of channeling impacts that were

¹⁴ A number of alternative specifications could be used, the most general of which would be separate regression equations for the treatment group and the control group in each site. See section D of Chapter V for a comparison of estimates obtained from the above model to those obtained from more general models.

¹⁵ There are econometric procedures to estimate such models. However, to obtain such estimates some screen/baseline explanatory variables must be excluded from some equations but not from others. If the analyst excludes from a given equation explanatory variables that truly belong, the coefficient estimates will be biased. Thus, unbiased estimates are obtained from these procedures only if the analyst correctly specifies the interrelationships among all of the endogenous variables and which exogenous variables affect which outcomes directly.

¹⁶ Inclusion of too many explanatory variables can result in a high degree of colinearity among them, which typically reduces precision and can produce anomalous estimates. However, this is unlikely to create much problem for the coefficients on the treatment variables, since random assignment ensures that treatment status will not be highly correlated with any of the regressors. Thus, the precision of the estimated treatment/control differences will not be seriously diminished by multicollinearity.

conducted by different individuals at different points in time, and associated economies in estimation through standardizing estimation programs.

The explanatory variables that were used in the regression model fell into six categories:

- Sample member's absolute level of need for assistance due to physical or mental disabilities
- The availability of informal caregivers to provide this assistance
- The amount of formal care received by sample members at baseline
- The sample member's ability to pay for additional services or nursing home care
- The availability of nursing home beds and other area-specific factors
- The sample member's -outlook on life and demographic characteristics.

These six categories of characteristics were represented in the regression model by variables obtained from the baseline and screen interviews. Need for assistance was reflected by sample members' impairment on activities of daily living (ADL) tasks (eating, dressing, toileting, mobility, bathing), continence, whether they had a recent change in health condition, whether they were cognitively impaired (i.e., whether they had behavioral problems or were disoriented), the number of unmet needs for assistance that they had and expected to continue for 6 months or more, the number of physician visits in the two months prior to baseline, and whether the individual was referred to channeling by a hospital or nursing home or home health agency.¹⁷ We also included variables indicating whether the sample member completed the baseline without help from a proxy, required some help from a proxy, or required a proxy to complete the entire baseline.

The availability of informal care was captured by two variables: sample members' living arrangement and the number of hours of care they were receiving care from visiting informal caregivers during a typical week at the time of the baseline. Living arrangement was defined by whether sample members lived alone but were receiving informal care at baseline, lived alone without such care, lived with one of their children, or lived with someone but not with their child.

The receipt of formal care is also represented by two variables: whether such care was received from visiting caregivers, and the number of hours of in-home care received from visiting formal caregivers during a "typical" week at the time of the baseline.

Sample members' outcomes also depended on the availability of hospital and nursing home beds and on other area characteristics such as the availability of formal services and case management, population density, what services the state Medicaid

¹⁷ These variables indicate whether sample members had a recent hospital or nursing home stay and serve as a proxy for serious health problems.

program covered, and any other city, state or regional differences that could affect outcomes. Since the area characteristics faced were the same for all sample members residing in a given site, binary variables indicating in which site the sample member resided were sufficient to capture the effects of any such differences across sites.¹⁸ Hence, we included 9 binary site variables in the regression.¹⁹

In addition to the amount of services received or available at baseline, another important factor affecting outcomes was the ability to pay for additional services, either in the community or in an institution. To capture these effects we included variables for whether sample members were eligible for Medicaid at baseline or would be eligible within a short period of time after entering a nursing home, based on their current income and assets.²⁰ Whether sample members were homeowners was also included as a measure of their wealth.

The attitudes of elderly individuals are also important in explaining outcomes; hence, we included the baseline measure of sample members' overall satisfaction with life. Variables indicating whether the sample members had already applied for admission to a nursing home or were in a nursing home at the screen were included, because they indicate individuals' predisposition towards institutionalization. Also included was a binary variable indicating whether the sample member had lost a close friend or relative to death within the two months prior to baseline, since major losses are felt by many to have serious effects on elderly individuals' health.

Gender, age, and ethnic background are demographic variables included in virtually every study of the impaired elderly. There may be differences between elderly men and women in ability to care for themselves, in the difficulty caregivers face in caring for them, and in the likelihood that they will have a surviving spouse to care for them. Age is included because individuals' health deteriorates with age. Furthermore, the older a sample member is, the older his or her children and friends are likely to be and the less able to provide informal care. Ethnicity was included to capture any cultural differences in the intergenerational dependency, informal support systems, or attitudes toward nursing homes of the aged.

¹⁸ The results obtained using binary site variables are exactly equal to what would be obtained if all variables (dependent and independent) were transformed into deviations around their site-specific means. Obviously, this nets out the effects of any site-specific factors on outcomes.

¹⁹ For convenience of interpretation, we actually include a binary variable indicating whether the sample member resided in a basic or financial control model site and 8 binary site variables (4 for each model). This is exactly equivalent to a specification with 9 binary site variables and no "model" variable but renormalizes the coefficients on the site variables. With the binary model variable included, the coefficients on a given site variable represents the regression-adjusted difference in mean outcomes between that site and the excluded site from the same model. With 9 site binaries, the coefficient on any one is interpreted as the differences in mean outcomes between that site and the only excluded site.

²⁰ In preliminary analyses separate variables were used for income and Medicaid eligibility. However, because of the high correlation between the two variables it was difficult to interpret the coefficients. Hence, a composite variable was defined.

TABLE III.1. Mean Values of Explanatory Variables Used in Regression Model			
Variable	Mean	Variable	Mean
Need for Assistance		Availability of Informal Care	
ADL Impairment (S)		Living Arrangement (B)	
Extremely Severe	0.233	Lives Alone, No Informal Support	0.073
Severe	0.348	Lives Alone, Informal Support	0.282
Moderate Impairment	0.223	Lives with Child	0.251
Mild or No Impairment	0.196	(Lives with Someone Other than Child)	0.377
Incontinence (S)		Missing Information	0.016
Incontinent	0.472	Hours of Care Received per Week from Visiting Informal Caregiver (B)	12.0
Needs Help with Colostomy Bag or Other Device (Continent)	0.102	Demographic Characteristics and Attitudes	
Cognitive Impairment (S)		Whether Male (B)	0.285
Severe (Moderate Impairment)	0.153	Age (B)	79.6
Mild or No Impairment	0.318	Ethnicity (S)	
Missing Data	0.471	Black	0.223
Unmet Needs (S)		Hispanic	0.037
High Unmet Needs	0.303	(White or Other)	0.740
(Moderate Unmet Needs)	0.340	Whether Currently Married (B)	0.318
Low Unmet Needs	0.302	Overall Satisfaction with Life (B)	
Missing Data	0.054	Completely	0.117
Whether Experienced Recent Change in Health (S)	0.818	(somewhat)	0.248
Whether Death of Close Friend or Relative Other Than Spouse (S)		Not Very	0.288
Death of Close Person (No Death)	0.244	Missing Data	0.348
Missing Data	0.406	Ability to Pay for Care	
Referral Source (S)		Whether Home Owner (B)	0.421
Hospital or Nursing Home	0.297	Medicaid Coverage (B)	
Home Health Agency	0.173	Currently Eligible	0.226
(Other)	0.531	Eligible Within 3 Months	0.304
Number of Physician Visits in Previous Two Months (B)	1.7	(Not Eligible in 3 Months)	0.401
Whether Waitlisted or Applied to Nursing Home, or in Nursing Home at Screen (B)	0.097	Missing Information	0.069
Type of Respondent at Baseline		Site	
Self Respondent (Mixed Proxy/Self Respondent)	0.417	Basic	
All Proxy Respondent	0.285	Baltimore	0.108
		Eastern Kentucky	0.079
		Houston	0.111
		Middlesex County	0.112
		(Southern Maine)	0.079

TABLE III.1 (continued)			
Variable	Mean	Variable	Mean
Receipt of Formal Services		Financial Control	
Whether Received Formal In-Home Care (B)	0.600	Cleveland	0.093
Hours of Formal I-Home Care Received per Week (B)	7.3	Greater Lynn	0.96
Model		Miami	0.118
(Basic)	0.488	Philadelphia	0.140
Financial Control	0.512	(Rensselaer County)	0.065
<p>NOTE: Means were computed for the Medicare sample, the largest analysis sample (N = 5,554) employing these standard control variables (see test for description of this sample). Letters in parentheses following variable names indicate whether data used were from the baseline (B) or screen (S) interviews. For variables represented by a set of binary indicators (e.g., ADL) one of the categories must be excluded from the regression to avoid perfect collinearity. Parentheses indicate which category was excluded, although this choice has no bearing on the estimates of treatment/control differences.</p>			

The means of the variables included in the model are given in Table III.1. All variables were obtained from the screen or baseline interviews, as designated in the table. Most of the variables are binary and self-explanatory. However, a few require some explanation. Impairment on activities of daily living (ADL) was defined according to sample members' most serious impairment, using the following hierarchy: eating, transfer or toileting, dressing, bathing. Thus, sample members impaired on eating were classified as extremely severe, those whose most serious impairment was transfer or toileting were severely impaired, those whose most serious impairment was dressing were moderately impaired, and others were classified as mildly impaired. Cognitive impairment was defined by whether sample members at screen exhibited behavioral problems or disorientation that required constant supervision (severe cognitive impairment), had behavioral problems that did not require daily supervision (moderate impairment), or had only mild or occasional problems with disorientation (mild or no cognitive impairment). Unmet needs was simply a count of the number of areas (0 to 5) in which the sample member needed more help and expected this need to continue for six months or more. "Change in health status" is a binary variable indicating whether the sample member reported experiencing the onset or worsening of any of several health conditions or illnesses.

Observations lacking data on one or more of the control variables were retained in the analysis by imputing values for missing variables. Data on some of the control variables were available from both the screen and baseline interviews; if data from the primary source for these variables was missing, values were imputed from the other source. Sample means were imputed in instances where no data were present on the desired variable from either the screen or the baseline, provided that less than 3 percent of the sample required imputation on that variable.²¹ If more than 3 percent of the

²¹ Four sets of mean values were computed and used for imputations, one each for treatments and controls in each model. Thus, the value imputed for any given observation depended on the treatment group and model to which the observation belonged. In addition, for variables such as hours of formal and informal care the value imputed depended upon whether the individual was known to have received some care. The conditional mean hours per recipient were imputed to those known to have received some care.

sample were missing data on a particular variable, zero values were imputed and a separate binary variable was created, indicating for which observations the data were missing on the control variable. This missing data indicator was included in the regression equation to capture differences in outcomes between those with and without available data on a particular control variable.

B. TESTING STRATEGY

The regression procedure described above provides estimates of treatment/control differences in outcomes, controlling for any initial differences between the two groups. These are our best estimates of channeling impacts. However, even if channeling had no impact the treatment and control groups may have somewhat different outcomes strictly by chance. Hence, we relied on statistical tests to determine whether the estimated differences were sufficiently large that they were unlikely to have occurred by chance.

Three types of tests were used in the analysis: t-tests on the estimates of a_B and a_F , the estimated treatment/control differences, to determine whether they were significantly different from zero; F-tests to test whether the estimated impacts in the basic and financial control models differed from each other by more than might have been expected to occur by chance; and multivariate F-tests to test whether estimates of channeling impacts on sets of related outcome measures were equal to zero. Each of these tests is described below.

1. Tests of Whether Channeling Impacts Existed

The widely used t-test simply tests whether an estimated regression coefficient differs from zero by more than might reasonably be expected to occur because of sample variation. In our application, the regression coefficients a_B and a_F estimate the treatment/control differences. If the true effect of channeling on some outcome is zero, the estimates of a_B and a_F should be relatively small. The test enables us to determine, with some known probability of error, whether channeling had some impact on the outcome examined.

Two criteria must be specified by the researcher in conducting t-tests: whether one-tailed or two-tailed tests are to be used and the significance level of the test. The choice of two-tailed or one-tailed tests depends on whether channeling is expected to affect the level of some outcome in a particular direction, or whether the impact could be in either direction. For most outcomes examined, the intention was that channeling would have a particular directional effect (e.g., reduction in nursing home use). However, for the vast majority of the outcomes, there were plausible reasons why the impact could be in the opposite direction, and for some important outcomes there was a high degree of uncertainty about the direction of channeling impacts to expect (e.g., informal caregiving and costs). Since we would clearly not ignore estimates that were of

the "wrong" sign but were large and statistically significant had a two-tailed test been conducted, the appropriate test to use is the two-tailed test.

To avoid the appearance of arbitrariness in the selection of tests and confusion in the minds of readers as to which type of test was being employed in any given table, particularly in reports covering multiple outcomes, two-tailed tests were used throughout the analysis, even for those few outcomes where the only plausible hypothesis about channeling's impact is unidirectional. The use of two-tailed t-tests also should result in greater consistency between these tests and the multivariate F-tests (described below), which are, by definition two-tailed.

The use of two-tailed tests did on occasion result in the inference that channeling's impact on a particular outcome in some time period was not significantly different from zero when a one-tailed test would have led to a different conclusion. In such cases, supporting evidence from other time periods and related outcome measures was used to obtain the correct inference about whether channeling appeared to have impacts on the behavior under examination. The magnitude of the estimate also was considered in drawing these inferences. The interpretation of coefficients and test statistics is described in more detail in Section C below.

The significance level at which to conduct the t-tests was the other testing decision. To make it relatively unlikely that chance differences between the two groups would be interpreted as channeling impacts, we followed customary conventions of statistical testing, conducting the t-tests at the .05 (5 percent) significance level. This means that based on the sample size and observed sample variation, there was a small prior probability that treatment/control differences of the magnitude estimated would have occurred by chance, and that such differences are therefore likely to be due to the effects of channeling. Tables in final reports on channeling impacts containing estimated impacts also indicated which estimates would still have been statistically significant had the test been conducted at the .01 level, implying an even smaller likelihood that the observed difference was due to chance sample variation.

Although we believe these decisions about one-tailed versus two-tailed tests and significance levels are the most appropriate, throughout the final technical reports on channeling impacts we provide the t-statistics along with the estimates. Readers can therefore determine for themselves whether and how inferences would change if alternative choices had been made.

2. Tests of Equivalence of Impacts Between Basic and Financial Control Models

In addition to determining whether the basic and financial control models affected specific outcome measures, we were also interested in knowing whether the models differed from each other in the-size of the impact. It was hypothesized that the greater resources and flexibility of funding available under the financial control model would result in larger impacts for this group. However, differences between the environments

into which the two models were introduced could also produce differences in the size of impacts achieved by the alternative models.²²

Simple F-tests of the equivalence of a_B and a_F (from the regression equation) provided the tests of this hypothesis. The tests were conducted at the .05 level, consistent with the significance level selected for the t-tests. To reduce the likelihood of inconsistencies in the test results (such as the estimate for one model being significantly different from zero and the other not, but an F-test indicating no significant difference between the two models), the F-tests were conducted in two stages. We first tested whether a_B and a_F were both equal to zero using a joint F-test. If that hypothesis could not be rejected, no further test of equivalence was necessary. If the test did indicate rejection of the hypothesis that both were equal to zero we then tested whether they were equal to each other.

3. Multivariate Tests of Whether Channeling Impacts Existed

The individual tests described above were conducted at a significance level that made it relatively unlikely that, for any particular outcome measure, chance differences between treatment and control groups would be interpreted incorrectly as channeling impacts. However, because so many outcomes were examined (each for 2 models and 3 time periods), the probability that such errors would occur in at least a few instances was very high. To lessen the probability of making such errors, multivariate tests were employed that simultaneously tested the hypothesis that channeling impacts on a set of related outcome measures were jointly equal to zero. For example, estimates of channeling impacts on nursing home days, the probability of being admitted to a nursing home, and nursing home expenditures were tested jointly to determine whether any were significantly different from zero. The advantage of this type of test is that if (for example) only one of the 6 impact estimates (3 for each model) were significantly different from zero using the individual t-tests, and the other impact estimates were all small and far from being statistically significant, it is probably unlikely that channeling really influenced nursing home use. The multivariate test in such cases would typically indicate (depending on the size and significance of the estimates) that we could not reject the hypothesis that channeling's impact on the set of nursing home outcomes was zero.

Tests that impacts on the set of outcomes being considered jointly were all zero were conducted for the basic model, the financial control model, and for both models together. We also used multivariate tests to determine whether impacts on given sets of outcomes in the basic model were equal to those in the financial control model. In each case the tests were conducted on related outcome measures, such as alternative

²² This problem could have been averted had each local channeling program implemented both models, with eligible applicants randomly assigned to the basic model, the financial model, or the control group. However, implementing both models in every channeling program would have led to serious problems as clients assigned to the basic model observed the much greater services provided to the clients who were under the financial model. Furthermore, making the complicated interagency arrangements necessary to set up the funds pool for the financial model in twice as many sites would have created a financial burden for the demonstration.

measures of well-being or informal care, for a given time period. Because the tests require that the same observations be used in all of the equations for which the coefficients are being tested jointly, outcomes in different time periods were tested separately.

The lower likelihood of erroneously concluding that channeling affected outcomes when the treatment/control differences were actually due to chance makes the use of multivariate tests attractive. Furthermore, it suggests that they should be used hierarchically, that is, that t-tests should only be examined if the multivariate tests indicate that not all channeling impacts in a given substantive area are zero. In this instance t-tests would indicate which of the outcomes channeling did appear to affect. However, strict adherence to test results in this fashion would increase the probability of making the opposite type of error--concluding that channeling had no impacts when in fact it had. The method of assessing and interpreting the many estimates and test statistics produced in the analysis, described in the section below, was designed to strike a balance between these two types of errors.

C. INTERPRETATION OF ESTIMATES

Performing statistical tests at the .05 significance level ensures a low probability of erroneously concluding that channeling affected a given outcome when the observed treatment/control difference is actually due to chance ("type I" errors). The use of multivariate tests further decreases the probability of such errors. The discussion of the power of the statistical tests presented earlier suggested that the sample sizes were sufficiently large that with t-tests performed at the .05 significance level we could be quite confident that large channeling impacts (i.e., those of policy-relevant magnitude) would not be misclassified as due to chance. However, strict adherence to the more stringent multivariate test to reduce further the probability of type I errors means that it is more likely that we will make the opposite error--concluding that channeling had no impact when the program was truly effective (type II errors).²³

Because of the desire to avoid both types of error, we do not rely solely on the hierarchical testing structure raised in the previous section, nor on any single statistic to ascertain whether channeling affected outcomes of interest. The sheer number of outcomes examined and test performed means that strict reliance on test statistics would result in a number of both type I and type II errors.

²³ This tradeoff between type I and type II errors is essentially reversed for informal care outcomes. That is, whereas the "conservative" approach for other outcomes is to ensure a low probability of erroneously concluding that there were channeling impacts where none exist (i.e., a low probability of type I errors), the conservative approach for informal care is to avoid concluding that channeling had no effect on informal care when in fact it did result in reductions in such care. This difference arises because, unlike other expected effects of channeling, the hypothesized reductions in informal care are generally regarded as adverse effects, because they imply that informal caregivers were substituting expensive formal care for their own time. Therefore, estimated reductions in informal care that were large, even if not statistically significant at the .05 level, were discussed in the report on informal impacts (Christianson, 1986).

To decrease the number of such errors, throughout the analysis we looked not only at the statistical significance of the estimates but also their magnitudes and patterns. Specifically, to assess whether channeling affected a given outcome, we looked for consistency in the direction, size, and statistical significance of estimated impacts on: (1) related outcome measures, (2) the same outcome in other time periods, and (3) the same outcome in the other channeling model.

We also examined the estimated impact at the site level, to see if the model level estimate was essentially due to one or two particular sites rather than being widespread. Dependence on patterns across model, time period, and site to verify whether impacts exist cannot be rigid, since there are reasons why effects may differ across these dimensions. Nevertheless, if patterns exist they provide evidence that the observed differences were due to channeling rather than to chance.

Finally, we also drew on theory and results from the process analysis (Carcagno et al., 1986) to assess the likelihood that the estimates obtained represented real impacts rather than chance differences. It is clearly inappropriate to conclude that only those estimates with the expected sign were due to real effects of channeling and all others were due to chance. However, awareness of how outcomes interrelate and the process by which channeling was likely to bring about effects on individuals, coupled with knowledge about how the local programs operated, were useful inputs to the assessment of whether impacts existed when the statistical evidence was mixed.

The following example demonstrates this strategy of evaluating the impact estimates. Suppose that estimated channeling impacts on nursing home days in the basic model at 6 months were statistically significant, and admissions were not significant and that the multivariate test was not significant. However, suppose that the t-statistics on these other impact estimates and the multivariate test statistics were quite near the critical values necessary for the estimates to be considered significant, and the estimates were all of the expected sign. Suppose further that the estimates for the 7 to 12 month period were also of the same sign and roughly comparable in magnitude, but not significant at the .05 level. In such cases, it would seem likely that channeling did have an impact on nursing home use, although perhaps not a strong effect or perhaps an effect that was concentrated in a few sites or in clients of a certain type. The magnitude of the estimates and effects on subgroups of sample members were examined to address these possibilities. Finally, since channeling-induced reductions in nursing home use were expected to be obtained by increases in formal community-based services, we would examine service impact estimates for confirmation. Thus, we used the collective evidence of several estimates and test statistics to determine whether channeling influenced outcomes in a given area.

In addition to ascertaining whether channeling affected outcomes, we also examined the size of the estimated impact. In general, it makes little difference whether channeling impacts in a given area were zero or just very modest in size. To obtain some indication of the proportionate magnitude of impacts, mean values of the outcome variables for the control group were presented alongside the estimated impacts on

these outcomes in tables displaying the results. Impacts exceeding 20 percent of the control group mean were generally felt to be large, although this could vary with the absolute magnitude of the control group mean (20 percent of a very small mean is still a very small impact). The dollar value of the impact also provided a useful way of assessing the importance of a given estimate for some outcome measures.

D. ESTIMATED IMPACTS FOR SUBGROUPS

In addition to determining whether channeling impacts differed by model, we also tested whether impacts on key channeling outcomes differed across sites and for various subsets of the sample defined by characteristics of the sample members. The baseline and screen characteristics used to form these subsets, described more fully in Grannemann et al. (1986), include:

- Impairment on activities of daily living (extremely severe, severe, moderate, mild/none)
- Continence (incontinent, need help with device to be continent, continent)
- Unmet needs (high, medium, low)
- Living arrangement (alone without current informal support, alone with current support, with own child(ren), with someone but not with child)
- Health system contact (in nursing home at randomization, on nursing home wait list, in a hospital or referred to channeling by a hospital or nursing home, referred by a home-health agency, referred by family or other source or self)
- Medicaid eligibility (eligible at baseline, not eligible but would be within 3 months after entering nursing home, would not be eligible)
- Cognitive impairment (severe, moderate, mild/none)
- Site

All of these characteristics were also explanatory variables in the standard regression model given in equation 1. (See Table III.1 for sample means of these variables.)

To obtain estimated impacts for the 3 to 5 subgroups formed by each of the classifying variables, the standard regression was modified as follows:

$$(2) \quad Y = a_0 + a_T T + a_1 X_1 + a_2 X_2 + a_{T1} T * X_1 + a_S S + a_{TS} T * S + e,$$

where X_1 is a vector that contains the binary variables representing the characteristics defining the subgroups and X_2 contains the other explanatory variables used in the

standard regression model.²⁴ This equation was estimated separately for the basic and financial control models, to reduce the number of parameters and simplify the calculation of impacts and standard errors.

The estimate of channeling's impact obtained from this model is

$$\text{Impact} = a_T + a_{T1}X_1 + a_{TS}S,$$

which depends on the set of 8 characteristics defining the subgroups. Estimated impacts for a particular subgroup were calculated by setting the variables in X_1 representing the classifying characteristic of interest at 1 for the category for which impact estimates were desired and 0 for the other categories of this characteristic, and setting all of the other characteristics in X_1 at the sample mean. Impacts were estimated in this way for each subgroup defined by each of the classifying variables. Standard errors of these estimated impacts were computed and used to form t-statistics to test whether impacts were significantly different from zero.

The primary tests conducted, however, were of whether the estimated impacts differed from each other across the subgroups defined by each of the classifying variables.²⁵ The hypothesis that no such difference occurred was tested by performing for each classifying characteristic an F-test of whether the coefficients in a_{T1} (or a_{TS} for tests of equivalence across sites) on the binary variables representing that characteristic were equal to zero. Given the large number of such tests, however, we first jointly tested all of the coefficients in a_{T1} to determine whether they were equal to zero. Rejection of this hypothesis indicated that channeling impacts on a given outcome did vary with at least one of the classifying characteristics. In such cases, the F-tests for each characteristic were then examined to determine with which of the characteristics channeling impacts varied. More details on the computation and interpretation of these test statistics is given in Grannemann et al. (1986).

²⁴ In order to estimate this model, binary variables for one of the categories of each of the classifying variables (X_1) were excluded from the model. (The results were unaffected by the choice of which category is dropped.) In some cases, data were missing on one or two classifying characteristics. For each of the 4 characteristics for which such missing data occurred, a separate binary variable indicating whether the necessary information on that characteristic was missing was included in X_1 . Estimated coefficients on these indicator variables were ignored; the procedure was intended solely to retain those observations with a small amount of missing information without assuming (perhaps erroneously) a value for the missing characteristic.

²⁵ An alternative way to define subgroups would involve using combinations of these 8 (or other) characteristics, e.g., individuals who live alone and are on Medicaid. Impacts for a number of multidimensional subgroups are examined in Grannemann et al (1986).

IV. SPECIFIC METHODOLOGICAL ISSUES EXAMINED

The randomized design, large samples, and straightforward estimation methodology eliminate the major reasons for questioning the evaluation results. However, despite these strengths a number of methodological issues arose that could cast doubt on the validity of the estimates obtained. The issues include:

- Actual equivalence of the treatment and control groups
- Comparability of the baseline data for treatment and control groups
- Sample attrition
- Validity of pooling observations across sites and models
- Differences between early and late cohorts of sample members
- Inappropriateness of regression procedure for some outcomes
- Effect of use of proxy respondents on impact estimates

Each of these issues was examined early in the analysis to determine whether it would lead to biased or distorted estimates of channeling impacts, and, when necessary, procedures were developed to avoid such distortions. For the major issues--comparability of the baseline data and attrition bias--separate reports were prepared (Brown and Mossel, 1984; Brown et al., 1986) that describe the analyses in detail. For the other issues, internal memoranda document the analyses. The results from these investigations are summarized below. The documents from which they were drawn are available from the author upon request.

A. THE EQUIVALENCE OF TREATMENT AND CONTROL GROUPS AT RANDOMIZATION

Due to the random assignment of eligible channeling applicants, the control and treatment groups should be composed of individuals that on average were very similar at the time of application on any observed or unobserved characteristic. Hence, the control group should yield reliable estimates of what would have happened to clients in the absence of channeling, and comparison of outcomes for the treatment and control groups therefore should yield reliable estimates of channeling impacts.

Only two factors (other than measurement error) could cause the mean values of the pre-application characteristics of the full treatment and control groups to differ: deviation from the randomization procedures and normal sampling variability. Deviations from the carefully developed randomization procedures could be either deliberate (e.g., intake workers purposely misrecording as treatments some applicants who were randomly assigned to the control group, but who had especially pressing needs for assistance) or accidental. The dedication and professionalism of the

channeling program staff at each site and the safeguards built into the assignment procedure made either occurrence very unlikely. Site staff were extremely cooperative in faithfully executing the procedures. (See Phillips et al., 1986, for details of the randomization procedures.)

Sampling variability, on the other hand, is the difference between the two groups that occurs simply by chance. For the sample sizes available at the model level, such differences between the two groups should be very small, and are expected to be statistically insignificant.

Despite the expected small, chance differences between the two groups, the implications of large chance differences for estimates of program impacts was so great that it was necessary to verify that in fact the two groups were comparable. This assessment was carried out by comparing mean values of screen characteristics for the treatment and control groups in each model, adjusting for the unequal distribution of the two groups across sites. The following screen characteristics were examined:

- Demographic: age, sex, ethnic background.
- Financial Resources: monthly income, types of insurance coverage.
- Living Arrangement: proportion in long-term care institution; proportion living alone, with spouse, with others, or with spouse and others.
- Health and Functioning (see below): activities of daily living (ADL) index, cognitive impairments affecting functioning, unmet needs for service.
- Help Received: whether help was received in the areas of meal preparation, housework or shopping, taking medicine, medical treatments at home, and personal care; expected lack of sufficient support from family and friends in coming months (fragile informal supports).
- Referral Source: whether referred to channeling by family, by a hospital, by a home health agency, etc.
- Nursing Home Application: whether had applied for admission to nursing home or were on a nursing home waiting list at screen.

Estimates of the differences between the treatment and control groups were obtained by regressing the screen characteristics on two binary variables representing treatment status (one each for basic and financial control models) and 10 binary site variables. The coefficients on the treatment variables provided the estimates of the treatment/control differences in means, controlling for the different distribution of the two groups across sites. Estimates of the treatment/control differences in means of these variables at each site were also examined. Both the model and site level differences

were tested to determine whether they were larger than could reasonably be expected to occur because of chance sample variation.

This analysis, presented in detail in Brown and Harrigan (1983), showed that there were very few variables for which treatment/control differences were statistically significant. Of the 53 screen variables examined for each model, there was only one characteristic for which differences were statistically significant in the basic model and four in the financial control model. Furthermore, even the significant differences were small in magnitude (three percentage points or less for binary variables) and (with one exception) occurred for characteristics possessed by less than seven percent of the sample. Treatment/control comparisons at the site level yielded similar conclusions: the number of statistically significant differences was no larger than would be expected by chance and no patterns of differences were found to indicate that noncomparable groups were obtained in any site.

Thus, although there may be unobserved differences between the two groups, the comparisons an observed characteristics provided no evidence of either systematic deviations from the random assignment procedures or important treatment/control differences arising by chance. We concluded that the control group provided a reliable measure of what would have happened to the treatment group in the absence of channeling, and therefore, simple comparisons of outcomes for treatment and control groups (controlling for differences in distribution across sites) should yield unbiased estimates of channeling impacts.

B. THE COMPARABILITY OF BASELINE DATA FOR TREATMENT AND CONTROL GROUPS

Another aspect of the evaluation design which could have raised questions about the accuracy of the estimates of channeling impacts was that the baseline data were collected by different types of interviewers for the treatment and control groups. The combination of several factor--conflicts between research needs and good case management practices, data collection costs, and the desire to minimize the burden on sample members--led to the decision that baseline data would be collected by channeling staff for members of the treatment group, and by research interviewers for the control group. For a variety of reasons, this difference in data collection could result in differences between the two groups on observed data for some characteristics, when in fact no real differences exist between the two groups on these baseline characteristics. Estimates of channeling impacts that are obtained from regression models which use these baseline data as explanatory variables could then be distorted, because these artificial differences between the two groups are treated as real pre-treatment differences that must be accounted for (netted out) by the regression.

Brown and Mossel (1984) conducted an extensive analysis to determine whether the baseline data for treatments and controls were comparable and, if not, what needed to be done to ensure that regression estimates of channeling impacts would not be

biased by such differences. Reasons why baseline data may differ for the two groups were identified, including:

- True differences at randomization due to chance
- True differences due to different patterns of attrition between randomization and baseline
- Spurious differences due to differences in the length of time between randomization and baseline for the two groups
- Spurious differences due to incentives of clients or their proxy respondents to overreport needs and impairments to channeling staff (who used the baseline to prepare a care plan for the client), and to underreport ability to pay for needed services
- Spurious differences due to differences between research interviewers and channeling staff in how questions were asked (including clarifications and probing), and how answers were recorded
- Treatment-induced differences due to anticipated or actual effects of channeling on the treatment group prior to baseline (and known lack of assistance from channeling for the control group)
- Spurious differences due to the differential usage of proxy respondents

As indicated in the previous section, comparison of treatment and control groups on screen variables for the full sample indicated virtually no differences outside the range of normal chance variation. Comparison of the screen characteristics of treatments and controls for baseline respondents indicated that attrition at baseline had led to very few differences between the remaining treatment and control groups. A model of baseline attrition confirmed that only for a few screen variables was the relationship between sample member characteristics and the probability of response significantly different between treatment and control groups.

Despite the overwhelming evidence, based on screen characteristics, that there were essentially no true treatment/control differences at randomization due to chance, and only minor differences due to differential attrition, Brown and Mossel (1984) found a substantial number of large and statistically significant differences between the two groups on baseline variables, including some of the same variables for which no differences were found on the screen. Although real differences between the two groups (either due to differential attrition, or to pre-existing differences not detected by screen measures) could not be ruled out entirely, they concluded that differential measurement was largely responsible for the observed baseline differences between treatments and controls. This conclusion was based on several pieces of evidence:

- The finding that very few screen variables exhibited statistically significant differences between treatments and controls among baseline respondents
- The finding that few screen variables exhibited a significantly different impact on the probability of baseline response for treatments than for controls

- The many statistically significant and occasionally very large treatment/control differences found on baseline variables, including some for which no difference was found on the screen version of the same variable
- The general correspondence of results with a priori expectations about which variables were likely to be affected by noncomparable measurement and the direction of the treatment/control differences
- The timing and proxy use differences that were known to exist at baseline and which were obviously responsible for the observed differences on some of the baseline variables and probably responsible for the differences on some others
- The general correspondence of treatment/control differences at baseline with baseline-reinterview differences observed for a subsample of treatment group members who were given a second baseline by research interviewers

Brown and Mossel then showed how regression estimates of channeling impacts would be affected by the use of noncomparable data items as explanatory variables in the regression. The expressions for bias induced by noncomparable data suggested two types of tests of baseline variables to determine whether the baseline differences were so large that it was unlikely that they represented true treatment/control differences and therefore might cause significant bias in estimates of channeling impacts, or small enough that they may well be due to chance and were unlikely to affect impact estimates. The two tests—one for baseline variables for which comparable measures were available on the screen, and one for variables that had no such screen counterparts—made use of all of the available information.

For baseline variables with screen counterparts, the test was for whether the treatment/control differences at baseline were significantly different from the treatment/control differences in the screen version of the variable for the same individuals. For those variables for which the hypothesis of no differential was rejected, the baseline version of the variable was considered noncomparable, and only the screen version was used in future analyses. Variables for which no significant differential was found were considered to be comparably measured at baseline and therefore the baseline version could be included as a control variable in later analyses. The conclusions based on this procedure were then compared to the results obtained from the reinterview sample, which were based on comparison of baseline and reinterview responses on these same questions. The two sets of results were found to be broadly consistent in terms of which variables appeared to be noncomparable, and the direction of the differences.

For baseline variables that had no screen counterpart, the procedure used was to regress these variables on treatment status, site, and the variables selected from the group with screen counterparts, and test whether the coefficients on the two treatment status variables (for basic and financial control models) were significantly different from zero. This was a test of whether there were treatment/control differences in these baseline variables beyond what could be explained by the small observed differences at

screen in a set of other variables. Variables for which this hypothesis was rejected were then considered noncomparable, under the assumption (based on the evidence cited above) that any such remaining differences were more likely to be due to noncomparable data rather than real differences. Again, the results obtained were found to be broadly consistent with the reinterview sample comparisons of baseline and reinterview responses.

The two sets of tests yielded the following conclusions regarding the comparability of the baseline variables that were used as control variables in a preliminary analysis of channeling impacts at 6 months (Kemper et al., 1985) and were then being considered for use in the final analyses:

Comparable Baseline Variables	Noncomparable Baseline Variables
Age (*)	Ethnicity (*)
Sex (*)	Income (*)
Insurance (*)	ADL (*)
Living arrangement (*)	IADL
Nursing home waiting list (*)	Unmet needs (*)
Home ownership	Attitude toward nursing home
Stressful events	Education
Hours of informal care received (per week)	Assets
Hours of formal care received (per week)	SPMSQ
Number of physician visits	Medical conditions
Global life satisfaction	Self-rating of health
	Restricted days last 2 months
	Hospital days last 2 months
	Nursing home days last 2 months
(*) indicates that a screen version of the variable exists	

Only those baseline variables found to be comparable were included as control variables in the final channeling analyses. For noncomparable baseline variables with screen counterparts, the screen version was used as a control variable in its place. The other noncomparable baseline variables were excluded from the set of control variables, with the exception of hospital and nursing home days, which were replaced with information from the screen on whether the sample member was in a hospital or nursing home at screen or referred to channeling by hospital or nursing home staff.

The exclusion of the noncomparable variables is not likely to have caused serious problems for the analysis. Estimates of channeling impacts obtained from regressions with control variables drawn only from the screen were found to be different for some outcome measures from those obtained from regressions using the (comparable and noncomparable) baseline control variables, as expected, but the standard errors of these impact estimates were virtually unaffected by this difference in regressors. Thus, the argument that increased precision would be obtained if the more complete baseline data were used as control variables was not borne out in this case. It is the case, however, that any attrition-induced differences between treatments and controls on excluded characteristics were not controlled for in estimating channeling impacts. The evidence in Brown and Mossel suggests that real differences between the two groups are likely to be considerably smaller than the observed differences in the

data. Thus, failure to control for such real differences, if they exist, is likely to have caused less bias than attempting to account for them by using control variables that were not comparably measured for the two groups. However, the inability to examine impacts for subgroups defined on potentially important but noncomparable variables such as SPMSQ, medical conditions, attitude toward nursing home, and IADL weakens that analysis.

C. THE EFFECTS OF SAMPLE ATTRITION

The experimental design of the channeling evaluation was chosen to ensure that the experience of the control group would provide a reliable estimate of what would have occurred to treatment group members in the absence of the demonstration. However, as noted above, attrition from the carefully drawn channeling sample could thwart these intentions if the sample available for analysis after attrition were not comparable for the two groups. Regression models were used in the evaluation to control for observable differences between the treatment and control groups that could arise because of attrition, but estimates may still be biased if the two groups differ on unobservable characteristics. Bias occurs if (1) those sample members for whom data are available differ on unobservable characteristics from those for whom data are not available, (2) those unobservable factors also affect outcomes of interest, and (3) rates or patterns of attrition differ for treatment and control groups.

For each of the major areas of analysis in the evaluation, an analysis sample was defined which included those observations in the research sample for which the data necessary for analysis were available. Thus, the following analysis samples were defined:

- 6/12 and 18 month Medicare samples (for hospital outcomes)
- 6, 12, and 18 month nursing home samples (nursing home outcomes)
- 6, 12, and 18 month followup samples (well-being outcomes)
- 6, 12, and 18 month in-community samples (formal and informal care outcomes)

As shown in Table IV.1, the percent of the full sample included in most of these analysis samples was somewhat greater (about 6 to 14 percentage points) for treatments than for controls, especially in the financial control model. Thus, one of the conditions that, in combination with the other two, could lead to bias was present. These differences were due primarily to treatment/control differences in response rates at the baseline interview. However, despite this difference in rates of attrition, the analysis samples exhibited only minor treatment/control differences on initial screen characteristics.

TABLE IV.1. Percent of Full Sample Included in Analysis Samples									
	Basic Model			Financial Control Model			Full Sample		
	Treatments	Controls	Total	Treatments	Controls	Total	Treatments	Controls	Total
Number of Observations in Full Sample	1,779	1,345	3,124	1,923	1,279	3,202	3,702	2,624	6,326
6 Month Outcomes									
Percent of Full Sample Included in:									
Medicare sample	50.4	82.1	86.8	93.3	81.9	88.8	91.9	82.0	87.8
Nursing home sample	72.0	67.1	69.9	80.5	67.3	75.2	76.4	67.2	72.6
Followup sample	66.4	62.0	64.5	73.1	59.2	67.5	69.9	60.6	66.0
In-community sample	54.8	51.5	53.3	62.3	48.9	56.9	58.7	50.2	55.2
12 Month Outcomes									
Percent of Full Sample Included in:									
Medicare sample	90.4	82.1	86.8	93.3	81.9	88.8	91.9	82.0	87.8
Nursing home sample	76.4	69.5	73.4	82.0	68.9	76.8	79.3	69.2	75.1
Followup sample	59.1	52.1	56.1	63.0	51.4	58.4	61.2	51.8	57.3
In-community sample	47.1	41.0	44.5	50.7	40.7	46.7	49.0	40.9	45.6
18 Month Outcomes									
Number of Observations in 18-month Cohort	922	697	1,619	926	620	1,546	1,848	1,317	3,165
Percent of Cohort Included in:									
Medicare sample	89.3	84.9	87.4	94.1	80.8	88.8	91.7	83.0	88.1
Nursing home sample	69.8	68.1	69.1	78.8	64.4	73.0	74.4	66.4	71.0
Followup sample	43.8	40.3	42.3	50.9	40.2	46.6	47.4	40.2	44.4
In-community sample	33.6	31.3	32.6	38.8	31.5	35.8	36.2	31.4	34.2

To investigate whether impact estimates based on these analysis samples were likely to be biased because of attrition, two types of analyses were performed during the evaluation and reported on in Brown et al. (1986)--a heuristic approach and a statistical modeling approach. Under the heuristic approach, Medicare data, which were available for virtually the entire research sample, were used to construct several variables measuring the amount of Medicare-covered services used, including hospital days and expenditures, nursing home days and expenditures, and several types of formal community-based and physician services. Channeling impacts on these Medicare-only variables were then estimated on the full sample, and again on the various analysis samples. These two sets of estimates were then compared to determine whether limiting the analysis to observations in the analysis samples produced different estimates than the full sample.

For the variables examined, the impact estimates obtained on the analysis samples rarely differed substantively from those for the full sample. This was especially true for the Medicare sample. Since over 98 percent of all hospital use by sample members was covered by Medicare, it was clear that attrition led to no bias in estimated impacts on hospital outcomes. For other outcomes and samples, however, this type of comparison was less compelling: although there were few instances of noteworthy differences between the full and analysis samples on the Medicare-covered variables examined, the Medicare data covered only a fraction of the total use of nursing homes and formal services and contained no information at all on other key outcomes, including well-being and informal care. Thus, it was possible that estimated impacts on these other outcomes would be biased by attrition, even though the estimates on Medicare-covered outcomes were not. Alternative procedures were required to determine whether attrition bias for these outcomes was present.

A statistical model developed by Heckman (1979) to control for the nonrandom selection of an analysis sample was used for this purpose. For each analysis sample, a model was estimated to predict which of the full sample observations were retained in the analysis, as a function of personal characteristics measured on the screening interview. Each estimated "sample inclusion" model was then used to construct for each member of the corresponding analysis sample a new variable that, when included as an additional explanatory variable in the regression equation used to estimate channeling impacts, controls for the effects of attrition. The coefficient on the constructed attrition bias term was then tested for statistical significance to determine whether the condition necessary for regression estimates to be biased by sample attrition was met.

This procedure was implemented for the 6-, 12-, and 18-month measures of the following key outcomes:

- Nursing home outcomes (nursing homes samples)
 - whether admitted
 - number of days in nursing homes
 - nursing home expenditures

- Well-being outcomes (followup samples)
 - number of unmet needs
 - number of impairments on activities of daily living
 - whether dissatisfied with life

- Formal and informal care outcomes (in-community samples)
 - whether received care from visiting formal caregivers
 - hours of formal in-home care received
 - number of visits from formal caregivers
 - whether received care from visiting informal caregivers
 - hours of care received from visiting informal caregivers
 - number of visits from informal caregivers

In general, this procedure yielded very little evidence of attrition bias. The estimated correlations between unobserved factors affecting attrition and those affecting a given outcome variable were typically small and rarely significantly different from zero. Impact estimates obtained from the regressions which included the control variable for the effects of attrition were very similar to the impact estimates obtained without this correction term.

Finally, to ensure that the results obtained from the statistical correction procedure were not distorted by overly restrictive assumptions, Brown et al. (1986) developed a somewhat more general model that would take into account two possible differences between treatments and controls and between models: differences in the relationship between observed (screen) characteristics and attrition, and differences in the covariance between unobserved factors affecting attrition and those affecting the outcome variable under examination. Use of this more general procedure showed (1) that the attrition models were not very different for treatments and controls or for basic and financial control models, and (2) that although there were some substantive differences between the 4 treatment/model groups in the correlations between unobserved factors, controlling for them separately yielded no convincing evidence that the unadjusted estimates were biased by attrition.

Although both the heuristic and statistical approaches led us ultimately to conclude that attrition bias was not a major problem, there were a number of isolated results that, if viewed alone, would have caused greater concern about attrition. To further ensure that no important evidence of attrition bias was being overlooked, the results from the heuristic Medicare data analysis were compared to those obtained from the statistical analyses for each outcome area to see if the alternative approaches both indicated that attrition bias might be a problem for any given set of outcomes. The specific patterns of attrition implied by the two approaches were also compared for consistency.

Estimates of impacts on hospital outcomes were shown conclusively to be unaffected by attrition, based on Medicare data alone. For nursing home outcomes, the Medicare comparison showed no evidence of bias in the estimates, and the only

evidence to the contrary from the statistical procedure was two cases in which impact estimates changed in statistical significance. However, in both of these instances, the impact estimates changed only marginally after controlling for the effects of attrition, going from slightly below the critical value for statistical significance to slightly above it (and vice versa). Furthermore, the results that ostensibly controlled for the effects of attrition had the implausible implication that the bias was in one direction at 6 months and in the opposite direction at 12 months, and occurred only in the basic model. Finally, some sensitivity tests were performed which showed that estimates of channeling impacts changed only slightly under a variety of different assumptions about the use of nursing homes by those with missing data. Thus, it seemed clear that estimates of impacts on nursing home outcomes were not biased by attrition, and it was virtually certain that conclusions about the lack of channeling impacts on nursing home use would not change even if some bias did exist.

For well-being outcomes, the Medicare data could provide no direct evidence concerning attrition bias, but comparison of the full and followup sample estimates of impacts on Medicare-covered services suggested that bias was potentially a problem only for the basic model, and only at six months. However, the results from the statistical procedure to measure attrition bias implied that there was no bias in any of the well-being outcome measures examined in any time period for either model.

For formal care outcomes, the in-community sample estimates of impacts on use of Medicare-covered services were very similar to the estimates obtained on the full sample in all three time periods for the financial control model, and at 12 and 18 months in the basic model. However, at 6 months in the basic model, estimated impacts on skilled nursing visits and reimbursements were statistically significant for the analysis sample but not for the full sample. This suggested that the in-community sample estimates of impacts on use of formal care might be overstated in this time period for the basic model because of attrition. However, the statistical significance of the impact estimates did not differ between the two samples for several other outcomes even in this period, nor was the magnitude of the difference that great even for skilled nursing (13 percent of the control group mean for the full sample estimate compared to about 24 percent of the control group mean for the analysis sample estimate). The lack of evidence of bias at 12 months and in the other model led us to doubt further that attrition bias was a major problem for the estimates of impacts of formal care. This conclusion was further supported by the results from the statistical analyses, which indicated an absence of the conditions necessary for attrition bias and strong similarity between impact estimates obtained using the procedure to control for the possible effects of attrition and estimates obtained without such control.

For informal care outcomes the evidence was less clear cut. The above comparison of estimated impacts on Medicare-covered services for the full and in-community samples suggested that attrition from the in-community sample used in the informal care analysis was not systematic. However, because the Medicare claims lack data on informal care outcomes, this analysis provided only weak evidence that no bias occurred in estimates of impacts on informal care. The results from the initial statistical

procedure showed no evidence of bias, but the other, less restrictive statistical approach of controlling for attrition effects led to results that implied serious bias in the estimates for both channeling models. Whereas the unadjusted results implied no effect of channeling on informal care in the basic model, and (at most) modest reductions in the financial control model, the latter adjusted estimates showed large, statistically significant reductions in informal care in the basic model and no reductions in the financial control. Also, both the Medicare and more general statistical approaches implied similar patterns of attrition, i.e., that the systematic attrition occurred mainly for the treatment group in the basic model. However, a number of factors were identified by Brown et al. which suggested that this result was a statistical anomaly rather than credible evidence of severe attrition bias. Hence, we concluded that informal care impact estimates were probably not biased by attrition either.

The two approaches used in this analysis of attrition each have their flaws. The heuristic approach of seeing how estimated impacts on some variables changed when the analysis was restricted to a subset of the full sample is appealing because it provides a direct measure of attrition bias, albeit for variables other than those in which we are most interested. Reliance on these results as proof that there is no attrition bias in the estimated impacts on those outcomes in which we are interested requires belief that any unobserved factors affecting both attrition and the outcomes of interest also affected the Medicare outcomes. Although this assumption may be plausible, it obviously cannot be verified.

The statistical approach is also appealing, but for different reasons-it pertains to precisely the outcome variables of interest, provides a direct test of whether there is bias in the estimates obtained on the analysis sample, and also offers a way to obtain unbiased estimates of impacts on any outcome. The more general model developed and used in Brown et al. adds to the attractiveness of this approach by making the results sensitive to potentially different observed and unobserved patterns of attrition for treatment and control groups. However, in either statistical model the estimates may be quite sensitive to the assumptions of the model (bivariate normal disturbance terms in the outcome and sample inclusion equations), may reflect other nonlinear relationships between the outcome and control variables that have nothing to do with attrition, and are sensitive to colinearity between the correction term and the control variables in the outcome equations.

Despite these flaws, the analyses that were conducted on attrition from the channeling sample greatly exceed what is normally done or is possible to do to examine attrition bias, because the data available from the screen and Medicare claims on nonrespondents greatly exceeds what is usually available on sample dropouts. By definition, it is never possible to know with certainty what results would have been obtained had no sample attrition occurred. The heuristic and statistical approaches were the best methods available to assess the effects of attrition on our impact estimates, and both approaches provided convincing evidence that the inferences drawn from the analysis samples about the existence and magnitude of channeling impacts were no

different from what would have been drawn had the full sample been available for analysis.

D. THE VALIDITY OF POOLING OBSERVATIONS

In selecting a regression model to estimate channeling impacts, a key issue was “pooling”--i.e., whether channeling impacts for each model could be accurately estimated by a single parameter in a single regression equation estimated on the full sample, or whether segments of the sample were so different from each other that a single equation or parameter would not accurately or adequately reflect the real relationships and would produce distorted impact estimates. Three pooling issues were examined:

- Can valid estimates of channeling impacts at the model level be obtained by treating observations from any site implementing the model as if they were all from the same site, or must separate impact estimates be obtained for each site and then explicitly averaged to obtain model impacts?
- Can a single regression equation be used to estimate channeling impacts at the model level, or are separate equations necessary for each site and/or treatment group in order to obtain valid estimates?
- Can valid estimates of impacts at the site level be obtained from a single regression equation, or are separate equations necessary for each site?

The regression model specified in Chapter III is based on the assumption that the above types of pooling are appropriate. That is, a single equation was estimated using all observations, with impacts for each model represented by a single parameter. The advantage of pooling is that if the restrictions on regression estimates implied by pooling are true, much more precise estimates (i.e., estimates with smaller variances) can be obtained because only one estimate is being made for each model rather than one for each site. The possible disadvantage of pooling is that if the implied restrictions are not true, pooling observations could produce biased and misleading estimates of the model or site level impacts. The analysis described below was conducted to determine whether the smaller variances produced by pooling observations could be obtained without distorting estimates of channeling impacts.

1. Were Separate Impact Estimates for Each Site Necessary to Accurately, Estimate Model Impacts?

The type of pooling of greatest concern for this evaluation was whether a single parameter would be sufficient to estimate the effects of a channeling model or whether impacts were so different across sites that separate impact estimates were required for

each site. In the latter case, model impacts would be obtained by computing a weighted average of the estimated impacts for the five sites implementing the model.²⁶

The restriction implicit in using a single parameter is that impacts are the same in all sites implementing a given model. This restriction was tested by estimating an unrestricted version of this--an equation with 10 site* treatment interaction terms in place of the two binary treatment status variables used in equation 1--and testing whether the coefficients on the 5 site* treatment terms involving a given channeling model were equal to each other.

This test was conducted for a set of 14 key outcome variables at 6-, 12-, and 18-months, including hospital and nursing home use (whether admitted, number of days), receipt of case management,²⁷ receipt of formal and informal care (whether received, hours received),²⁸ sample member wellbeing (number of unmet needs, number of impairment on activities of daily living, degree of global life satisfaction), and sample members' living arrangement (in community, hospital, nursing home, or, deceased). Of the 82 tests (41 for each model), the hypothesis that impacts were equal across sites was rejected in eight cases. The eight cases included whether case management was received, for both the 6- and 12- month measures in both models, and four scattered outcome measures at 18 months (for which the sample sizes were smaller by half). This is a relatively small proportion of the tests and the fact that results were strongest for case management outcomes made it less troubling, since impacts were large and statistically significant in all sites. Even more compelling was the finding that even for those eight outcomes, impacts at the model level computed from the equation yielding separate site impact estimates tended to differ little from model impacts computed from the equation without site-specific impacts. Thus, even if channeling impacts differed across sites, model level impact estimates were not distorted by the implicit assumption to the contrary in the pooled specification (equation 1). The smaller standard errors led us to prefer the pooled specification.

2. Were Separate Equations for Each Site and/or Treatment Group Necessary to Estimate Channeling Model Impacts?

Estimating a single equation on all of the observations combined implicitly constrains the estimated relationship between client characteristics and outcomes to be the same in all sites. However, if this assumption were not true, the estimated impacts at the model level from the pooled data could be distorted. The test described in the

²⁶ Another type of pooling that was considered was pooling the control groups from the two models. However, this would undo much of the benefits of randomization in that the control group would be obtained from 10 sites and the treatment group from only 5 of these sites for each model. Actual program effects would be confounded with differences among the sites in the estimates of channeling impacts. Coefficients on the binary site indicators in the regressions were nearly always large and statistically significant; hence, formal tests of whether control groups could be pooled would have failed for virtually every outcome measure examined.

²⁷ The case management measure used was not available at 18 months.

²⁸ Two measures of whether informal care was received were examined: receipt of any informal care, and receipt of care from a visiting informal caregiver.

previous section addressed only the issue of whether separate impacts for each site were required, and was based on the assumption that a single equation for all sites was appropriate. If the assumption were incorrect, the results from the above tests could be erroneous as well.

To test the implicit constraints implied by pooling observations from all of the sites, separate equations were estimated for each site and the sum of squared residuals from these regressions was compared to the sum of squared residuals from the single equation. The F-tests constructed from these two sums for each outcome variable showed that in 10 of the 41 instances, the constraints on the regression coefficients implied by pooling were rejected. However, since our concern was only with whether estimates of channeling impacts were distorted by estimating a single equation rather than separate equations, we used the site-specific equations to construct an estimate of channeling impacts at the model level²⁹ and then compared this estimate to the impact obtained from a single equation, for each of the key outcome measures. For each of the 82 comparisons the difference between the two alternative estimates was slight. Thus, despite the greater than chance incidence of formal rejection of the constraints on regression coefficients implied by pooling, the primary estimates of interest for the evaluation (channeling model impacts) were unaffected by estimation of a single equation rather than site-specific equations.

We also tested another set of restrictions that are implicit in the use of a single equation: that the relationship between outcomes and sample member characteristics were the same for treatments and controls. As always, the concern was with whether these implicit constraints, if not appropriate, would lead to different estimates of channeling impacts. Performing statistical tests of these restrictions indicated that for only 3 of the 41 outcomes examined were the implied restrictions rejected. Again, even for the 3 outcomes for which the constraints on the coefficients on explanatory variables were formally rejected, the impact estimates obtained from the separate equations were very similar to those obtained from the single equation.

Based on the above findings, we concluded that use of a single equation provided the best estimates of channeling impacts at the model level. The single equation yielded very similar impact estimates with considerably (up to 20 percent) smaller standard errors, thereby reducing the probability of erroneous inferences of the types discussed in Chapter III.

3. Can Valid Estimates of Site-Specific Impacts be Obtained from A Single Equation?

Despite the widespread findings that impacts at the model level did not seem to be distorted by pooling, there was still some concern that the site-specific impact

²⁹ Impact estimates at the model level were obtained from the separate, unpooled equations by first using the latter to compute predicted outcomes (at the sample mean of the client characteristics used in the regression) for treatments and for controls in each site, then taking a weighted average of the treatment/control difference in expected outcomes at the five sites comprising the model.

estimates to be computed (see Applebaum et al., 1986) might be distorted if they were obtained from a single equation (with site*treatment interaction terms) rather than from separate equations for each site. Comparison of the two alternative estimates showed that of the 530 impact estimates,³⁰ 438 were not significantly different from zero whether the single or multiple equation variant was used. Of the remaining 92 estimates, 65 were statistically significant under both procedures and in all but 2 of these cases the estimate was quite similar in magnitude. There were 19 cases in which the single equation estimate was statistically significant but the separate equation impact estimate was not. In over half of these cases however, the estimates were quite close in magnitude, and the insignificant estimate had t-values very close to the critical value. The reduction in standard errors achieved by pooling was the primary reason for these differences in significance. Finally, there were 8 instances in which the separate equations produced statistically significant impact estimates at the site level, but the single equation did not. In most of these cases the two estimates differed substantially in size as well as significance.

We concluded that estimates of impacts at the site level obtained from a single regression equation would only rarely yield different conclusions about channeling impacts than would the estimates obtained from the unpooled model. Furthermore, even when different it may well be the case that the pooled estimate would be preferred because the standard errors would be smaller.

E. DIFFERENCES BETWEEN EARLY AND LATE COHORTS OF SAMPLE MEMBERS

From the outset of the demonstration it was recognized that the impacts of channeling might vary with the length of time since the client entered the program, as clients' needs and health status change and as case managers and clients become more familiar with each other. However, comparing estimates of channeling impacts at 18 months to those obtained at 12 months could result in misleading inferences about such changes because, as pointed out in Chapter II, only half of the sample was followed up at 18 months, and time constraints led to defining this group as the half who entered the sample earliest. Erroneous inferences would occur if channeling's effectiveness changed with calendar time (because of specific changes in the environment in which channeling operates or in the program itself) rather than with the length of time the sample member was in the program. Alternatively, program effectiveness could change if the type of clients served by channeling changed over time. Since the 18-month cohort consists of those enrolling earliest, we must ensure that any differences between 12- and 18-month results are not due to differences in the

³⁰ The 530 impact estimates arise from examining 18 outcome variables at 6 and 12 months and 17 variables at 18 months, with impacts computed for each of the 10 sites. These 18 variables are the same at the 14 examined above, except that "living arrangements" (in the community, a hospital, a nursing home, or deceased on followup reference date) is treated here as 4 separate variables rather than as one for testing purposes, and the life satisfaction variable is treated as two variables (whether very satisfied, whether somewhat satisfied) rather than as one.

calendar period covered by the early and late cohorts or to differences between the cohorts rather than to the length of time spent in channeling.

To distinguish changes in impacts due to length of time in the program from those due to cohort effects such as those just described, estimated impacts on a set of 14 key outcomes (those used in the attrition and pooling analyses) at 6 and 12 months for the early cohort were compared to the corresponding estimates for the late cohort. Equivalence of the impacts at these earlier points would suggest that comparison of 18-month estimates obtained on only the early cohort to estimated impacts at 12 months based on the full sample should be interpreted as effects of the length of time in channeling. A finding of statistically significant differences between cohorts in impacts during the 1-6 and 7-12 month periods would indicate that 18-month results should be compared to 6- and 12-month results estimated on only the early cohort.³¹ While such cohort differences for the early periods would not necessarily imply that any differences in estimated impacts between 12 and 18 months would be due to cohort effects rather than to the length of time in channeling, it would suggest that possibility.

To investigate this issue, the standard regression model shown in equation 1 was modified in order to estimate separate impacts of channeling for each cohort on the key outcome variables listed in Section D above. The modification was to replace each of the binary treatment status variables in equation 1 with two new binary variables, the first equal to 1 only for treatment group members in that model in the early cohort and the second equal to 1 for treatments in the late cohort for that model. Two additional binary variables were also added to the regression equation, one for each channeling model, indicating whether the sample member was in the late cohort. The coefficients on the four new treatment variables provided estimates of channeling impacts for the two cohorts for each channeling model. The coefficients on the cohort indicator variables provided estimates of the differences in mean outcomes between cohorts for the control group in each model, controlling for possible differences between the cohorts on other explanatory variables.

For each key outcome measure, the revised regression equation was estimated and an F-test was performed (separately for basic and financial control models) to test for significant differences between the impact of channeling for the early cohort and the impact for the late cohort. In addition, multivariate tests were conducted on groups of related outcome measures to determine whether jointly, across the set of outcomes, impacts for the early cohort differed from those estimated for the late cohort.

The tests indicated that channeling impacts differed very little between cohorts at 6 and 12 months after randomization. Of the five instances of significantly different estimates (out of 72 tests), two were for receipt of case management at 6 months, for which the impact estimates were large, positive, and highly significant for both cohorts. Thus, even though the estimates were statistically different, the inferences to be drawn

³¹ Before comparing the impact estimates for the early and late cohorts, we compared the two groups on baseline and screen characteristics to determine whether they differed in composition prior to entering the sample. We found that the early and late cohorts differed very little for the control group, but somewhat more for the treatment group.

from the case management results were the same for both cohorts. The fact that it was changes in the control group which were responsible for the observed differences between cohorts in impacts on case management suggests that channeling may have changed relatively little, but the availability of non-channeling case management may have changed over time.

The remaining three instances of significant differences by cohort were isolated, and two of these occurred at 6 months. This is important, since it is the comparison of impacts at 12 and 18 months that we were most concerned about being distorted by cohort effects. "Whether formal care was received" was the only 12-month outcome for which a statistically significant difference across cohorts was found, and only for the basic model (although the cohort differential in the financial control model was nearly as large and had a test statistic only slightly smaller than the critical value for significance at the .05 level). The difference in impacts was due entirely to the significant difference (decline) between the early and late cohorts in the proportion of the control group receiving formal care. Whether this drop was due to different attrition of controls for the two cohorts, to changes in the types of clients attracted, or to changes in the local availability of formal services is not clear. However, the two former explanations do not seem likely given that the proportion of controls receiving formal care at baseline was very similar for the two cohorts in the basic model--57 and 55 percent for early and late cohorts, respectively. The fact that estimated impacts on hours of formal care did not differ significantly across cohorts further increased our confidence that cohort differences did not distort the comparison of 12- and 18-month impacts in general.

We concluded that estimates at 18 months on the early cohort could be compared to those at 12 months for the full sample with little concern that the comparison would be distorted by differences between the cohorts. The exception to this conclusion was that if such comparisons for formal care outcomes suggested sizeable changes in impacts between 12 and 18 months, it would be important to interpret these changes in light of the cohort difference identified here. In the final analysis of channeling impacts on use of formal community services, Corson et al. (1986) did in fact find a marked decline in impacts between 12 and 18 months, which was attributed to this cohort effect.

F. POTENTIAL PROBLEMS WITH REGRESSION ANALYSIS

Under certain statistical assumptions, the regression procedure described in Chapter III will provide unbiased estimates of channeling impacts. The assumption on which unbiasedness depends is that the disturbance term representing the unobserved factors affecting outcomes be uncorrelated with the screen/baseline control variables and treatment status. This condition is not definitely verifiable, but the fact that sample members were randomly assigned to treatment and control groups makes it unlikely that the disturbance term is correlated with treatment status; hence, estimates of channeling impacts obtained by regression are expected to be unbiased.

Unbiasedness is not the only desirable property of the estimates, however. When outcome variables are not normally distributed, regression estimates lose some of their other desirable properties and may exhibit other characteristics that are undesirable. Two types of channeling outcome variables that had non-normal distributions were those that were binary or truncated at zero, and those that were skewed (i.e., that had extremely large values for a small number of observations). Analyses were conducted to determine whether the regression estimates of impacts on these two types of outcomes were distorted or less reliable in some way than alternative estimates.

1. *The Validity of Regression Estimates of Channeling Impacts for Binary and Truncated Dependent Variables*

Estimates that are unbiased are known to be accurate on average; however, we also want impact estimates that in any particular instance are unlikely to deviate greatly from true impacts. The smaller the variance of the estimates, the narrower the confidence intervals around the estimates and the lower the probability of failing to detect important channeling impacts. However, the requirement for regression estimates to have minimum variance--homoscedasticity of the disturbance terms--will not be met for many of the dependent variables examined in the channeling evaluation because they are binary (e.g., whether admitted to a nursing home) or bounded at zero (e.g., number of days spent in the hospital). Furthermore, if the disturbance term is not homoscedastic, the test statistics calculated by the regression program will not be strictly correct. Finally, the predicted value for some observations may be less than zero when regression is used for binary or bounded dependent variables, which is obviously inappropriate. (Predicted values may also be greater than one, which is equally inappropriate for binary variables.)

For cases such as these, econometric procedures have been developed to provide estimates with desirable properties (under certain assumptions). Probit and logit models are the estimation procedures most widely used for binary dependent variables and Tobit analysis is used by economists for bounded variables. (See Maddala, 1983, for a discussion of these procedures, their statistical properties, and the assumptions on which they are based.) In practice however, these more complex and expensive estimation procedures typically provide estimates of the effects of explanatory variables on dependent variables which closely resemble in size and significance the estimated effects obtained from least squares regression. This result has been demonstrated in several previous applied studies (Corson et al., 1985; Grossman, et al., 1986; Hollister, et al., 1985; and others) as well as in the recent econometric literature (Greene, 1981, 1983). Furthermore, all of the statistical properties of the probit and Tobit estimators, including unbiasedness, depend on the assumption that the disturbance term is normally distributed, a condition not required by regression.

The much greater ease with which statistical tests can be performed with least squares regression and the much lower computational cost compared to probit, logit, and Tobit (which require iterative maximum likelihood estimation) led us to strongly prefer least squares as an estimation strategy. However, to ensure that computational

ease and cost savings were not achieved at the cost of seriously distorted impact estimates or test statistics, we compared estimates of channeling impacts obtained from regression to estimates obtained from the more complex procedures, using key outcome variables that were binary or truncated at zero.³²

Comparison of the probit model estimates to least squares estimates for binary dependent variables.³³ The probit model is based on the assumptions that individuals will take a given action (e.g., enter a nursing home) when a certain unobserved threshold is reached, that this threshold is determined by observed and unobserved factors, and that the threshold differs across individuals. Consider, for example, the decision to enter a nursing home. The probit model for this outcome is written as:

$$Y^* = a_0 + a_B T_B + a_F T_F + a_S S + a_X X - e$$

$$Y = \begin{cases} 1 & \text{if } Y^* > 0 \\ 0 & \text{if } Y^* \leq 0. \end{cases}$$

where Y^* is the unobserved indicator of the propensity to enter a nursing home, which depends on the set of variables specified as explanatory variables in the standard regression equation given in Chapter III. The disturbance term e is the unobserved individual-specific threshold, for example, the individual's unwillingness to enter nursing homes.³⁴ Sample members whose unmet need for services is so great that it outweighs their distaste for nursing homes are assumed to enter such institutions (given the availability of beds). The observed binary dependent variable (Y) is equal to 1 for those who enter nursing homes and 0 for those who do not. The parameters of this probit model (the a_i 's) are estimated by maximum likelihood, i.e., by choosing the values that maximize the product of predicted probabilities of entering a nursing home (for actual entrants) or not entering (for nonentrants). Predicted probabilities from this model will always be between zero and one, and if the assumed model is correct, the resulting estimates have the minimum variance possible. The estimated impacts of channeling are obtained by computing the predicted probability of entering for a treatment group member, with all of the other characteristics X set at the sample mean, and subtracting the predicted probability for controls computed at the same values of X .

³² Some of the regression estimates presented in this section differ somewhat from those presented in final channeling reports because of various changes in samples or variables between the early analysis performed to address methodological issues and the final analysis.

³³ Probit and logit estimates of the effects of a given explanatory variable on the dependent variable are virtually indistinguishable from each other in most applications. We have used probit in this comparison because it was somewhat easier to obtain the desired test statistics from our probit program than our logit program.

³⁴ The disturbance term is subtracted from rather than added to the equation to facilitate the interpretation of e as a threshold (see text). Obviously, the sign on e and its interpretation could both be changed with no effect on the results.

**TABLE IV.2. Impact Estimates from Least Squares Regression and from Probit for Selected Binary Outcome Measures
(In percentage points; t-statistics in parentheses)**

	Basic Model				Financial Control				Sample Size
	Regression		Probit ^a		Regression		Probit ^a		
Whether Received Any Formal Care - 6 months	6.96**	(3.49)	7.35**	(3.49)	16.31**	(8.09)	17.23**	(8.12)	4,974
Whether Had Any Visiting Informal Caregiver - 6 months	-2.33	(-1.22)	-2.34	(1.18)	-2.57	(-1.33)	-2.77	(-1.33)	4,899
Whether Received Any Informal Care - 6 months	-2.97	(-1.50)	-3.12	(-1.44)	-2.64	(-1.32)	-2.92	(-1.38)	4,899
Whether Received Comprehensive Case Management - months 1-6	51.17**	(26.33)	52.67**	(26.44)	56.34**	(28.93)	58.35**	(29.36)	3,955
Whether Admitted to Hospital -									
months 1-6	-2.80	(-1.44)	-2.93	(-1.47)	2.04	(1.04)	2.12	(1.07)	5,554
months 7-12	-0.36	(-0.20)	-0.43	(-0.23)	0.37	(0.20)	0.48	(0.26)	5,554
Whether Admitted to Nursing Home -									
months 1-6	-0.52	(-0.37)	-0.20	(-0.15)	-0.37	(-0.27)	-0.16	(-0.12)	4,593
months 7-12	-2.23	(-1.88)	-2.22	(-1.93)	0.29	(0.25)	0.40	(0.36)	4,752

NOTE: Regression estimates and sample sizes do not in all cases correspond exactly with those presented in final channeling reports, because some changes may have taken place between the time that this analysis was conducted and the final analyses were completed.

- a. Estimates of channeling impacts were obtained from the probit coefficients by computing the predicted probability of the dependent variable for treatments and for controls (with all of the explanatory variables set at their overall sample means) and subtracting. Thus, $\text{impact} = F(Xb + a) - F(Xb)$, where F is the cumulative normal distribution function, X is the mean of the explanatory variables for treatments and controls combined, b is the vector of estimated probit coefficients on the explanatory variables, and “ a ” is the estimated probit coefficient on the treatment status indicator. The standard error of this difference was then calculated using the usual formula for approximating the variance of a nonlinear a combination of estimators. (Kmenta, 1971; p. 444). The t-statistics is simply the ratio of the estimated impact to the estimated standard error of the impact.

** Significantly different from zero at the .01 level (2-tailed test).

The least squares and probit estimates of channeling impacts on a set of key binary outcome variables are compared in Table IV.2. The impact estimates and t-statistics were very similar for all six of the variables examined, for both models. For no outcome was there a change in the statistical significance when probit was used. Even estimates that were statistically insignificant exhibited only small changes in magnitude.

Comparison of Tobit estimates to least squares regression estimates. When the dependent variable is truncated at zero but not binary, such as nursing home expenditures or days, regression estimates lose some of their desirable properties. The Tobit procedure, which is closely related to the probit procedure, was designed to overcome these weaknesses. A Tobit model of the number of days spent in nursing homes, for example, would be written as:

$$Y^* = a_0 + a_B T_B + a_F T_F + a_S S + a_X X - e$$

$$Y = \begin{cases} Y^* & \text{if } Y^* > 0 \\ 0 & \text{if } Y^* \leq 0. \end{cases}$$

where observed nursing home days (Y) is equal to the expression given for Y* for individuals whose need for nursing home care outweighs their unobserved unwillingness to enter nursing homes (e), and equal to zero for others. Again, maximum likelihood methods are used to estimate the coefficients and the standard error of e. The effects of channeling are estimated by computing the expected value of the outcome Y for treatments and for controls, both at the point of means of the other explanatory variables, and taking the difference. (See Moffitt and McDonald, 1980, for the correct expression for obtaining predicted outcomes from Tobit models.)

The regression and Tobit estimates of channeling impacts on a set of key outcome variables that are bounded at zero are contained in Table IV.3. For most of the 24 comparisons, the differences between the two alternative estimates were quite small (though somewhat greater than the differences observed between probit and regression). However, in 3 instances, the differences were fairly large and resulted in a change in the statistical significance of the impact estimates: hours of formal care at 6 and 12 months in the basic model and nursing home expenditures at 6 months in the basic model. The impact of channeling on formal care in the basic model went from essentially zero using the regression model to nearly 1 hour per week at 6 months (about 15 percent of the control group mean) using the Tobit model, with the latter being statistically significant at the .05 level. The same change in statistical significance occurred at 12 months for this outcome in the basic model, although the two estimates were not that different in magnitude. The effect on nursing home expenditures went in the opposite direction. The regression estimate was a reduction of 165 dollars (about 25 percent of the control group mean), which dropped to 47 dollars when Tobit was used.

**TABLE IV.3. Impact Estimates from Least Squares Regression and from Tobit for Selected Truncated Outcome Measures
(t-statistics in parentheses)**

		Basic Model				Financial Control				Sample Size
		Regression		Tobit ^a		Regression		Tobit ^a		
Hours of Formal Care										
6 Months:	impact	0.14	(0.22)	0.92*	(2.00)	5.35**	(8.15)	5.09**	(9.99)	4,974
	control mean ^b	6.4		6.2		4.8		6.3		
12 Months:	Impact	1.14	(1.78)	1.46**	(3.38)	3.58**	(5.56)	3.39**	(6.62)	5,040
	control mean	5.2		4.8		4.5		6.0		
Hours of Informal Care										
6 Months:	impact	-0.98	(-1.29)	-0.74	(-1.42)	-0.31	(-0.41)	-0.59	(-1.02)	4,899
	control mean	6.02		6.27		6.31		7.08		
12 Months:	Impact	-0.03	(-0.04)	0.08	(0.21)	0.07	(0.12)	-0.29	(-0.62)	4,998
	control mean	3.69		3.96		4.56		5.15		
Hospital Days										
6 Months:	Impact	-0.35	(-0.41)	-0.59	(-0.83)	-0.71	(-0.83)	-0.00	(-0.01)	5,554
	control mean	11.5		12.8		16.2		14.3		
12 Months:	Impact	-0.18	(-0.25)	-0.20	(-0.33)	-0.56	(-0.75)	-0.20	(-0.33)	5,554
	control mean	7.0		8.1		9.0		8.6		
Nursing Home Days										
6 Months:	impact	-2.36	(-1.93)	-0.59	(-0.67)	-1.14	(-0.94)	-0.27	(-0.33)	4,593
	control mean	12.2		6.4		9.6		5.6		
12 Months:	Impact	-1.19	(-0.63)	-2.56	(-1.59)	-2.19	(-1.15)	-0.02	(-0.02)	4,752
	control mean	16.3		12.8		16.7		10.1		
Hospital Expenditures										
6 Months:	Impact	-119	(-0.45)	-206	(-0.94)	-68	(-0.25)	89	(0.36)	5,554
	control mean	3,412		3,869		4,899		4,643		
12 Months:	Impact	59	(0.29)	-11	(-0.06)	-161	(-0.79)	-63	(-0.34)	5,554
	control mean	2,015		2,307		2,706		2,641		
Nursing Home Expenditures										
6 Months:	impact	-165*	(2.15)	-47	(-0.92)	-8	(-0.11)	6	(0.12)	4,593
	control mean	666		369		560		332		
12 Months:	Impact	-58	(-0.56)	-120	(-1.42)	-103	(-0.99)	1	(0.01)	4,752
	control mean	819		657		894		546		

TABLE IV.3 (continued)

NOTE: Regression estimates and sample sizes do not in all cases correspond exactly with those presented in final channeling reports, because some changes may have taken place between the time that this analysis was conducted and the final analyses were completed.

- a. Estimates of channeling impacts were obtained from the tobit coefficients by computing the predicted value of the outcome variable for treatments and controls (with all of the explanatory variables set at their overall sample means) and subtracting. Using the expression given by Moffitt and McDonald (1980) for the expected value of the dependent variable in a tobit model, the estimated impact was:

$$\text{Impact} = (\bar{X} b + a) * F((\bar{X} b + a)/a) = s * f((\bar{X} b + a)/s) - [\bar{X} b * F(\bar{X} b/s) + s * f(\bar{X} b/s)],$$

where \bar{X} is the mean of the explanatory variables for the treatment and control groups combined; b and a are the estimated tobit coefficients on the explanatory variables and treatment status indicators respectively; s is the estimated standard error of the disturbance term in the tobit model; f(.) is the standard normal density function; and F(.) is the cumulative distribution function of the standard normal (the predicted probability that the dependent variable is greater than zero). The standard error of the estimated impact was calculating using the usual formula for approximating the variance of a nonlinear combination of estimators (Kmenta, 1971: p 444). The t-statistic (in parentheses) is simply the ratio of the estimated impact to the estimated standard error of the impact.

* Significantly different from zero at the .05 level.

** Significantly different from zero at the .01 level.

Despite these differences, it was not clear that the Tobit procedure produced better estimates than regression even in these two instances. The predicted nursing home expenditures for controls was far below the actual mean, suggesting that Tobit may not have provided reliable estimates. Furthermore, for both the variables for which least squares and Tobit produced substantially different estimates there was evidence that the Tobit estimates reflected the probability of any use of these services more strongly than the extent of use. Both of these problems were due to outliers, cases with extremely large values of the outcome variable, which affect Tobit estimates somewhat differently than least squares estimates. Although less sensitivity to outliers would be a desirable feature, the distorting effects of outliers on Tobit estimates may be even greater than their effects on least squares estimates, especially if there are treatment/control differences in the number of outliers. These potential problems, combined with the greater expense and difficulty of hypothesis testing with the Tobit model, again led us to prefer least squares regression as the estimation procedure, and to analyze the effects of outliers on these estimates directly.

2. The Effects of Outliers on Regression Estimates of Channeling Impacts

The effects of outliers (i.e., extremely large values of the outcome variable that are not simply data errors) on estimates of population means and regression coefficients are well-known, but there is much less documentation about what should be done when confronted by such problems. A common "solution", discarding the outliers, may distort estimates of program impacts more than leaving them in, since one of the effects of the program may be to reduce extreme use of or expenditures on services. This effect would be totally missed if outliers are discarded. However, it may be the case that differences between the two groups in the very small proportion of outliers could arise strictly by chance and affect the estimated treatment/control difference so greatly that it no longer provides a reliable estimate of channeling impacts.

Duan et al. (1983) cite examples of -how even estimates which are unbiased can yield very misleading inferences about program impacts in cases where the outcome variable is zero for a substantial fraction of the sample but has extremely large values for a small fraction of the remaining cases. They then propose an alternative estimator for such situations. This procedure seemed potentially appropriate for the channeling evaluation, since several of the key outcome variables exhibit these characteristics, especially hospital and nursing home days and expenses.

The procedure advocated, by Duan et al. is to break such service use variables (measured either in physical units or expenditures) into two separate variables: whether the service is used at all, and for those who use it, the amount of such services. The expected value of use is the product of the probability of use and the expected amount of use given that some occurred. Thus, a probit model is estimated first for whether any use occurred, as a function of treatment status and other explanatory variables. Then, using only observations that had some service use, a regression model is estimated to predict the amount of use (again dependent on treatment status and control variables), with the amount being expressed in logarithmic form to reduce the influence of outliers

on the estimates. These two equations are then used to obtain predicted probabilities of use and amounts of use by service users for treatments and for controls with the same characteristics. These estimates in turn are used to compute overall expected use for the treatment and control groups and the difference between them.

This procedure was used on a set of key hospital and nursing home outcome variables with skewed distributions. Table IV.4 contains a comparison of the 2-part, least squares and Tobit estimates of channeling impacts. The 2-part method yielded estimates which differed somewhat from the regression estimates, but not by enough to change the inference about whether channeling affected hospital and nursing home outcomes. The 2-part estimates were also generally closer to the least squares estimates than to the Tobit estimate, especially for the outcomes exhibiting the largest discrepancy between least squares and Tobit.

These results suggested that the more cumbersome two-part method was not necessary, at least for hospital and nursing home outcomes where outliers were most likely to occur. However, the results from the Tobit analysis suggested that estimates of channeling impacts on hours of formal care received at 6 months was also affected by outliers. To investigate this, the 2-part method was used for this outcome variable as well. In the financial control model, estimated impacts from least squares and the 2part methods were both large and statistically significant. In the basic model, however, the estimated impact from regression was small (.14 hours) and not statistically significant, but the 2-part method estimate was much larger (2.5 hours) and the impact on both the probability of receiving care and the amount of care received by service recipients were statistically significant.

The nonsignificant effect on hours was unexpected because other estimates indicated that the basic model led to an increased proportion of sample members receiving any services. Thus, to have no effect on hours channeling would have had to decrease the average amount of services received by those who would have received some services even in channeling's absence. Further examination of the data showed that the small regression estimate of treatment/control differences was heavily influenced by the receipt of continuous (24 hours per day) formal care by 7 control group members (representing 20 percent of total use by the 1,000 controls in the sample) but only 2 treatment group members. Use of the 2part method dampened the effect of these outliers on the estimated treatment/control difference, and completely reversed the inference about channeling's effects on the average amount of care received by recipients. The estimate in column 7 of Table IV.4 indicates that treatment group recipients received significantly (2.8) more hours of care than recipients in the control group.

TABLE IV.4. Comparison of Least Squares, Tobit, and 2-Part Estimates of Channeling Impacts for Skewed Outcome Variables

Outcome	Alternative Estimates of Impacts			Control Group Mean	Components of 2-Part Method Estimate				Sample Size
	Tobit	Least Squares	2-Part Method ^s		Probability of Use		Quantity of Users		
					Impact	Control Mean	Impact	Control Mean	
6 Month Outcomes									
Hospital Days									
Basic	-0.59	-0.35	-0.74	11.5	-0.024	0.539	-0.4	22.19	5,554
Financial Control	0.00	-0.71	-0.77	16.2	0.018	0.546	-2.3	29.03	
Hospital Expenditures									
Basic	-206	-119	-227	\$3,412	-0.024	0.539	-131	6,632	5,554
Financial Control	89	-68	-178	\$4,889	0.018	0.546	-596	8,813	
Nursing Home Days									
Basic	-0.59	-2.36	-2.42	12.2	-0.004	0.113	-19.2*	81.30	4,593
Financial Control	-0.27	-1.14	-0.08	9.6	0.001	0.107	-1.4	68.37	
Nursing Home Expenditures									
Basic	-47	-165*	-131	\$666	-0.004	0.113	-1035	4,521	4,593
Financial Control	6	-8	-30	\$560	-0.001	0.107	-320	4,158	
Hours of Formal Care									
Basic	0.92*	0.14	2.50*	6.50	0.074**	0.400	2.82*	16.24	4,974
Financial Control	5.09**	5.35**	8.41**	5.02	0.172**	0.474	10.20**	10.60	
6 Month Outcomes									
Hospital Days									
Basic	-0.20	-0.18	0.40	7.0	-0.005	0.339	1.5	21.06	5,554
Financial Control	-0.20	-0.56	-0.44	9.0	-0.0003	0.350	-1.2	25.17	
Hospital Expenditures									
Basic	-11	59	139	\$2,015	-0.005	0.339	506	6,079	5,554
Financial Control	-63	-161	-132	\$2,706	-0.0003	0.350	-370	7,597	
Nursing Home Days									
Basic	-2.56	-1.19	-0.78	16.3	-0.025	0.129	19.3	111.41	4,752
Financial Control	-0.02	-2.19	-2.43	16.7	0.004	0.103	-27.5	128.66	
Nursing Home Expenditures									
Basic	-120	-58	-4	\$819	-0.025	0.129	1,345	5,757	4,752
Financial Control	1	-103	-124	\$894	0.004	0.103	-1,420	6,910	

TABLE IV.4 (continued)

a. The impact estimate obtained from the two-part method was calculated as follows:

$$\begin{aligned} \text{Impact} &= (\text{proportion of control group with } Y > 0 + \text{estimated channeling impact on proportion}) \\ &* (\text{average value of } Y \text{ for control group members with } Y > 0 + \text{estimated impact on } Y \text{ for those with } Y > 0) \\ &- (\text{proportion of controls with } Y > 0) * (\text{average } Y \text{ for controls with } Y > 0). \end{aligned}$$

where Y is the value of the outcome variable examined. The impact on the proportion for sample members with Y > 0 was estimated from a probit model. The impact on outcomes for those with Y > 0 was estimated by first regressing the logarithm of the outcome variable on binary treatment indicators and the standard control variables, using only those cases with Y > 0. The coefficients (b) on the treatment status variables from this log regression were then used to calculate impacts on expenditures:

$$\text{Impact on those with } Y > 0 = (e^b - 1) * (\text{control group mean for those with } Y > 0).$$

These four components used to construct the overall impact are presented in columns 5 through 8 of this table.

* Significantly different from zero at the .05 level (2-tailed test).

** Significantly different from zero at the .01 level (2-tailed test).

Given the similarity of the 2-part estimates to the ordinary least squares regression estimates for nursing home and hospital days and expenditures, the final reports on these outcomes relied upon the ordinary regression results. This was done because the standard errors of impacts from the 2-part method are more cumbersome to calculate, and multivariate tests would be especially difficult to conduct. Even for hours of formal care, we chose in the final reports to rely on least squares estimates (computed both with and without the outliers), despite the fact that the 2-part method did yield estimates that were less sensitive to outliers than the ordinary least squares estimates. The reason for this decision was that if channeling did in fact reduce the service use of a small number of cases who would otherwise have used large amounts of services, the savings from such effects could be very substantial. The two-part method may understate the importance of such cases.

The 2-part method therefore may never give the most appropriate estimates. If important channeling effects occur for outliers, the two-part method may mask them. On the other hand, if treatment/control differences in outliers were due strictly to chance, the optimal approach is to drop them, rather than to just reduce their influence. Thus, throughout the evaluation, least squares regression was used to estimate channeling impacts. As shown in Table IV.4, this yields the same inferences about impacts on hospital and nursing home outcomes as the 2-part method. For formal care at 6 months, impacts were estimated in the final report with outliers included and then with them excluded. Evidence was presented indicating which estimates provided the most accurate indication of channeling impacts. (See Corson et al., 1986 for further discussion of those results.) No other outcome measures appeared to have skewed distributions; hence, no other analyses of the effects of outliers were conducted.

G. THE EFFECTS ON IMPACT ESTIMATES OF USING PROXY RESPONDENTS

Because of the frailty of the sample, many sample members required the help of others (family, friend, nurse, caregiver) to complete the interview. However, proxies' responses to questions may differ considerably from those that the sample members would have given, especially to questions about attitudes or feelings. This issue raised concerns from the beginning of the evaluation about whether use of proxies at followup would distort our estimates of channeling impacts.

In order for proxy use at followup to bias impact estimates, it must be true that proxies for either the treatment group, the control group, or both respond differently than sample members would. There are three ways in which proxy use at followup could affect impact estimates:

1. If proxies over- or underreported (relative to sample members) to the same extent for treatment and control groups, but rates of proxy use differed for treatment and controls.

2. If proxies for the treatment group over- or underreported more or less than did proxies for control group members (whether rates of proxy use differed or not).
3. If proxies over- or underreported to the same extent for both groups and rates of proxy use were the same. (In this case, the bias will be proportional because if the dependent variable mean is, say, overstated by a certain proportion for both treatments and controls, then the treatment/control difference is overstated by the same proportion.)

Of these, the first was considered to be the most likely to occur, and the second the least likely. The third situation would be clearly less serious than the other two, since proportional misreporting for both treatments and controls implies that impacts expressed as a percent of the control group mean will be unaffected. Therefore, we looked first at rates of proxy use and compared them for treatment and control groups, and then we compared impact estimates for self-respondents and proxy respondents.

Rates of proxy use for treatment and control groups were remarkably similar for the two groups at all 3 followup interviews, both in answering specific questions and in overall response to the interview. Overall, about 40 to 45 percent of the interviews were completed without any assistance from proxies, while another 40 to 45 percent were completed entirely by proxies. For 45 to 50 percent of the sample members, a proxy answered the specific interview questions about the sample member's attitudes about satisfaction and contentment with life and with service arrangement.

The similarity of rates for the two groups made it, less likely that proxy use distorted estimates of channeling impacts. However, it was still possible, unless proxies responded no differently from sample members on average. To examine this question, the mean responses of proxies and self-respondents to several key questions at followup were compared. These comparisons showed that sample members with proxy respondents were recorded as being more impaired (on ADL and IADL tasks), less satisfied with life, and lonelier than sample members who responded themselves. However, examination of records data showed that sample members requiring proxies also had many more hospital and nursing home days, which suggests that the reported differences on interview items between those with and those without proxies may be real differences rather than the result of differential reporting by proxies and sample members. However, this conclusion could not be drawn without direct investigation of the effect on impact estimates of using proxy respondents.

To provide an indication of whether impact estimates were affected by proxy use, we estimated impacts on key outcomes separately for sample members with proxy respondents and those who responded themselves. We did this by modifying the standard regression model, replacing the binary treatment variables (T) with interaction terms (T* respondent type), then testing to see if impacts (the coefficients on T* respondent type) were equal.

We found relatively few significant differences in impacts (16 out of 90) between these two groups, but more than would be expected by chance. For impairment/health status outcomes (ADL, IADL, hospital days, nursing home days) we found a few significant differences but no systematic pattern. Among the formal and informal care measures, we found statistically significant differences in impacts across types of respondents only for the outcome variable indicating whether any informal care was received. The treatment group had a significantly lower proportion receiving informal care (from visiting caregivers or from anyone) than the control group among self respondents, but not among proxy respondents. However, it was unclear whether this difference was due to differences in physical or cognitive impairment between the types of clients who required proxy respondents and those who did not or to responses by proxy members that were not accurate reflections of what the sample members would have given themselves.

Six variables measuring sample members' attitudes were also examined, including their loneliness, overall satisfaction with life, confidence about receipt of care, contentment, self rating of health, and degree of concern about receiving needed care. Again we found relatively little difference in impacts across respondent types, except for the global life satisfaction variable. Among sample members with proxy respondents, the proportion reporting low satisfaction at 6 and 12 months was significantly smaller for the treatment group than for controls in both models, but no such pattern occurred for self respondents. Again, the relevant question was whether these results were due to differential reporting by proxies, or whether they, perhaps reflected the fact that proxy users were the most impaired (and presumably, least satisfied initially) and channeling may have had the biggest impact on the morale of those who were originally the most impaired/least satisfied (perhaps because they were not receiving needed services).

To distinguish between these two alternative explanations for the differences in impacts between self and proxy respondent cases, the regression model used to estimate impacts for the two groups was modified by including additional interaction terms involving treatment status and baseline measures of other factors that could affect channeling impacts. These factors were ones that were used in the analysis of channeling impacts on particular subgroups (see Chapter III): ADL, continence, unmet needs, referral source, Medicaid eligibility, living arrangement, whether on a nursing home waiting list, cognitive impairment, and site. Respondent type was added to this model as an additional set of subgroups. If the apparent differences in impacts across proxy use categories observed for informal care and global life satisfaction were in fact due to differences in impacts across impairment levels, impacts estimated from the revised subgroup regression for these two outcomes should no longer differ significantly across proxy use category, because the differences in channeling's effects across impairment subgroups would now be controlled for.

Once these other interactions were entered, impacts on informal care were no longer significantly different across types of respondents. Thus, it appeared that for informal care, proxy use did not affect impact estimates.

For the outcome variable representing sample members' satisfaction with life, however, the difference in impacts by respondent type remained statistically significant. Differences without controlling for subgroup effects were statistically significant for both models at 12 months and for the financial control model at 6 months. After controlling for other subgroup effects, only the 6-month basic model results indicated- significantly different impacts by respondent type. However, it was clear that in all three cases, the overall significant improvement in life satisfaction was driven by the treatment/control difference for those with proxy respondents. Thus, for this outcome the difference in impacts across types of respondent were not merely reflecting impact differentials across baseline impairment or unmet need categories.

From the set of analyses conducted we concluded that with one possible exception the use of proxy respondents did not result in distorted estimates of channeling impacts. The potential exception to this was the result for life satisfaction, for which it was difficult to distinguish between two plausible alternatives. It is possible that, as caregivers, proxies for treatment group members were so pleased with the additional help channeling provided that their response reflected the proxy's own satisfaction more than that of the sample member. On the other hand, it may have been the case that sample members requiring proxies at followup were those most dissatisfied with life at baseline and it was this dissatisfied group for which channeling had the biggest effect on reported life satisfaction. Yet another possible explanation is that those who required proxies at followup but were not highly impaired at baseline may be the group whose health or ability to function deteriorated, the most over the six months. Channeling impacts on satisfaction could be greatest for this group. In any case, however, channeling appears to have had an impact on satisfaction. Whether these impacts were for a certain set of sample members or for the caregivers of those sample members is unclear.

V. CONCLUSION

A great deal of care went into the design of the sample and estimation procedures used in the channeling evaluation to ensure that the validity of the results would not be subject to doubt. There were a few uncertainties inherent in the design of the demonstration that lead to questions about the generalizability of the results, including:

- Whether the sites in which the demonstration was conducted were representative (they were not randomly selected, were heavily concentrated in the eastern part of the country, and appeared to have somewhat fewer nursing home beds and more community services available than most areas)
- Whether the sample members were representative of the population at high risk of institutionalization (sample members may have been more strongly opposed to living in nursing homes or may have had better informal support systems than comparably impaired community residents who did not apply to channeling)
- Whether the limited duration of the demonstration affected either the types of individuals who applied or were referred to channeling or the effectiveness of the program
- Whether following sample members for only 12 or 18 months was too short a period to observe the effects of channeling (channeling impacts on outcomes such as institutionalization may only become apparent only over a period of years)

We do not believe that the limited duration of the demonstration or the shortness of the followup period led to misleading inferences about channeling impacts. At the time they were referred or chose to apply to channeling, many of the sample members were at a critical point where they needed assistance, either because they had just left or were about to leave a hospital or nursing home, or because they had an informal support system that was inadequate for their current needs or in danger of collapse. Furthermore, a large fraction of sample members (about 40 percent) were admitted to a hospital during the first 6 months after randomization. These are precisely the circumstances in which the case management and services provided by channeling were expected to benefit clients. Thus, the argument that the effects of the program were unlikely to be observed during the 12 or 18 months immediately after entering the program has little credence. Nor is it likely that individuals who potentially would have benefitted greatly from participation in channeling neither applied nor were referred to channeling because it was only guaranteed to last a few years.

The basic design questions of the representativeness of the sites and applicants, on the other hand, are potentially more serious. However, they cannot be answered with the available data. Whether program impacts would be different in different

environments or for different types of individuals than those analyzed are questions that can rarely be answered in any evaluation.

The potential effect of the local environment on channeling impacts raises another important limitation of the design: comparison of impacts for the basic and financial control models confound such environmental effects with effects of differences between the two models. Since one of the goals of the evaluation was to determine whether the additional resources and control over funds available to case managers in the financial model led to greater impacts, this was a potentially serious shortcoming. However, because of the general lack of significant impacts for either model, the importance of this limitation is muted.

In addition to these design features, another potential limitation of the evaluation was the fact that the estimates of the effects of channeling were attenuated in two ways: first not all treatment group members actually participated in channeling, and second, some controls received case management from other agencies or sources, and many controls received community-based services. While the case management may not have been as comprehensive as that offered by channeling and the services may not have been as extensive, control group receipt of services means that the evaluation did not provide estimates of the absolute effects of the case management and services provided by channeling but rather of how the effects of channeling compare to the effects of services that were already in existence. Furthermore, estimated impacts will understate the actual program effects on participants, because the outcomes for the treatment group are averaged over all treatment group members, including the 20 percent who never had a care plan completed.

These dampening effects have two implications for the interpretation of impact estimates. First, the impacts of channeling on those who actually received the treatment were 25 percent larger than the estimated treatment/control differences (as are the average costs of providing these services). However, since relatively few of the impacts were significantly different from zero, and costs were also measured on a per treatment group member basis, this has little importance for the evaluation. Second, and more important, the general lack of statistically significant impact estimates does not mean that case management and formal services in general or the channeling program in particular have no effects on impaired elderly people. A separate report (Brown and Phillips, 1986) addresses the broader issue of the effects of case management and services per se on nursing home use.

What we have focused on in this report are issues that were known about at the beginning or that arose during the course of the evaluation and which could be mitigated by analytic methods. We have shown that given the basic design, the estimated treatment/control differences provided valid, robust estimates of channeling impacts. The random assignment of eligible sample members to treatment and control groups ensured that the comparison of the two groups provides an unbiased estimate of the difference between treatment group members' actual outcomes and what would have happened to them in the absence of the program. The large sample sizes,

statistical testing procedures, and methods of piecing together evidence across time periods, models, and outcome measures provide a high degree of confidence that the evaluation neither concluded that channeling influenced some outcome when no impact actually occurred nor failed to detect important impacts that did occur.

A variety of potential threats to the validity of the results were identified and assessed including sample composition issues (whether the treatment and control groups had similar initial characteristics and whether the expected equivalence of the two groups was distorted by sample attrition), data issues (whether the differences between the two groups in who collected the baseline data led to differential measurement of those data, and whether the use of outcome data collected from proxy respondents distorted estimates of channeling impacts), and estimation issues (whether observations from all sites and from both treatment and control groups should be combined to obtain a single regression estimate of channeling impacts for each model, whether impacts of channeling were the same for the early and late cohorts of sample members, and whether regression provided robust estimates of impacts for the outcome variables that were not distributed normally). Of all these potential problems, only the noncomparability of the baseline data was determined to be likely to distort estimates of program impacts. To avoid this distortion, baseline variables judged to be noncomparably measured were excluded from use as control variables in the regression equation. (Where they existed, screen counterparts to these noncomparable baseline variables were used as substitutes.) All of the other potential problems with the data or regression estimation approach were found to have little or no actual effect on impact estimates or their interpretation, so it was not necessary to implement special procedures broadly. The isolated cases in which there was some evidence of a potential problem for specific outcome variables were identified and examined in detail in technical reports dealing with those outcomes, and where appropriate, alternative estimates were presented.

REFERENCES

- Applebaum, Robert, Robert S. Brown, and Peter Kemper. "Evaluation of the National Long Term Care Demonstration: An Analysis of Site-Specific Results." Channeling Evaluation Supplementary Report 86-05, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Brown, Randall S., et al. "Final Report on the Effects of Sample Attrition on Estimates of Channeling's Impacts." Channeling Evaluation Technical Report TR-86B-13, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Brown, Randall S., and Margaret Harrigan. "The Comparability of Treatment and Control Groups at Randomization." Channeling Evaluation Technical Report TR-84B-04, Princeton, NJ: Mathematica Policy Research, Inc., 1983.
- Brown, Randall S. and Peter A. Mossel. "Examination of the Equivalence of Treatment and Control Groups and the Comparability of Baseline Data." Princeton, NJ: Channeling Evaluation Technical Report TR-84B-05, Mathematica Policy Research, Inc., October 1984.
- Brown, Randall S. and Barbara Phillips. "The Effects of Case Management and Community Services on the Impaired Elderly." Channeling Evaluation Technical Report TR-86B-02, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Carcagno, George J., et al. "The Evaluation of the National Long Term Care Demonstration: The Planning and Operational Experience of the Channeling Projects." Channeling Evaluation Technical Report TR-86B-05. Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Christianson, Jon B. "Channeling Effects on Informal Care." Channeling Evaluation Technical Report TR-86B-07, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Corson, Walter, et al. "Channeling Effects on Formal Community Based Services and Housing." Channeling Evaluation Technical Report TR-86B-10, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Corson, Walter, Sharon Long, and Rebecca Maynard. "An Impact Evaluation of the Buffalo Dislocated Worker Demonstration Program." Princeton, NJ: Mathematica Policy Research, Inc., 1985.
- Duan, Naihou, et al. "A Comparison of Alternative Models for the Demand for Medical Care." Journal of Business and Economic Statistics, vol. 1, no.2, April 1983.

- Grannemann, Thomas, et al. "Differential Impacts Among Subgroups of Channeling Enrollees." Channeling Evaluation Supplementary Report 86-04, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Greene, W.H. "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model." Econometrics, vol. 49, no. 2, March 1981.
- _____. "Estimation of Limited Dependent Variable Models by Ordinary Least Squares and the Method of Moments." Journal of Econometrics, vol. 21, no. 2, February 1983.
- Grossman, Jean, Rebecca Maynard and Judy Roberts. "Reanalysis of the Effects of Selected Employment and Training Programs for Welfare Recipients." Princeton, NJ: Mathematica Policy Research, Inc., 1985.
- Hackman, James. "Sample Selection Bias as a Specification Error." Econometrics, vol. 47, no. 1, January 1979.
- Hollister, Robert, Peter Kemper, and Rebecca Maynard. The National Supported Work Demonstration. Madison, WI: The University of Wisconsin Press, 1984.
- Kemper, Peter, et al. "Initial Research Design of the National Long Term Care Demonstration." Channeling Evaluation Technical Report TR-84B-02, Princeton, NJ: Mathematica Policy Research, Inc., 1982.
- Kemper, Peter et al. "Channeling Effects for an Early Sample at 6-Month Followup." Channeling Evaluation Supplementary Report 85-01, Princeton, NJ: Mathematica Policy Research, Inc., 1985.
- Kemper, Peter, et al. "The Evaluation of the National Long Term Care Demonstration: Final Report." Channeling Evaluation Final Report 86-01, Princeton, NJ: Mathematica Policy Research, Inc., 1986.
- Kmenta, Jan. Elements of Econometrics. New York, NY: The Macmillan Company, 1971.
- Maddala, G.S. Limited-dependent and Qualitative Variables in Econometrics. New York, NY: Cambridge University Press, 1983.
- Moffitt, Robert A. and J. McDonald. "The Uses of Tobit Analysis." Review of Economics and Statistics, May 1980.
- Phillips, Barbara et al. "The Evaluation of the National Long Term Care Demonstration: Survey Data Collection Design and Procedures." Channeling Evaluation Technical Report TR-86B-03, Princeton, NJ: Mathematica Policy Research, Inc., 1986.

Wooldridge, Judith, and Jennifer Schore. "Channeling Effects on Hospital, Nursing Home, and Other Medical Services." Channeling Evaluation Technical Report TR-86B-09, Princeton, NJ: Mathematica Policy Research, Inc., 1986.

CHANNELING EVALUATION REPORTS AND DATA COLLECTION INSTRUMENTS

FINAL REPORT

- 86-01 The Evaluation of the National Long Term Care Demonstration: Final Report. Peter Kemper et al. 1986.

TECHNICAL REPORTS

- TR-84B-01 Issues in Developing the Client Assessment Instrument for the National Long Term Care Demonstration. Barbara Phillips, Raymond J. Baxter, at Susan A. Stephens. January 1981, 167 pages, \$11.70.
- TR-84B-02 Initial Research Design of the National Long Term Care Demonstration. Peter Kemper et al. November 15, 1982, 230 pages, \$16.10.
- TR-84B-03 Informal Care to the Impaired Elderly: Report of the National Long Term Care Demonstration Survey of Informal Caregivers. Jon B. Christianson and Susan A. Stephens. June 6, 1984, 187 pages, \$13.10.
- TR-84B-04 The Comparability of Treatment and Control Groups at Randomization. Randall S. Brown and Margaret Harrigan. October 27, 1983, 38 pages, \$2.65.
- TR-84B-05 Examination of the Equivalence of Treatment and Control Groups and the Comparability of Baseline Data. Randall S. Brown and Peter A. Mossel. October 1984, 151 pages, \$11.30.
- TR-86B-01 Methodological Issues in the Evaluation of the National Long Term Care Demonstration. Randall S. Brown. 1986.
- TR-86B-02 The Effects of Case Management and Community Services on the Impaired Elderly. Randall S. Brown and Barbara Phillips. 1986.
- TR-86B-03 The Evaluation of the National Long Term Care Demonstration: Survey Data Collection Design and Procedures. Barbara Phillips, Susan Stephens, and Joanna Cerf (with others). March 1986, 292 pages, \$20.45.

- TR-86B-04 The Evaluation of the National Long Term Care Demonstration: Analysis of Channeling Project Costs. Craig Thornton, Joanna Will and Mark Davies. 1986.
- TR-86B-05 The Evaluation of the National Long Term Care Demonstration: The Planning and Operational Experience of the Channeling Projects, Volume 1 and 2. George Carcagno, Robert Applebaum, Jon Christianson, Barbara Phillips, Craig Thornton and Joanna Will. 1986.
- TR-86B-06 The Evaluation of the National Long Term Care Demonstration: The Effects of Sample Attrition on Estimates of Channeling's Impacts. Peter A. Mossel and Randall S. Brown. 1986.
- TR-86B-07 Channeling Effects on Informal Care. Jon B. Christianson. May 1986, 310 pages, \$21.70.
- TR-86B-08 Channeling Effects on Informal Care: Appendixes. Jon B. Christianson. May 1986, 230 pages, \$16.10.
- TR-86B-09 Channeling Effects on Hospital, Nursing Home, and Other Medical Services. Judith Wooldridge and Jennifer Schore. 1986.
- TR-86B-10 Channeling Effects on Formal Community Based Services and Housing. Walter Corson, Thomas Grannemann, Nancy Holden and Craig Thornton. 1986.
- TR-86B-11 Channeling Effects on the Quality of Clients' Lives. Robert A. Applebaum and Margaret Harrigan. April 1986, 137 pages, \$9.60.
- TR-86B-12 The Evaluation of the National Long Term Care Demonstration: Analysis of the Benefits and Costs of Channeling. Craig Thornton and Shari Miller Dunstan. 1986.
- TR-86B-13 Final Report on the Effects of Sample Attrition on Estimates of Channeling's Impact. Randall S. Brown, et al. 1986.
- TR-86B-14 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part I: Background." Judith Wooldridge, Shari Miller Dunstan, and Nancy Holden. May 1986.
- TR-86B-15 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #1: Screen File." Judith Wooldridge and Daniel J. Buckley. May 1986.

- TR-86B-16 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #2: Baseline File." Judith Wooldridge and Daniel J. Buckley. May 1986.
- TR-86B-17 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #3: Client Tracking/Status Change File." Judith Wooldridge, Nancy Holden and Margaret Harrigan. May 1986.
- TR-86B-18 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #4: Sample Member Followup Files." Judith Wooldridge and Daniel J. Buckley. May 1986.
- TR-86B-19 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #5: Status File." Judith Wooldridge, Nancy Holden and Margaret Harrigan. May 1986.
- TR-86B-20 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #6: Caregiver Baseline." Judith Wooldridge and Daniel J. Buckley. May 1986.
- TR-86B-21 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #7: Caregiver Followup Files." Judith Wooldridge and Richard Ross. May 1986.
- TR-86B-22 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #8: Formal Community Services, Housing and Transfers, and Case Management File." Judith Wooldridge, Shari Miller Dunstan and Daniel J. Buckley. May 1986.
- TR-86B-23 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #9: Informal Care File." Judith Wooldridge, Shari Miller Dunstan, and Richard Ross. May 1986.
- TR-86B-24 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #10: Hospital, Nursing Home, and Other Medical Services File." Judith Wooldridge and Daniel J. Buckley. May 1986.

TR-86B-25 "Data Base Documentation, National Long Term Care Channeling Evaluation, Part II: Individual Public Use File Documentation, Report #11: Well Being File" Judith Wooldridge, Shari Miller Dunstan and Richard Ross. May 1986.

SUPPLEMENTARY REPORTS

- 86-02 The Evaluation of the National Long Term Care Demonstration: Tables Comparing Channeling to Other Community Care Demonstrations. Robert A. Applebaum, Margaret Harrigan, and Peter Kemper. 1986.
- 86-04 Differential Impacts Among Subgroups of Channeling Enrollees. Thomas W. Grannemann and Jean Baldwin Grossman. 1986.
- 86-05 The Evaluation of the National Long Term Care Demonstration: An Analysis of Site-Specific Results. Robert A. Applebaum, Randall S. Brown, and Peter Kemper. May 1986, 66 pages, \$4.65.

PRELIMINARY REPORTS

- 83-15 The Planning and Implementation of Channeling: Early Experiences of the National Long Term Care Demonstration. Raymond J. Baxter et al. April 15, 1983, 259 pages.³⁵
- 83-21 Implementation of Early Operation of the National Long Term Care Demonstration: Overview. Raymond J. Baxter et al. December 1983, 36 pages, \$2.55.
- 85-01 Channeling Effects for an Early Sample at 6-Month Followup. Peter Kemper et al. May 1985, 1984 pages, \$13.50.
- 85-05 Channeling Effects for an Early Sample at 6-Month Followup: Executive Summary. Peter Kemper et al. May 1985, 12 pages, no charge.

DATA COLLECTION INSTRUMENTS

- 82-08 National Long Term Care Demonstration Applicant Screen. March 1982, 20 pages, \$2.00.

³⁵ This report is available free of charge from from the Office of the Assistant Secretary for Planning and Evaluation, Department of Disability Aging and Long Term Care Policy, Division of Health and Human Services, HHH Building, 200 Independence Avenue, S.W., Washington, D.C. 20201.

- 82-09 National Long Term Care Demonstration Clinical Assessment and Research Baseline Instrument: Community Version.³⁶ March 1982, 108 pages, \$6.30.
- 82-10 National Long Term Care Demonstration Clinical Assessment and Research Baseline Instrument: Institutional Version.³⁶ March 1982, 112 pages, \$7.85.
- 82-11 National Long Term Care Demonstration Followup Instrument. November 1982, 78 pages, \$5.45.
- 82-12 National Long Term Care Demonstration Client Tracking/Status Change Form. March 1982, 3 pages, no charge.
- 82-13 National Long Term Care Demonstration Evaluation: Channeling Demonstration Project Instructions Manual for Reporting Financial Status, September 1982, 79 pages, \$5.55.
- 83-22 National Long Term Care Demonstration Informal Caregiver Survey Baseline. January 1983, 44 pages, \$3.10.
- 83-23 National Long Term Care Demonstration Informal Caregiver Followup Instrument. August 1983, 102 pages, \$7.15.
- 84-13 National Long Term Care Demonstration Institutional Provider Discussion Guide. January 1984, 22 pages, \$2.00.
- 84-14 National Long Term Care Demonstration Community-Based Provider Discussion Guide. February 1984, 68 pages, \$4.75.
- 84-15 National Long Term Care Demonstration Provider Characteristics Instrument. January 1984, 26 pages, \$2.00.

All papers are reproduced in an 8-1/2 x 11 format. Each publication is priced separately, based on the number of pages it contains. In addition a charge will be made for postage and handling costs. Please make all checks payable to Mathematica Policy Research, Inc. Requests for publications should be sent to:

Mathematica Policy Research, Inc.
Office of Publications
Room 247
P.O. Box 2393
Princeton, New Jersey 08543-2393
(609) 799-3535

³⁶ These instruments were used for both research and clinical purposes. After the research sample intake was completed, a clinical version of the instrument was subsequently developed by Temple University's Institute on Aging and is available from it.

CHANNELING TECHNICAL ASSISTANCE REPORTS

A Guide to Memorandum of Understanding Negotiation and Development -- M. Johnson and L. Sterthous, October 1981

Liability: Issues of Negligence and Liability, prepared for the National Channeling Demonstration Projects -- E. Cohen and L. Staroscik, October 1982

Channeling Case Management Manual -- B. Schneider, L. Gottesman, P. Kohn, B. Morrell, L. Staroscik, L. Sterthous, January 1983

Care Plan Study Final Report (1986)

Trainer's Manual for Client Assessment Training (1986)

Trainer's Manual for Screening Training (1986)

Trainer's Manual for Case Management Training (1986)

Program Product: Clinical Baseline Assessment Instrument and Instruction Manual (1986)

Program Product: Care Plan Forms and Instructions (1986)

TECHNICAL ASSISTANCE REPORTS ARE AVAILABLE FROM:

Temple University
Institute on Aging 083-52
1601 North Broad Street
Philadelphia, Pennsylvania 19122

NATIONAL LONG-TERM CARE CHANNELING DEMONSTRATION

REPORTS AVAILABLE

A Guide to Memorandum of Understanding Negotiation and Development

HTML: <http://aspe.hhs.gov/daltcp/reports/mouguide.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/mouguide.pdf>

An Analysis of Site-Specific Results

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/sitees.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/sitees.pdf>

Analysis of Channeling Project Costs

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/projctes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/projctes.pdf>

Analysis of the Benefits and Costs of Channeling

Executive Summary: <http://aspe.hhs.gov/daltcp/reports/1986/costes.htm>

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/cost.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/cost.pdf>

Applicant Screen Set

HTML: <http://aspe.hhs.gov/daltcp/reports/1982/appscset.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1982/appscset.pdf>

Assessment and Care Planning for the Frail Elderly: A Problem Specific Approach

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/asmtcare.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/asmtcare.pdf>

Assessment Training for Case Managers: A Trainer's Guide

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/asmttran.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1985/asmttran.pdf>

Case Management Forms Set

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/cmforms.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1985/cmforms.pdf>

Case Management Training for Case Managers: A Trainer's Guide

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/cmtrain.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1985/cmtrain.pdf>

Channeling Effects for an Early Sample at 6-Month Follow-up

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/6monthes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1985/6monthes.pdf>

Channeling Effects on Formal Community-Based Services and Housing

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/commtyes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/commtyes.pdf>

Channeling Effects on Hospital, Nursing Home and Other Medical Services

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/hospites.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/hospites.pdf>

Channeling Effects on Informal Care

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/informes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/informes.pdf>

Channeling Effects on the Quality of Clients' Lives

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/qualtyes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/qualtyes.pdf>

Clinical Baseline Assessment Instrument Set

HTML: <http://aspe.hhs.gov/daltcp/reports/cbainstr.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/cbainstr.pdf>

Community Services and Long-Term Care: Issues of Negligence and Liability

HTML: <http://aspe.hhs.gov/daltcp/reports/negliab.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/negliab.pdf>

Differential Impacts Among Subgroups of Channeling Enrollees

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/enrolles.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/enrolles.pdf>

Differential Impacts Among Subgroups of Channeling Enrollees Six Months After Randomization

HTML: <http://aspe.hhs.gov/daltcp/reports/1984/difimpes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1984/difimpes.pdf>

Examination of the Equivalence of Treatment and Control Groups and the Comparability of Baseline Data

HTML: <http://aspe.hhs.gov/daltcp/reports/1984/baslines.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1984/baslines.pdf>

Final Report on the Effects of Sample Attrition on Estimates of Channeling's Impacts

Executive Summary: <http://aspe.hhs.gov/daltcp/reports/1986/atrines.htm>

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/atrtn.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/atrtn.pdf>

Informal Care to the Impaired Elderly: Report of the National Long-Term Care Demonstration Survey of Informal Caregivers

HTML: <http://aspe.hhs.gov/daltcp/reports/1984/impaires.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1984/impaires.pdf>

Informal Services and Supports

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/infserv.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1985/infserv.pdf>

Initial Research Design of the National Long-Term Care Demonstration

HTML: <http://aspe.hhs.gov/daltcp/reports/designes.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/designes.pdf>

Issues in Developing the Client Assessment Instrument for the National Long-Term Care Channeling Demonstration

HTML: <http://aspe.hhs.gov/daltcp/reports/1981/instrues.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1981/instrues.pdf>

Methodological Issues in the Evaluation of the National Long-Term Care Demonstration

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/methodes.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1986/methodes.pdf>

National Long-Term Care Channeling Demonstration: Summary of Demonstration and Reports

HTML: <http://aspe.hhs.gov/daltcp/reports/1991/chansum.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1991/chansum.pdf>

Screening Training for Screeners: A Trainer's Guide

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/scretrai.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1985/scretrai.pdf>

Survey Data Collection Design and Procedures

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/sydataes.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1986/sydataes.pdf>

Tables Comparing Channeling to Other Community Care Demonstrations

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/tablees.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1986/tablees.pdf>

The Channeling Case Management Manual

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/cmmanual.htm>
PDF: <http://aspe.hhs.gov/daltcp/reports/1986/cmmanual.pdf>

The Channeling Financial Control System

HTML: <http://aspe.hhs.gov/daltcp/reports/1985/chanfcs.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1985/chanfcs.pdf>

The Comparability of Treatment and Control Groups at Randomization

HTML: <http://aspe.hhs.gov/daltcp/reports/compares.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/compares.pdf>

The Effects of Case Management and Community Services on the Impaired Elderly

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/casmanes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/casmanes.pdf>

The Effects of Sample Attrition on Estimates of Channeling's Impacts for an Early Sample

HTML: <http://aspe.hhs.gov/daltcp/reports/1984/earlyes.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1984/earlyes.pdf>

The Evaluation of the National Long-Term Care Demonstration: Final Report

Executive Summary: <http://aspe.hhs.gov/daltcp/reports/1986/chanes.htm>

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/chan.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/chan.pdf>

The Evaluation of the National Long-Term Care Demonstration

Executive Summary: <http://aspe.hhs.gov/daltcp/reports/1988/hsres.htm>

HTML: <http://aspe.hhs.gov/daltcp/reports/1988/hsre.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1988/hsre.pdf>

The Planning and Implementation of Channeling: Early Experiences of the National Long-Term Care Demonstration

Executive Summary: <http://aspe.hhs.gov/daltcp/reports/1983/implees.htm>

HTML: <http://aspe.hhs.gov/daltcp/reports/1983/imple.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1983/imple.pdf>

The Planning and Operational Experience of the Channeling Projects (2 volumes)

HTML: <http://aspe.hhs.gov/daltcp/reports/1986/proceses.htm>

PDF: <http://aspe.hhs.gov/daltcp/reports/1986/proceses.pdf>

DATA COLLECTION INSTRUMENTS

Applicant Screen

HTML: <http://aspe.hhs.gov/daltcp/instruments/1981/AppSc.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/1981/AppSc.pdf>

Client Contact Log

HTML: <http://aspe.hhs.gov/daltcp/instruments/CIConLog.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/CIConLog.pdf>

Client Tracking Form

HTML: <http://aspe.hhs.gov/daltcp/instruments/1982/CITracFm.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/1982/CITracFm.pdf>

Clinical Assessment and Research Baseline Instrument: Community Version

HTML: <http://aspe.hhs.gov/daltcp/instruments/carbicv.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/carbicv.pdf>

Clinical Baseline Assessment Instrument: Community Version

HTML: <http://aspe.hhs.gov/daltcp/instruments/1983/cbaicv.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/1983/cbaicv.pdf>

Clinical Baseline Assessment Instrument: Institutional Version

HTML: <http://aspe.hhs.gov/daltcp/instruments/1983/cbaiiv.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/1983/cbaiiv.pdf>

Eighteen Month Followup Instrument

HTML: <http://aspe.hhs.gov/daltcp/instruments/18mfi.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/18mfi.pdf>

Followup Instrument

HTML: <http://aspe.hhs.gov/daltcp/instruments/FollInst.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/FollInst.pdf>

Informal Caregiver Followup Instrument

HTML: <http://aspe.hhs.gov/daltcp/instruments/ICFollIns.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/ICFollIns.pdf>

Informal Caregiver Survey Baseline

HTML: <http://aspe.hhs.gov/daltcp/instruments/ICSurvey.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/ICSurvey.pdf>

Screening Identification Sheet

HTML: <http://aspe.hhs.gov/daltcp/instruments/1982/ScrIDSh.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/1982/ScrIDSh.pdf>

Time Sheet

HTML: <http://aspe.hhs.gov/daltcp/instruments/TimeSh.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/TimeSh.pdf>

Twelve Month Followup Instrument

HTML: <http://aspe.hhs.gov/daltcp/instruments/12mfi.htm>

PDF: <http://aspe.hhs.gov/daltcp/instruments/12mfi.pdf>

To obtain a printed copy of this report, send the full report title and your mailing information to:

U.S. Department of Health and Human Services
Office of Disability, Aging and Long-Term Care Policy
Room 424E, H.H. Humphrey Building
200 Independence Avenue, S.W.
Washington, D.C. 20201
FAX: 202-401-7733
Email: webmaster.DALTCP@hhs.gov

RETURN TO:

Office of Disability, Aging and Long-Term Care Policy (DALTCP) Home
[\[http://aspe.hhs.gov/office_specific/daltcp.cfm\]](http://aspe.hhs.gov/office_specific/daltcp.cfm)

Assistant Secretary for Planning and Evaluation (ASPE) Home
[\[http://aspe.hhs.gov\]](http://aspe.hhs.gov)

U.S. Department of Health and Human Services Home
[\[http://www.hhs.gov\]](http://www.hhs.gov)