

# The Research Design and Methodological Issues for the Cash and Counseling Evaluation

*Randall S. Brown and Stacy B. Dale*

---

**Objective.** To provide an overview of the design, research questions, data sources, and methods used to evaluate the Cash and Counseling Demonstration and resolution of analytic concerns that arose. The methodology was designed to provide statistically rigorous estimates while presenting the findings in a manner easily accessible to a broad, nontechnical audience.

**Study Setting.** Eligible Medicaid beneficiaries in Arkansas, Florida, and New Jersey who volunteered to participate in the demonstration were randomly assigned to receive an allowance and direct their own Medicaid supportive services as Cash and Counseling consumers (the treatment group) or to rely on Medicaid services as usual (the control group). The demonstration included elderly and nonelderly adults in all three states and children in Florida. Both age groups in Arkansas and New Jersey and the elderly adults group in Florida primarily included individuals with physical disabilities. In Florida, the children and nonelderly adults primarily included individuals with developmental disabilities. The intervention was conducted from 1999 through 2003.

**Data Sources.** Data included baseline and 9-month follow-up surveys of consumers, surveys of the primary informal caregiver and the primary paid worker for sample members, program data, interviews with program staff, and Medicaid and Medicare claims data.

**Methods.** Descriptive data analyses were conducted on program participation, program implementation, and the experiences of hired workers. Program impacts on consumers, caregivers, and costs were estimated using an intent-to-treat approach, comparing the regression-adjusted means of outcomes for the full treatment and control groups. A broad set of control variables from the baseline interview and prior Medicaid claims data controlled for possible preexisting differences. Ordinal scale responses were converted to binary outcome indicators for high and for low values for ease of presentation and interpretation of effects. Two-tailed statistical tests of the estimated effects were conducted at the .05 level. Separate estimates were provided for each state and for each age group. Sensitivity tests were conducted of the robustness of estimates to outliers (for continuous outcome measures) and to proxy use.

**Principal Findings/Conclusion.** The experimental design, high survey response rates, and available sample sizes led to valid, unbiased estimates of program impacts, with adequate power to detect moderate-size impacts on most outcomes for the key age

subgroups examined. For certain survey-based outcome measures related to satisfaction with paid care, the sample had to be restricted to those who received care and those without proxy respondents who were also hired workers. Sensitivity tests suggest that these necessary restrictions were unlikely to have led to overstatement of favorable program effects on these outcome measures. The high proportions of sample members with proxy respondents reflect the frailty of the sample members. Similar rates for treatment and control groups cases with proxy respondents suggest the high use of proxy respondents has not biased the estimated program effects on survey measures.

**Key Words.** Consumer-directed care, random assignment, evaluation design

---

The Cash and Counseling evaluation described why individuals chose to participate in the demonstration, how the three demonstration states (Arkansas, Florida, and New Jersey) implemented their programs, and how each state's program affected consumers, caregivers, hired workers, and Medicaid costs. To estimate the program's impacts, eligible beneficiaries who applied to participate were randomly assigned to either have the option to participate in Cash and Counseling in lieu of receiving traditional Medicaid services (the treatment group) or to receive supportive services as usual from Medicaid-certified providers (the control group). Within each state, differences in outcomes between the treatment and control groups provide estimates of the program's impacts. This article describes demonstration enrollment and random assignment, the analyses that were conducted, the survey instruments and other data sources, methods of estimating program effects, and possible limitations and methodological concerns. More detailed outcome-specific methodological issues are presented in other papers in this volume that address those outcomes.

## DEMONSTRATION ENROLLMENT

Beneficiaries who enrolled in the demonstration completed a baseline telephone interview and were then randomly assigned to the treatment or control group, in a one to one ratio. Random assignment was stratified by state, age group (elderly adults, nonelderly adults, and children), and whether the

---

Address correspondence to Randall Brown, Ph.D., Vice President, Director of Health Research, Mathematica Policy Research Inc., PO Box 2393, Princeton, NJ 08543. Stacy B. Dale, M.P.A., is a researcher at Mathematica Policy Research Inc., Princeton, NJ.

consumer was a new applicant for Medicaid personal call services (PCS) or home- and community-based services (HCBS) or was already receiving PCS/HCBS at the time they enrolled in the study. Random assignments were reported back to each state on a daily basis, and sample members were notified whether they were assigned to the treatment group or the control group. Those assigned to the treatment group were contacted by a counselor (or “consultant” in New Jersey and Florida) who interacted with them to: (1) develop and revise an allowance spending plan, (2) offer advice about hiring workers, and (3) monitor allowance use and well-being.

Initially, enrollment targets were set to be 3,100 adults in Arkansas and New Jersey and 4,650 adults and children in Florida for a 12-month intake period. After discovering it was more difficult than anticipated to recruit enrollees, the program extended the intake period in each state, and reduced the target sample sizes to approximately 2,000 adults in each state and 1,000 children (in Florida). Both Florida and New Jersey sought to have their adult samples comprised equally of elderly and nonelderly beneficiaries. Programs stopped enrolling when they reached their targets, or in July 2002, whichever came first. Arkansas began enrollment in December 1998 and enrolled 2,008 adults; New Jersey began intake in November 1999 and enrolled 1,755 adults; Florida enrolled 1,818 adults and 1,002 children, beginning in June 2000. (See Foster et al. [2007], for further discussion of enrollment patterns.)

## RESEARCH QUESTIONS AND ANALYSES

The research questions examined fell into four broad categories: (1) what types of consumers participated in the Cash and Counseling program; (2) how was the program implemented; (3) what were the program’s effects on consumers, caregivers, and costs; and (4) how did hired workers fare. Table 1 provides a summary of the research questions, measures, data sources, and methods and directs the reader to the paper in this issue that addresses each question in greater detail.

For each state, separate estimates were calculated for beneficiary subgroups defined by age (18–64, or 65 and older [Arkansas and New Jersey]; and under 18, 18–59, or 60 and older [Florida]).<sup>1</sup> The elderly and nonelderly age groups in Arkansas and New Jersey included adults with physical disabilities. In Florida, the elderly primarily included beneficiaries with physical disabilities, and the nonelderly adults and children included primarily those with developmental disabilities.<sup>2</sup> Separate estimates for elderly adults enabled tests

Table 1: Summary of Research Questions, Measures, Data Sources, and Methods

<i>Research Questions</i>	<i>Key Measures</i>	<i>Primary Data Sources</i>	<i>Methodology</i>	<i>This Issue:</i>
Beneficiary participation What proportion of eligible beneficiaries chose to enroll, and how do the characteristics of participants and nonparticipants differ?	Consumer's demographic characteristics; Medicaid costs for PCS or HCBS before the demonstration	Medicaid claims data and program enrollment data	Comparison of characteristics of participants to eligible nonparticipants	Schore et al.
Why did individuals agree (or decline) to participate?	Reasons for agreeing (or declining) to participate	Participation questionnaire	Descriptive tabulation of reasons for participation or nonparticipation	Schore et al.
Did the demonstration lead some eligible beneficiaries to use PCS or HCBS who otherwise would not have?	Ratio of new PCS (or HCBS) enrollees to all PCS (or HCBS) users	Medicaid claims and enrollment data	Examination of trends in ratio of new PCS (or HCBS) users to all PCS (or HCBS users) before and after demonstration intake	Schore et al.
Program implementation What were the key program features? What lessons were learned about how to best structure and operate a consumer-directed program?	Responses of program staff about the implementation of program elements (consumer eligibility criteria, outreach and enrollment, benefit and rate determination, content of counseling, uses of cash permitted, fiscal intermediary function, consumer	Interviews with key state officials, program staff members, personal care industry officials, and organizations providing outreach, enrollment, consulting and fiscal services	Qualitative analysis	Phillips and Schneider

*continued*

Table 1. *Continued*

<i>Research Questions</i>	<i>Key Measures</i>	<i>Primary Data Sources</i>	<i>Methodology</i>	<i>This Issue:</i>
How was the program implemented, according to consumers and the consultants who work with them?	monitoring); agency concerns; lessons Time until consumer received allowance; types of goods and services consumer used allowance for; consumer's use of and satisfaction with program services; methods consumer used to recruit workers; reasons for disenrollment; activities performed by consultants	Early (4-6 month) survey of treatment group; consultant survey; program data on consumers' uses of cash allowance in month 8 after enrollment; claims data on allowance payments	Descriptive tabulation of consumers' and consultants' experiences	Schore et al.
Program impacts How did the program affect consumers' use of personal assistance services?	Whether received paid care, hours of paid and unpaid care received, type and timing of care received, equipment and modifications made to meet personal assistance needs Consumer's satisfaction with paid caregivers, satisfaction with overall care arrangements, unmet needs for personal assistance, functioning measures, adverse events, health status and problems, and quality of life Caregiver's emotional, physical, and financial strain; amount, type, and timing of assistance	9-month consumer follow-up survey	Regression adjusted comparison of treatment group and control group means	Carlson et al.
How did the program affect consumers' well-being?	Consumer's satisfaction with paid caregivers, satisfaction with overall care arrangements, unmet needs for personal assistance, functioning measures, adverse events, health status and problems, and quality of life Caregiver's emotional, physical, and financial strain; amount, type, and timing of assistance	9-month consumer follow-up survey	Regression adjusted comparison of treatment group and control group means	Carlson et al.
How did the program affect consumers'	Caregiver's emotional, physical, and financial strain; amount, type, and timing of assistance	Caregiver survey	Regression adjusted comparison of treatment group	Foster et al.

primary informal caregivers?	caregiver provided; caregiver's satisfaction with care recipient's supportive services; caregiver's own health status, and satisfaction with life	and control group means	Dale and Brown
How did the program affect Medicaid and Medicare costs and service utilization?	Costs of PCS or HCBS, costs of other Medicaid services, costs of Medicare services, uses of specific services (hospital, nursing home, etc.)	Regression adjusted comparison of treatment group and control group means	Dale and Brown
Outcomes for workers	Amount, type and timing of care	Descriptive comparison	Foster et al.
How did hired workers fare?	worker provided; training worker received; worker's compensation; satisfaction with benefits and working conditions; physical, emotional, and financial strain	of hired workers for the treatment group and agency workers for the control group	

HCBS, home and community based services; PCS, personal care services.

of the hypothesis that consumer direction would not work for elderly consumers, and captured any differences in impacts across age groups that could have arisen due to the sizable differences between these populations in their needs, in the control group's likelihood of receiving paid care, and in the treatment group's participation in the program. A number of findings did differ in meaningful ways across these subgroups within states, so the distinction was meaningful. (While the sample was originally stratified by whether the consumer was a new applicant for or continuing user of PCS/HCBS, this distinction turned out to only be relevant in Arkansas, as it was the only state that allowed those who were not already enrolled in one of the feeder programs to enroll in Cash and Counseling. Therefore, we only reported results for this subgroup in our early reports on Arkansas. See Dale, Brown, and Phillips 2004; Dale et al. 2004; Foster et al. 2003.)

## DATA SOURCES

The implementation analyses drew from site visits, program data, and surveys of study participants, nonparticipants, and consultants. The major sources of evaluation data for the participation and impact analyses were (1) telephone surveys with demonstration participants and their caregivers, and (2) Medicare and Medicaid eligibility and claims data.

### *Participation Questionnaire*

The purpose of the participation questionnaire was to examine why some eligible beneficiaries chose to participate in Cash and Counseling, and others did not. Throughout the evaluation intake period, the outreach and enrollment workers in each state administered anonymous hard-copy participation questionnaires about eligible beneficiaries' reasons for agreeing or declining to participate in Cash and Counseling. Workers administered the questionnaire to beneficiaries who requested informational telephone calls or home visits, after the beneficiary decided whether or not to participate, yielding the following sample sizes of completed interviews for participants and decliners. The percentage of demonstration participants that responded was 47 percent in Arkansas, 67 percent in Florida, and 54 percent in New Jersey. It was not possible to calculate a questionnaire completion rate for nonparticipants, because the number of nonparticipants asked to complete the survey is not known.

Table 2: Questionnaire Respondents, by Participation Decision and State

	<i>Arkansa's</i>		<i>Florida</i>		<i>New Jersey</i>	
	<i>Partici- pants</i>	<i>Non- participants</i>	<i>Partici- pants</i>	<i>Non- participants</i>	<i>Partici- pants</i>	<i>Non- participant</i>
Number of respondents	953	585	1,877	2,792	950	1,735

In addition to collecting data on the participation decision, the questionnaire included questions about the beneficiary’s age, sex, race, and county of residence, how the demonstration was explained (in person or by telephone), who made the participation decision (the beneficiary alone or with others), whether the decision maker had ever supervised anyone, and how long the beneficiary had been receiving PCS or HCBS.

*Baseline Survey*

The baseline survey of demonstration participants was conducted by Mathematica at the time of consumers’ enrollment. It collected information on the consumer’s demographic characteristics, health and functioning, use of paid and unpaid supportive services, hiring and supervisory experience, satisfaction with overall care arrangements, perception of unmet needs, and attitudes about consumer direction. It also collected information on the familial relationship between the consumer and the primary informal caregiver, whether caregivers were employed, and whether caregivers were interested in being paid for caregiving, as reported by the consumer (or proxy). By design, the response rate was 100 percent, because individuals could only participate in the demonstration if they completed the baseline survey. (Individuals were assigned to the treatment or control group after responding to the baseline survey.) The baseline interview was conducted with adult participants and parents of child participants whenever possible, although proxies were often used due to the high proportion of sample members who had difficulty speaking, hearing, or understanding. Proxy rates for the baseline survey were 25 and 30 percent for the nonelderly in Arkansas and New Jersey, respectively, but were much higher (78 percent) in Florida, because 89 percent of the nonelderly sample members there were beneficiaries with developmental

disabilities. For the elderly, proxy rates ranged from 52 percent (in New Jersey) to 61 percent (in Florida).

#### *Early (4- or 6-Month) Survey of Treatment Group Members*

We conducted a telephone interview with treatment group members about their early experiences in the Cash and Counseling intervention. This interview was conducted 4 months after random assignment in Arkansas, between April 1999 and September 2001. Because Florida and New Jersey treatment group members started receiving allowances later than their Arkansas counterparts, this survey was conducted 6 months after random assignment in Florida (between January 2001 and February 2003) and in New Jersey (between June 2000 and February 2003).<sup>3</sup> It collected data on how consumers planned to use their allowance, difficulties they may have had with their employer responsibilities, their participation in Food Stamp and SSI programs, and reasons for dropping out (for disenrollees). Over 90 percent of treatment group members in each state responded to this survey. Control group members were sent a letter 4 (or 6) months after enrollment requesting any changes in contact information, to minimize treatment–control differences in response rates to the 9-month survey.

#### *Consumer 9-Month Survey*

We attempted to contact each treatment and control group member for a 30-minute telephone survey 9 months after enrollment (which fell between September 1999 and May 2003). This survey included questions on the consumer's use of personal care services and other supportive services, hours of paid and unpaid care received, unmet needs for care, satisfaction with care, health and functioning, and quality of life. The main reference period was defined as the most recent 2-week period that the sample member was living at home. (For those who were hospitalized or institutionalized, this would be the 2-week period before those events.) For questions about frequent occurrences (such as missing a dose of medicine), we asked about the previous week. For questions about less frequent occurrences (such as respiratory infections), we asked about the previous month. Response rates were approximately 85 percent in each state and were slightly higher for the treatment group than the control group, in each state (Table 3a).

In the spirit of consumer direction, we encouraged sample members to respond to our surveys themselves, if possible. However, even though individuals with mild to moderate cognitive impairments can state consistent

Table 3a: Sample Sizes and Response Rates for Cash and Counseling Surveys

	Arkansas		Florida		New Jersey	
	Treatment	Control	Treatment	Control	Treatment	Control
Overall sample size						
Nonelderly	279	277	456	458	404	413
Elderly	725	727	453	451	467	471
All adults	1,004	1,004	909	909	871	884
Children	—	—	501	501	—	—
9-month survey response						
Nonelderly	87.1	83.0	91.9	85.6	85.4	81.6
Elderly	88.6	85.8	82.3	80.5	86.1	80.9
Children (parents responded)	—	—	86.8	81.5	—	—
Percentage of 9-month surveys completed by proxies						
Nonelderly	30.0	27.4	82.6	82.4	35.7	38.0
Elderly	71.7	70.5	65.7	65.6	61.9	59.8
Percentage of 9-month surveys completed by proxies who were paid caregivers						
Nonelderly	9.7	2.2	12.5	1.1	8.7	1.2
Elderly	26.1	2.3	16.8	6.2	19.1	1.1
Children	—	—	13.9	0.0	—	—
Caregiver survey response rate						
All adults	83.6	81.7	86.7	79.3	79.2	70.0
Children	—	—	86.8	81.5	—	—
Percent of primary informal caregiver respondents also in primary paid worker sample						
Adults	24.1	—	17.0	—	32.2	—

*continued*

Table 3a. *Continued*

	Arkansas		Florida		New Jersey	
	Treatment	Control	Treatment	Control	Treatment	Control
Children	—	—	13.6	—	—	—
Paid worker survey						
Response rate among	95.5	77.9	94.1	78.1	96.3	79.7
workers for adults						
Response rate among	—	—	91.6	83.6	—	—
workers for children						
Percent of paid worker respondents						
also in primary informal caregiver						
sample						
Adult	44.5	NA	35.2	NA	46.1	NA
Children			25.8			

Informal caregivers who were also paid workers are included in both the Paid Worker Survey response rates and in the Caregiver Survey response rates.

preferences about their care (see McHorney 1996; Feinberg and Whitlach 2001), many consumers in our sample were too cognitively or physically impaired to respond to the detailed survey that we administered. Therefore, many consumers (29–83 percent) had proxies respond for them during this interview. Proxy respondents were generally the sample member's relative or informal caregiver. In some cases, the proxy respondent was also the respondent's paid caregiver (as many informal caregivers were hired by consumers). The use of proxy respondents enabled the most impaired consumers (who otherwise would have not responded to the survey at all) to be retained in our analyses.

If the proxy was also a paid caregiver, we omitted questions about consumers' unmet needs, satisfaction with personal care, and paid caregiver performance in order to avoid possibly self-serving responses. The percentage of the treatment group with a paid caregiver as the proxy respondent ranged from 9 percent (for the nonelderly in New Jersey) to 26 percent (for the elderly in Arkansas; Table 3a). Questions about satisfaction with paid care were (obviously) not posed to sample members who did not receive such care. Also, questions about adverse events, health problems, self-care, and quality of life were not posed to the proxies of sample members who died before the reference period in question. Table 3b shows how the sample sizes for each type of outcome are affected by each of these restrictions. We discuss sensitivity tests that we performed to assess the bias imposed by these restrictions in the final section of this paper.

### *Caregiver Survey*

The caregiver survey was conducted about 10 months after consumer's enrollment. The same basic survey instrument was administered to consumer's primary informal caregivers (the person providing the most unpaid care) and to their primary paid workers (the person providing the most paid care), but some questions were only asked of paid workers and others were only asked of unpaid workers. Those primary informal caregivers of treatment group members who were hired by the consumer and became primary paid workers were asked all of the questions. The percentage of the informal caregiver sample that was also included in the primary paid worker sample ranged from 14 percent (among children in Florida) to 32 percent (among New Jersey adults; see Table 3a).

The reference period for this survey was the most recent 2-week period that the consumer was at home (for caregivers) and for the most recent 2-week

Table 3b: Samples Sizes and Restrictions for Analyses Using Consumer 9-Month Follow-up Survey

	Type of Restriction	Arkansas		Florida		New Jersey	
		Treatment Group (N)	Control Group (N)	Treatment Group (N)	Control Group (N)	Treatment Group (N)	Control Group (N)
Service use	None	243	230	419	392	341	332
Nonelderly		642	624	372	362	402	381
Elderly				441	418		
Children							
Satisfaction with paid caregivers	Consumer had no paid care	186	136	204	200	255	219
Nonelderly	Proxy was worker	338	381	216	236	248	255
Elderly	Proxy could not answer			289	273		
Children							
Unmet needs	Proxy was worker	216	224	362	387	310	327
Nonelderly		453	607	296	330	308	373
Elderly				378	418		
Children							
Adverse events and health problems; quality of life	Deceased	234	226	418	391	339	329
Nonelderly		574	569	362	345	374	368
Elderly				439	418		
Children							

Overall satisfaction with care	Proxy was worker					
Nonelderly	Proxy could not answer	207	212	275	292	291
Elderly		418	501	245	278	291
Children				378	418	

Sample sizes represent maximum possible sample for outcome measure. Analyses for particular outcome measures could have had slightly smaller samples due to respondents failing to answer a particular question.

period for which the worker provided in-home care to the consumer (for paid workers). Proxy respondents were not allowed for caregivers.

*Primary Informal Caregivers.* The baseline survey asked consumers to identify and provide contact information for their primary informal caregiver, defined as the person who provided them with the most unpaid care during the week preceding the survey. Over 80 percent of consumers identified and provided contact information for a primary informal caregiver. These primary informal caregivers were surveyed about 10 months after the baseline interview about the types and amounts of care provided, the extent to which they worried about the beneficiaries' health and safety, and measures of the caregivers' physical, emotional, and financial well-being. The overall response rate was 84 percent for caregivers associated with treatment group members and 78 percent for those associated with control group members.

*Primary Paid Workers.* Primary paid workers were identified during the 9-month follow-up consumer survey and then interviewed. The consumer survey asked respondents receiving paid care at that time to identify and provide contact information for the person from whom they received the most paid hours for personal care, chores, activities, and routine health care during the week preceding the survey. We attempted to interview all these primary paid workers who had been directly hired by treatment group members and identified after August 2000, when this survey was begun.<sup>4</sup> Only a subset of the agency workers identified by control group members were surveyed—the target sample sizes were 300 completed interviews in Arkansas and New Jersey and 400 in Florida. We attempted to interview all agency workers identified after August 2000 until the targets were met.

The survey for primary paid workers was typically administered within 1 month after the consumer 9-month survey. It included questions about the type and timing of care the worker provided to the consumer in our sample, their compensation, satisfaction with working conditions, training and preparedness for work, and well-being. Response rates were similar in each state, averaging 79 percent for agency workers and 95 percent for directly hired workers.

### *Consultant Survey*

About 18 months after enrollment began MPR sent a 26-page mail survey, including many open-ended questions, to all currently active consultants (those providing counseling assistance to treatment group members). The

questionnaire contained sections on the consultant's background, caseload, consultant activities, perceived misuse of the allowance, awareness of whether any participants or their workers have been abused, recommended changes to consultant activities and the program itself, and the consultant's assessment of the program. In Arkansas, seven out of nine consultants responded; in New Jersey, 37 out of 50 (74 percent) responded, and in Florida, 180 out of 213 (85 percent) responded. Consultant's responses to this survey were used to inform the implementation analysis.

#### *Medicaid and Medicare Claims Data*

Medicaid and Medicare claims data were used to measure the costs and service utilization of all demonstration enrollees (thus there was no sample attrition for these analyses). Regression control variables for the cost analysis were constructed from sample members' Medicaid expenditures in the year before enrollment, preenrollment diagnoses (in Arkansas) and predicted expenditures based on their preenrollment diagnoses (in Florida and New Jersey). Outcome measures for the full sample were based on Medicaid and Medicare claims data for the first 12 months after enrollment. We followed the cohort of individuals who enrolled earliest in the demonstration for 2 post-enrollment years. This cohort includes 65 percent of the sample members in Arkansas, 82 percent in New Jersey, and 78 percent (of adults) in Florida.

## OUTCOME MEASURES

In several of our analyses, many of our outcome measures were derived from survey questions with four-point scales (e.g., degree of satisfaction). After first examining frequencies and determining that binary measures would not obscure important findings, we generally converted each four-point scale into two binary measures—one for the most favorable rating (very satisfied) and one for an unfavorable rating (somewhat or very dissatisfied). While impacts could be estimated with one multinomial logit model, such estimates would be imprecise because of the relatively large number of parameters estimated. Ordered logit models are designed for such ordinal outcome measures, but may mask important nonmonotonic impact patterns, such as the treatment group being more likely than the control group to be very satisfied and also more likely to be dissatisfied (and less likely to be between the two extremes). Thus, binary outcomes were defined for both ends of the spectrum on each variable.

Despite the fact that each analysis had a large number of outcome measures, we chose to estimate impacts on each of them rather than on indexes formed from multiple, related measures. We did this because the meaning of what is being measured is clearer when responses to actual survey questions are examined. For example, the reader can easily assess whether consumer direction increased the proportion of highly satisfied consumers, reduced the proportion of dissatisfied ones, or had both (or neither) effects. Also, the magnitude of the impacts is simple for readers to assess from binary measures based on actual survey questions. Finally, indexes typically assign arbitrary equal weights to the component measures, treat ordinal measures as if they were cardinal, and can mask important effects on component measures.

Outcome measures constructed from Medicaid and Medicare claims data also require explanation, given that many sample members were not at risk for such costs in every month. To avoid introducing selection bias, most of our analyses of these data were based on the Medicaid and Medicare expenditures of all treatment group and all control group members, including even those who had died or who were no longer enrolled in Medicaid or Medicare. Those who were never enrolled in Medicare (about 10 percent of the elderly and 60 percent of the nonelderly) in each of the three states had zero Medicare expenditures for the entire follow-up period.

Individuals enrolled in Medicaid and/or Medicare managed care programs were also retained in the analysis, even though there are no claims data for the services individuals receive in managed care programs that use capitated payments. None of the states used capitated managed care programs to cover Medicare long-term care services (i.e., nursing facility, home health, personal care, or long-term care waiver programs). However, at baseline, 15 percent of the sample in Florida and 11 percent of the sample in New Jersey was enrolled in Medicaid managed care programs that covered some Medicaid inpatient or outpatient services (but not long-term care), and 15 percent of the sample in Florida and 2 percent of the sample in New Jersey was enrolled in Medicare managed care. In our regression analyses, we did control for whether an individual was enrolled in managed care at baseline in case there were any chance differences in the managed care enrollment of treatment and control group members. The percentages of treatment group members and control group members enrolled in managed care in each state were similar at baseline (as would be expected under random assignment).

In summary, our cost estimates represent the program's average effects on total Medicaid and total Medicare costs per person over the year (or 2 years) after enrollment. However, Cash and Counseling would have little effect on the

Medicaid expenditures of those who are only enrolled in Medicaid or Medicare for a few months, and would not affect the public expenditures for services covered by capitated managed care programs; thus, our estimates should not be interpreted as expenditures per-Medicaid (or Medicare) enrollee per month at risk.

## METHODS

### *Methods for Impact Analyses*

Separate analyses were conducted for each state and for different age groups within each state. We estimated program effects separately for elderly and nonelderly adult consumers to evaluate the concern expressed by agencies and some policy makers that consumer direction is not appropriate for elderly people. (For the caregiver and worker analysis, however, we only report results for the full sample, as results did not vary by age group.) In Florida (the only program which included children), we estimated effects separately for children because children may have very different experiences under consumer direction than adults.

Impacts were evaluated by comparing outcomes for the treatment and control group, using regression analysis to increase precision and control for any observable characteristics on which the two groups differ by chance or due to differences in their survey response patterns. All of our models controlled for the consumer's preenrollment characteristics drawn from the baseline survey, including the consumer's demographic characteristics, health and functioning, and use of, satisfaction with, and unmet needs for personal care, and whether the consumer used a proxy at baseline (which was highly correlated with proxy use at follow-up, but not endogenous). The caregiver analysis also controlled for the caregiver's demographic characteristics that were drawn from the caregiver survey. Control variables for the cost analysis included the sample member's preenrollment costs and diagnoses (according to the Medicaid claims data), as well as a subset of the baseline survey variables. Each set of impact analyses included those control variables that might reasonably be expected to affect any of the outcome measures, as shown in Table 4. Such "reduced form" models are commonly used in evaluating impacts from randomized trials, given the complex causal relationships that are likely to exist among the numerous outcome measures.

As would be expected under random assignment, the baseline means for treatment group members were generally similar to those of control group members, though there were a few chance differences. We also examined the

Table 4: Regression Control Variables

	Impact Analysis		
	Consumers	Caregivers	Costs
<i>Baseline (Preenrollment) Variables</i>			
Consumer's demographic characteristics	X	X	X
Consumer's health and functioning	X	X	X
Types and amounts of personal assistance consumer received (2 weeks before enrollment)	X	X	X
Consumer's satisfaction with care and unmet needs	X	X	
Consumer's reasons for enrolling in the demonstration	X	X	
Consumer's work experience	X	X	
Consumer appointment of a proxy and/or a representative	X	X	
Primary informal caregiver's demographic characteristics*	X	X	
Primary informal caregiver's employment status and interest in becoming paid	X	X	
Consumer's baseline care plan value	X		X
Medicaid expenditures during preenrollment year <sup>†</sup>			X
Predicted cost of preenrollment diagnoses <sup>†</sup>			X
Consumer's enrollment in Medicaid or Medicare HMO <sup>†</sup>			X
Consumer's enrollment in Medicare, Medicaid, and other Medicaid waiver programs <sup>†</sup>			X

\*Drawn from the Caregiver Survey.

<sup>†</sup>Drawn from preenrollment Medicaid claims data and enrollment files.

<sup>‡</sup>Program intake data.

All variables drawn from the consumer baseline survey unless otherwise noted. HMO, health maintenance organization.

baseline means of 50 measures for treatment and control group members in the restricted samples that were used in our analyses of consumer well-being. In the most restricted sample (the sample used to analyze satisfaction with paid care in Arkansas) there were six significant differences (at the .10 level) in the nonelderly restricted sample (versus one for the full nonelderly sample) and 10 significant differences in the elderly sample (versus six for the full elderly sample) between the baseline means for the treatment and control groups. See Carlson et al. (2005); Dale and Brown (2005); Dale et al. (2005), and Foster et al. (2005) for complete sets of baseline means separately for the treatment and for the control group for the consumer, caregiver, cost, and paid worker samples. Means for selected baseline characteristics for the 9-month full sample are shown by state and age group in Table 5. Schore et al. (2007) describes how the characteristics of enrollees differ from those of eligible consumers who did not enroll.

For binary outcome measures (such as whether a respondent was very satisfied with care), treatment–control differences in outcomes were estimated with logistic regression models. We calculated the magnitude of the treatment–control difference by using the estimated model to compute the average predicted probability of the outcome occurring across all sample members under the assumption that each sample member was in the treatment group, and then repeating the calculation under the assumption that each member was in the control group. The difference between the two mean probabilities is the estimated mean impact on the probability of the outcome occurring. (We reported predicted probabilities rather than odds-ratios because predicted probabilities are easier to interpret for nontechnical readers.) For continuous outcome measures (such as hours of care provided or expenditures), ordinary least squares regression models were used to estimate impacts and to calculate predicted means for the treatment and control groups with all regressors set at their sample means.

Only outcomes for which the treatment–control difference was significantly different from zero at the .05 level, using two-tailed tests, are considered to have been affected by the Cash and Counseling program. For each type of model, we used the *p*-values of the estimated coefficients on the treatment status variable to assess the statistical significance of the impacts. This conservative approach may have resulted in failure to detect small program effects on some outcomes. However, by design, the sample sizes for the consumers are sufficiently large that we can be 80 percent certain of correctly concluding from our tests that the program had an impact if the true effect of the program is about 10 percentage points or greater for binary

Table 5: Selected Baseline Characteristics of Demonstration Enrollees (Percentage, Unless Noted)

	Arkansas		Florida			New Jersey	
	18-64	65+	3-17	18-59	60+	18-64	65+
Age (years)							
3-12	--	--	71.2	--	--	--	--
13-17	--	--	28.8	--	--	--	--
18-39	27.5	--	--	75.5	--	33.8	--
40-64 (40-59 in Florida)	72.5	--	--	24.5	--	66.2	--
65-79 (60-79 in Florida)	--	50.4	--	--	52.9	--	58.5
80+	--	49.6	--	--	47.1	--	41.5
Female	67.6	81.5	37.0	45.4	79.0	66.1	78.7
Self-identified as:							
White (versus black)	63.5	59.2	79.7	77.6	69.5	49.9	58.5
Hispanic (regardless of race)	0.9	1.3	18.6	21.2	34.6	29.8	40.4
Lived alone	38.8	29.6	0.0	9.1	28.4	34.3	36.5
Had at least one informal caregiver	89.6	91.5	99.7	94.1	84.9	84.6	84.2
Lived in area that was:							
Rural	36.5	39.7	18.2	14.6	10.7	9.5	10.6
Not independent in:*							
Transferring	62.1	67.4	61.0	51.8	65.2	66.3	66.5
Using toilet	59.2	68.0	84.2	63.6	66.8	69.4	65.7
Bathing	86.3	91.9	91.4	77.8	88.4	85.7	86.9
Needed more help with personal care <sup>†</sup>	69.3	61.9	66.0	53.7	64.2	74.1	73.6
Proxy respondent completed at least half of baseline interview	24.3	58.2	99.9	78.0	60.2	30.1	51.7
Had representative at baseline	26.6	48.3	100.0	84.9	70.0	NA	NA
Had cognitive impairment (inferred) <sup>‡</sup>	15.4	28.8	NA	57.9	45.7	NA	NA
Was receiving publicly funded home care <sup>§</sup>	60.6	79.1	59.1	65.2	70.2	43.7	46.4

Prospective monthly allowance amount (dollars)	358	298	1,165	1,610	817	1,087	1,072
Enrolled in Medicare	43.5	96.4	NA	41.5	93.3	42.4	90.5
Sample size	556	1,452	1,002	914	904	817	938

\*Received hands-on or stand-by help or did not perform activity at all.

†Personal care includes bathing, transferring, eating, and using the toilet.

‡We inferred the presence of a cognitive impairment if the sample member appointed a representative at enrollment and was physically or mentally unable to respond to the baseline survey.

§For Arkansas, the percentages reflect whether consumers were receiving publicly funded home care at baseline, regardless of how long they had been receiving it. For Florida and New Jersey, the percentages reflect whether consumers were receiving publicly funded home care for 6 months or longer at baseline. The measures differ because nearly all demonstration enrollees in Florida and New Jersey, by design, were already receiving traditional supportive services at the time they enrolled in the demonstration; in Arkansas, many were not receiving these services.

Source: Program records from each demonstration state and baseline telephone interviews by Mathematica Policy Research Inc.

outcomes with means of .4–.6, for the full sample of consumers or caregivers for each of the state-age groups (Table 6). This should be adequate precision for assessing binary outcomes; smaller favorable effects may not be that important to policy makers.<sup>5</sup> For the sample used to analyze satisfaction with paid care (the most restricted sample), minimum detectable effects ranged from 11 to 16 percentage points. We still observed many statistically significant impacts on satisfaction with paid care in spite of the lower power for these analyses.

The power of our statistical tests for impacts on Medicaid and Medicare costs is lower than for binary outcomes, because the variances of cost outcomes are larger. Because we did not want to overlook any unfavorable cost outcomes, treatment–control costs differences with *p*-values of .10 were considered statistically significant. Even so, minimum detectable impacts for cost outcomes were large, ranging from 7 to 25 percent of the control group mean, depending on the state and age group. This lower precision means that estimates must be interpreted cautiously. Failure to reject the hypothesis of equal costs for the treatment and control groups does not necessarily mean that costs were unaffected by the program; it only implies that the differences were not large enough for us to be confident that they are due to the program rather than to chance. In our cost article, we do discuss treatment–control differences (particularly adverse impacts) that are of a sizeable magnitude, even if they are not statistically significant.

For all impact analyses (consumer, caregiver, and costs), effects were estimated by comparing the subsequent outcomes for the full treatment and control groups (or the full set of survey respondents in these two groups),<sup>6</sup> regardless of whether the treatment group members actually received the monthly allowance.<sup>7</sup> The estimated treatment–control differences therefore reflect the effects on interested beneficiaries of being offered the *opportunity* to manage the allowance. Because some beneficiaries never received an allowance, this intent-to-treat approach understates the impacts of actual participation in the program.

Finally, we assessed the effects of outliers on our hours of care analyses and costs analyses. In the few cases where outliers appear to have substantially affected the treatment–control difference in means, we reported the distribution of the outcome measure (in addition to the mean value) for both the treatment and control groups.

#### *Methods for Primary Paid Worker Descriptive Analysis*

The analysis of primary paid workers is descriptive. While we do compare the experiences of directly hired workers with those of agency workers

Table 6: Minimum Detectable Impacts for Analysis Samples

	Arkansas			Florida			New Jersey		
	Sample Size	Detectable Difference*	Sample Size	Detectable Difference*	Sample Size	Detectable Difference*	Sample Size	Detectable Difference*	
Minimum detectable impacts on consumers service use									
Nonelderly	473	12.9	811	9.8	682	10.8			
Elderly	1,266	7.9	736	10.3	783	10.0			
Children	—	—	859	9.6	—	—			
Adverse events									
Nonelderly	460	13.1	809	9.8	668	10.8			
Elderly	1,143	8.3	707	10.5	742	10.3			
Children	—	—	857	9.6	—	—			
Satisfaction with paid care									
Nonelderly	322	15.8	404	13.9	474	12.9			
Elderly	719	10.5	452	13.2	503	12.5			
Children	—	—	562	11.8	—	—			
Minimum detectable impacts on caregivers:									
All adults	1,433	7.4	1,193	8.1	1,042	8.7			
Children	—	—	829	9.7	—	—			
Minimum detectable impacts on costs: <sup>+</sup>									
Year 1 results									
Nonelderly	556	16.4	913	11.9	817	20.2			
Elderly	1,452	6.9	904	12.0	938	12.3			
All adults	2,008	7.2	1,817	7.7	1,755	13.4			

*continued*

Table 6. Continued

	Arkansas		Florida		New Jersey	
	Sample Size	Detectable Difference*	Sample Size	Detectable Difference*	Sample Size	Detectable Difference*
Children	—	—	1,002	15.7	—	—
Year 2 results						
Nonelderly	383	19.1	833	14.1	671	23.1
Elderly	928	8.0	591	25.3	776	20.1
All adults	1,311	8.1	1,424	12.5	1,447	15.3
Children	—	—	1,002	15.7	—	—
Half width of 95% confidence intervals for paid workers						
Adults						
Hired by consumer	391	± 5.0	298	± 5.7	382	± 5.0
Agency worker	281	± 5.8	255	± 6.1	308	± 5.6
Children						
Hired by consumer	—	—	—	—	222	± 6.6
Agency worker	—	—	—	—	164	± 7.7

Calculations assume 80 percent power for two-tailed tests at the .05 significance level (for survey-based impacts) and the .10 level (for cost impacts). Outcome variables (other than costs) are assumed to be binary, and have a mean of .5. Minimum detectable effect calculations for costs are based on actual standard errors of the estimates of program impacts on total Medicaid costs for the year after enrollment.

\*Detectable differences are expressed as percentage points for impacts on binary outcomes for consumers and caregivers, and as a percent of the control group mean for impacts on costs. Confidence intervals are also in terms of percentage points for a binary variable with mean of .50.

†The anomalous result of larger detectable differences in costs when estimated on all adults than when estimated on just the elderly, despite the larger sample, is due to the use of actual standard errors from the regressions for each sample in these calculations. The large difference in means between the elderly and onelderly subgroups in both Arkansas and New Jersey leads to a much larger variance for the all adults group, offsetting the larger sample size.

serving control group members, the latter group does not reflect the counterfactual for the directly hired workers—i.e., it does not represent expected mean values of outcomes for the hired workers had the program not existed. However, the agency workers provide a useful benchmark for interpreting the estimates obtained for hired workers. We compared the means values and distributions for directly hired workers and agency workers using  $t$ -tests and  $\chi^2$  tests to identify differences greater than might be expected to occur by chance in samples of this size if the population means were identical.

## DRAWING INFERENCES ABOUT PROGRAM IMPACTS

Given the large number of outcome measures examined in this study, several approaches were taken for drawing inferences about which statistically significant treatment–control differences were likely to be due to the influence of the intervention and which were type-1 errors due to chance. In general, to ensure that the approach taken was conservative, statistically significant estimates suggesting *favorable* effects of the program were not considered to be true program effects unless one or more of the following patterns were observed:

- Estimated effects for related measures were also statistically significant in the same direction within that same state (e.g., higher satisfaction with care received should be accompanied by a reduction in unmet needs and higher satisfaction with one or more of various aspects of the care received).
- Estimated effects on this outcome were statistically significant for one or both of the other states (unless there were reasons to expect such differences across states).
- The  $p$ -value was less than .01.

On the other hand, any results suggesting the program had *adverse* effects were not subject to these more rigorous criteria, because we did not want to dismiss any potential negative consequences for consumers, caregivers, or states. We found very few adverse effects of the program on either consumers or caregivers, but Medicaid costs were consistently higher for the treatment group in each state.

## LIMITATIONS AND POTENTIAL METHODOLOGICAL CONCERNS

The study has a number of limitations. These limitations include potential selection bias for some outcomes due to missing data, self-reported data, uncertain generalizability, a short follow-up period, and limited precision for estimating subgroup effects.

The potential selection bias concern is that satisfaction with care and unmet needs outcome measures could not be observed for some sample members. Outcome measures pertaining to satisfaction with paid caregivers were not collected for sample members with proxy respondents who were also paid caregivers, because the respondent would be asked about the quality of the care they personally provided. This restriction disproportionately affected the treatment group. We also lacked data on satisfaction with paid care for those who did not receive such care, which disproportionately affected the control group. These differential sample selection mechanisms may have resulted in bias in the impact estimates. However, we believe this bias was not substantial because (1) roughly equal proportions of the treatment and control groups were excluded for these reasons; (2) the restrictions have countervailing effects (i.e., the former might be expected to bias estimates upward, the latter downward); and (3) we control for a comprehensive set of baseline characteristics. Sensitivity tests showed that treatment–control differences are similar or smaller (in absolute value) for sample members with proxy respondents who are not paid workers, than for sample members who respond for themselves. Thus, if the excluded cases with proxy respondents who were hired workers had been included and had a similar pattern of outcomes, overall treatment–control differences may have been somewhat smaller for some of these outcomes, but still statistically significant and sizeable. See Foster et al. (2003) for a fuller discussion of these sensitivity tests. Absence of observations on satisfaction for sample members who received no paid care surely leads to underestimates of program effects on satisfaction with paid care, as many control group members in two states do not receive any paid care despite being eligible for it. These sample members must, by definition, be dissatisfied with the paid care they are eligible for, because they either could not get an agency to provide it or they considered it so unacceptable that they chose not to receive it.

A second concern is that the evaluators did not directly observe the care provided under the Cash and Counseling program, but instead relied on survey responses from beneficiaries or their proxies about their care. Because

personal care is nonmedical and the consumer is an important judge of its quality, the evaluation's reliance on self-reports of satisfaction, unmet needs, adverse outcomes, and health problems is appropriate. Nonetheless, it is possible that some control group members exaggerated their dissatisfaction, because they were disappointed about not being assigned to the treatment group.

A related concern is that we may have captured the proxy respondent's level of satisfaction, rather than the consumer's level of satisfaction, for many sample members. For children in Florida, we make clear that we are capturing the *parents'* satisfaction, as the parent was always the decision maker and the interviewee. However, for other subgroups that have high proxy response rates, our survey-based results pertaining to care quality often reflect the family member's or caregiver's inference about the sample member's opinion.

Third, there are some limitations related to this study's generalizability. The study pertained to programs implemented in only three states, and thus the findings may not apply to other programs featuring consumer-directed care. Also, the findings can be generalized only to the extent that demonstration participants are representative of those who would enroll in an ongoing program. Those who volunteered for the demonstration may have been particularly dissatisfied with the traditional system or especially well suited for consumer-directed care (perhaps more proactive in their approach to acquiring needed services). Finally, estimated program effects depend, in part, on whether the local supply of home care workers in the area was adequate to meet the demand for services during the period studied. Thus, the results may have been quite different had the evaluation been carried out a few years later than the period studied here (when the labor market was generally tight) or in states where the labor market was tighter or looser than in these three states.

Fourth, program effects were measured after a relatively short follow-up period. Some program effects may not persist over time, as consumers age or lose paid family caregivers. Moreover, consumers' experiences with consumer direction could be unusually positive during the first 9 months of the program because of the novelty of the service model. On the other hand, consumers may learn better ways to manage their care and become more independent over time, so their experiences might become more positive.

Fifth, we cannot completely sort out how much of the program's large effects on informal caregivers was due to the consumer-directed model, and how much was due to the fact that many informal caregivers became paid workers. Across key measures of caregiver satisfaction and well-being, treatment group caregivers who were paid for caregiving during the follow-up period had especially favorable outcomes. For most measures, even treatment

group caregivers who remained unpaid had significantly better outcomes than control group caregivers, but the differences were substantially smaller than they were for paid caregivers. Differences between paid (or unpaid) caregivers in the treatment group and all caregivers in the control group must be cautiously interpreted, however, because of the selection bias inherent in the decision to become paid or remain unpaid. For example, it may be that caregivers who became paid had had more responsibility, on average, for arranging care recipients' personal care than caregivers who remained unpaid, and that caregivers accustomed to this responsibility benefited most from Cash and Counseling. Nonetheless, the fact that unpaid treatment group caregivers had better outcomes than control group caregivers suggests that at least some of the program's effect on caregivers is not due to their becoming paid, but due to other features (such as flexibility) of the consumer-directed model.

Finally, program effects could vary by subgroups not examined here. For example, in Arkansas, beneficiaries who had not received PCS before the demonstration ("new PCS applicants") might be particularly eager to hire a relative under Cash and Counseling, as such beneficiaries had either *chosen* not to receive agency care or lived in an area where agency care is hard to get because of worker shortages. Thus, the program might have larger effects on the receipt of paid care among those who were new applicants than among those who were already receiving PCS at enrollment ("continuing PCS users"; by design, there were very few new PCS applicants in New Jersey and Florida, so this subgroup is only relevant for Arkansas). In early analyses (see Dale, Brown, and Phillips 2004; Dale et al. 2004; Foster et al. 2003, 2004), we did examine whether program effects varied by whether an individual previously received PCS, as well by other subgroups, such as whether the consumer was cognitively impaired, whether the consumer lived in a rural area, and whether the consumer had unmet needs at baseline. Only rarely were there significant differences between subgroups. While the statistical power for these subgroup analyses was limited due to the smaller sample sizes, the point estimates were generally quite similar across subgroups. The exceptions to this pattern were for the subgroups defined by previous receipt of PCS in Arkansas, and by whether had unmet needs for personal care at baseline. Estimated program impacts on a few outcomes were larger (more favorable) for new PCS applicants than for those who were already receiving PCS at enrollment, and for those who had an unmet need at baseline.

None of these potential methodological concerns cast doubt on the basic findings of the evaluation. As shown in the later papers in this volume, the quantitative results are robust, highly plausible, and internally consistent with

each other, with what was learned in the implementation analysis, and with program features.

## ACKNOWLEDGMENTS

This article was based on analyses conducted as part of the Evaluation of the National Cash and Counseling Demonstration, which was jointly funded by The Robert Wood Johnson Foundation (RWJF) and the Office of the Assistant Secretary for Planning and Evaluation (ASPE) at the U.S. Department of Health and Human Services. RWJF provided additional funding for the preparation of this article.

*Disclosures:* None.

*Disclaimers:* The views expressed here are those of the authors and do not necessarily reflect those of RWJF, ASPE, the Cash and Counseling National Program Office, the demonstration states, or the Centers for Medicare and Medicaid Services, whose waivers made the demonstration possible.

## NOTES

- 1 For Florida, elderly adults were defined as those age 60 or older, instead of the more conventional age 65, because nearly all of those age 60 or older were participants in the state's waiver program for older adults (age 60+) with physical disabilities, whereas the great majority of those under age 60 were beneficiaries with developmental disabilities.
- 2 Some adults in Arkansas and New Jersey (as well as elderly adults in Florida) had developmental disabilities, but these people cannot be identified from Medicaid enrollment files.
- 3 As explained in Phillips and Schneider (2007), consumers in the treatment group had to develop a spending plan and get it approved before they could begin receiving an allowance to manage. They could receive agency services until they started on the allowance.
- 4 Funding for this survey was not secured until August 2000, about 1 year after the 9-month consumer interviews had begun in Arkansas. To reach the target sample size in Arkansas, we called back some treatment group members who had already completed their 9-month follow-up before August 2000 to obtain the names and contact information for their primary paid workers.
- 5 We did examine whether we would have found any *adverse* program effects on caregivers or consumers if we had used a higher *p*-value to determine statistical significance (such as .10 or .20). However, treatment-control differences almost always favored the treatment group, and when they did not, the *p*-value was generally much higher than .20.

- 6 The rate of item nonresponse on both the baseline and followup surveys was low, ranging from 0–3 percent per measure. When data were missing for regression control variables, the mean value was filled in so that we could retain all observations for which outcome data were available. Sample members were excluded from analyses for which they did not have outcome data.
- 7 See Carlson et al. (2007) for the proportion of treatment group members who never received an allowance. This occurred for a variety of reasons, including death or entering a nursing home before a spending plan was developed, or because consumers changed their minds about participating or could not find anyone to hire as a worker.

## REFERENCES

- Carlson, B., S. Dale, L. Foster, R. Brown, B. Phillips, and J. Schore. 2005. *The Effect of Cash and Counseling on Service Use and Care Quality in Three States*. Princeton, NJ: Mathematica Policy Research Inc.
- Carlson, B., L. Foster, S. Dale, and R. Brown. 2007. “Effects of Cash and Counseling on Personal Care and Well-Being.” *Health Services Research* 42 (1, supplement 1) DOI: 10.1111/j.1475-6773.2006.00673.x
- Dale, S., and R. Brown. 2005. *The Effect of Cash and Counseling on Medicaid and Medicare Costs: Findings for Adults in Three States*. Princeton, NJ: Mathematica Policy Research Inc.
- . 2007. “How Does Cash and Counseling Affect Costs?” *Health Services Research* 42 (1, supplement 1) DOI: 10.1111/j.1475-6773.2006.00680.x
- Dale, S., R. Brown, and B. Phillips. 2004. *Does Arkansas’ Cash and Counseling Program Affect Service Use and Public Costs?* Princeton, NJ: Mathematica Policy Research Inc.
- Dale, S., R. Brown, B. Phillips, and B. Carlson. 2005. *Experiences of Workers Hired under Cash and Counseling in Arkansas, Florida, and New Jersey*. Princeton, NJ: Mathematica Policy Research Inc.
- Dale, S., R. Brown, B. Phillips, J. Schore, and B. Carlson. 2004. *The Effect of Consumer Direction on Personal Assistance Received in Arkansas*. Princeton, NJ: Mathematica Policy Research Inc.
- Feinberg, L. F., and Whitlach. 2001. “Are Persons with Cognitive Impairment Able to State Consistent Choices?” *Gerontologist* 41: 374–82.
- Foster, L., R. Brown, B. Phillips, and B. Carlson. 2005. *How Cash and Counseling Affects Informal Caregivers: Findings from Three States*. Princeton, NJ: Mathematica Policy Research Inc.
- Foster, L., R. Brown, B. Phillips, J. Schore, and B. Carlson. 2003. *Does Consumer Direction Affect the Quality of Medicaid Personal Assistance in Arkansas?* Princeton, NJ: Mathematica Policy Research Inc.
- Foster, L., S. Dale, and R. Brown. 2007. “How Caregivers and Workers Fared in Cash and Counseling.” *Health Services Research* 42 (1, supplement 1) DOI: 10.1111/j.1475-6773.2006.00672.x

- Foster, L., S. Dale, R. Brown, B. Phillips, J. Schore, and B. Carlson. 2004. *Do Consumer-Directed Supportive Services Work for Children with Developmental Disabilities?* Princeton, NJ: Mathematica Policy Research Inc.
- McHorney, C. A. 1996. "Measuring and Monitoring General Health Status in Elderly Persons: Practical and Methodological Issues in Using the SF-36 Health Survey." *Gerontologist* 36: 571-583.
- Phillips, B., and B. Schneider. 2007. "Commonalities and Variations in the Cash and Counseling Programs across the Three Demonstration States." *Health Services Research* 42 (1, supplement 1) DOI: 10.1111/j.1475-6773.2006.00677.x
- Schore, J., L. Foster, and B. Phillips. 2007. "Consumer Enrollment and Experiences in the Cash and Counseling Program." *Health Services Research* 42 (1, supplement 1) DOI: 10.1111/j.1475-6773.2006.00679.x