Appendix

This appendix provides more detailed information on data sources, creation of research data sets, variable definitions, research methods, research limitations, and analysis details that require further explanation.

Data Sources

The following section of text describes the data collected from each supplying agency.

Texas Department of Human Services (DHS) Data

The primary data source for the experiment was the DHS SAVERR data system, the main repository of client and case information. Most other DHS data systems interact with the SAVERR system, which was modified to include the collection of additional data elements related to the waiver evaluation. DHS developed procedures to extract the necessary data from the SAVERR system for the experimental and control groups, and store the records on an evaluation tape that was made available to RMC for analysis. This tape, together with the monthly TANF client files, constituted the master file of the waiver population. All other files were linked to it by client number or social security number (SSN).

RMC relied upon other DHS data sources as appropriate to supplement the data on the evaluation tape. These supplementary sources included:

- monthly TANF, Food Stamp, and Medicaid client strip tapes covering the demonstration period;
- cumulative warrant files containing historical records of actual cash assistance paid to cases, whether by check or by electronic benefits transfer (EBT), and also including information on amounts deducted for penalties or sanctions;
- transaction files describing the disposition of TANF applications and re-certifications, as well as other case changes;
- penalty files which indicated client spells in penalty status;

- benefits 'ticked' files containing the most accurate record of the exact months of benefit receipt that count toward clients' state time limits; and
- Child Care Management Services (CCMS) administrative and payment data for the time period when DHS was responsible for managing that system (prior to September 1996).

Texas Workforce Commission (TWC) Data

As the administrator for the Unemployment Insurance (UI) program, TWC maintains a wage database system that contains reported employee wages by employer for each calendar quarter. The data identify employees by SSN, and were thus easily linked to members of the experimental and control groups. RMC researchers used these data to measure employment and earnings.

The administration of subsidized child care (SCC) was transferred to TWC from DHS in September 1996. In order to answer research questions regarding the use of SCC, RMC requested Child Care Management Services (CCMS) and child care payment administrative data from TWC and added it to the earlier data from DHS. Case and individual SCC level data included, but was not limited to: 1) spells of SCC receipt, 2) number of children receiving SCC, and, 3) costs of subsidized care.

TWC is also the source of Choices participation administrative data.[1] These client-level data include monthly tallies of actual hours of participation in each Choices component activity. However, the Choices program is not the only source of education, training and job search services that are available to indigent unemployed persons. Other programs administered by TWC offer similar services. Because individuals who are denied Choices participation -- either by their remoteness or because of lack of availability of slots -- may seek services through one or more of these programs, it was necessary to measure participation in other workforce development services. The TWC programs for which workforce participation data were collected included:

[1] The Choices program replaced the Texas JOBS program as Texas' primary employment readiness and training program for TANF recipients.

- Job Training Partnership Act (JTPA) data (June 1994 through June 1999), which was then replaced by Workforce Investment Act (WIA) data from July 1999 onward, and
- Welfare-to-Work (WTW) program data.

Texas Office of the Attorney General Data

In Texas, the Office of the Attorney General (OAG) is responsible for helping custodial parents receive child support from the non-custodial parent of their children. The OAG has developed automated data systems to facilitate the administration of this program. These data systems include detail on case status and demographics, amounts of support paid and owed by non-custodial parents, and share of the support collected that is disbursed to the state and custodial parent. The data were keyed to both SSN and TANF client number when applicable.

With the passage of PRWORA, the OAG developed a new data system so as to fulfill their increased responsibilities required by that legislation. Information from the prior data system, regarding case actions and collections occurring before September 1997, was not available for this study.

Texas Higher Education Coordinating Board Data

Education and training is also available to members of the experimental and control groups through community, technical, and senior colleges throughout Texas. Data are available for both semester-length courses and non-semester length courses from the Texas Higher Education Coordinating Board. Individual-level measures included enrollment in any post-secondary program and whether or not a degree was awarded.

Texas Education Agency (TEA) Data

The Texas Education Agency (TEA) is the source of statewide primary and secondary public education data in Texas, including school enrollment and student assessment scores. Due to a policy interpretation preventing the release by TEA of identifiable individual-level student information for this research, a *deidentified* approach was used for this analysis. Individual-level data were requested and received, but were identified solely by their group membership rather than by individual student identifier. This approach is described more fully below, in the section on creation of research data sets.

TEA staff linked records in the Public Education Information Management System (PEIMS) database via the student's SSN, and other identifying data elements when necessary, including name, date of birth, and gender. The PEIMS data utilized in this study consisted of individual attendance data in the form of the number of days present and the number of days instruction was offered. Also used were PEIMS individual-level dropout information, as well as the number of different school campuses that individuals attended in a given school year, which served as a measure of mobility.

In addition to providing PEIMS data, TEA provided individual scores on the reading and mathematics sections of the TAAS achievement test for children in the experimental and control groups. The Texas Assessment of Academic Skills (TAAS) scores contained in TEA's Student Assessment System provide a measure of student achievement in various subject areas to be used as a gauge for institutional accountability. TAAS reading and mathematics scores and test passage indicators are available for tested students in grades three through eight on an annual basis.

Because of federal privacy laws governing student data, TEA stripped individual student identifiers from the research data file before returning it to RMC. As a result, group membership could be calculated for these children but their records could not be linked to those of their parents.

Texas Department of Health (TDH)

The State Immunization Registry was designed to access and utilize a statewide immunization database, known as ImmTrac. However, as was the case with primary and secondary education data, stringent privacy laws prevented the release of individual-level medical records for this research. These data were also retrieved using a deidentification strategy that preserved information on clients' group membership. Details on this strategy are available in the next section.

House Bill 3054, passed during the 75th Texas Legislature, required the reporting of immunizations to TDH and ImmTrac by all medical providers. Accordingly, the ImmTrac database has been populated with vaccination information from across the state, including input from the Bureau of Vital Statistics (BVS), Women, Infant and Children (WIC) clinics, and the Integrated Client Encounter System (ICES). ImmTrac includes complete base records for children born in March 1993 or later. Additional children are added to the database as they receive vaccinations by reporting providers. ImmTrac currently does not capture information on immunizations provided by private non-Medicaid clinics or physicians. However, because virtually all public health care providers (regional clinics, local clinics, WIC clinics, and other Medicaid providers) contributed information to ImmTrac, data was available for the vast majority of young Medicaid recipients and other poor children who were born in March 1993 or later.

The ImmTrac system is relatively new and its records are, overall, somewhat incomplete. However, records tend to be considerably more complete for very young children (the focus of the State's recent immunization efforts) than for older children. Moreover, because all children are required to have up-to-date immunizations when they begin kindergarten, pre-school aged children were the focus of this aspect of the analysis.

Although ImmTrac contains a wealth of client identification data elements, it does not require that the SSN of the immunized child be reported. Currently, it is estimated that less than one-third of all ImmTrac records are identified by the child's SSN, thus causing linkage between TANF case information and ImmTrac records to be problematic. This difficulty was overcome, however, through use of "probabilistic matching" on a number of non-SSN identifying data elements. In any case, the incidence of failed linkages can be presumed to fall equally on both the experimental and control groups, thus allowing for the computation of unbiased net effects of the treatments. However, since the linkage is not perfect, and since some children's parents did not consent to release of their data, the actual rates of immunization are underestimated.

Texas Department of Protective and Regulatory Services (DPRS) Data

The CAPS administrative data system of the Texas Department of Protective and Regulatory Services (DPRS) captures reports and results of investigations of child abuse and neglect. Records are person-based (a record is kept for each victim, perpetrator, and other household member) and include SSN and DHS client number among the identifying data elements, permitting straightforward linkage to TANF case files. For each investigation documented, information is collected on the nature of the suspected abuse or neglect being investigated, the relationship between the perpetrator and victim, the dates the investigation began and ended, and whether or not the investigation is substantiated. DPRS also provided individual-level data on placement of children into foster care.

The CAPS system was begun September 1996, although records from previous data collection systems were converted and added to CAPS. Data in CAPS with investigation closure dates earlier than September 1996 were considered less reliable than cases with later closure dates, and were thus excluded from analysis.

Creation of Research Data Sets

Outcomes Allowing for Identification of Clients

To conduct the analysis of most outcomes, RMC researchers linked and merged data files from the disparate data sources noted here. The first step in pulling this data together was to assimilate the DHS evaluation tape, together with TANF monthly case/client files, and extract a listing of the SSNs and client numbers of all experimental and control group participants and the dependents on their cases. These extracts of identifying information were sent to some of the agencies providing the data to be linked to records in their files. These linked records were placed in a file to be transmitted back to RMC. In other cases, the agencies sent data covering their entire universe of clients for the time periods of interest, and the linking and extracting was done at RMC. The research team at RMC then created a relational data engine that tied together several individual and/or case-level relational datasets to produce flat files for analysis. The unit of analysis of the flat files differed (case or child-based, monthly or quarterly) according to the research questions they were intended to answer.

Outcomes for Which Deidentification of Clients was Necessary

As noted above, in some cases, strict confidentiality requirements precluded obtaining fully identifiable individual-level data. These restricted data sources included measures on numerous education and immunization outcomes. In response, RMC developed an alternative research strategy that allowed limited measurement of these outcomes without compromising the identities of the individuals whose data were included. This strategy is referred to as *deidentification*.

To achieve deidentification, RMC researchers supplied agencies with lists of clients containing common identifiers (SSN, DHS client number, name, gender, birth date, etc) for all children in the experimental and control groups. Also included on each child's record was a *group number* containing codes that could be used to reconstruct essential group membership information about the child's parent. First, the parent's pilot indicator was represented by codes indicating experimental or control group membership for each of the four experiments (the Clint office in RER Choices was treated separately). Also incorporated into the coding scheme were the parent's pilot assignment date (date of entry into the experiment), which was aggregated to the year of random assignment; service level or "tier," which represents the parent's approximate job-readiness; and whether the case was a one-parent or two-parent case at random assignment. The goal of this aggregation was to produce groups that represented as much information about the case as possible, while also satisfying the deidentification criterion that the smallest group could contain no fewer than five persons.

These client lists were sent to the agencies, which then extracted the necessary records and provided them to RMC with all identifiers removed, and only the aforementioned group numbers attached. These group numbers were then decoded to yield experiment, experimental group, random assignment date, service level, and two-parent case status. This allowed estimation of experimental net effects overall and separately by tier and by two-parent case status. Inclusion of the assignment dates also served to ensure that only post-random-assignment outcomes would be measured. However, due to the inability to link records to demographic characteristics of the parents, no regression analysis was possible using this procedure. Thus, in experiments in which entry effects were expected, it was not possible to determine which portion of an observed effect was attributable to the entry effect (a differential eligibility requirement), and which portion was due to the experiment itself.[2]

Analyzed Variables

Summary Table of Analysis Variables

The following table summarizes the variables used in the analysis.

Table A-1: Summary Table of Analysis Variables

Demographic Control Variables
Caretaker has no record of high school attendance according to DHS records
Caretaker is male
Caretaker is of White race[*]
Caretaker is of Black race[*]
Caretaker is of Hispanic ethnicity[*]
Caretaker's race or ethnicity is not elsewhere classified[*]
Age of TANF caretaker
Number of months on TANF out of the last year
Number of months employed out of the last year
Sum of caretaker's wages over previous year
Two-parent TANF case (formerly AFDC-UP)
Case was recorded by DHS as contaminated by contrary group's treatment
Welfare Dynamics
TANF receipt
Percent of time spent on TANF out of maximum possible
Penalties
Percent of months in penalty status
Child support
Any Choices penalty
Drug abuse
Texas Health Steps
Immunization
Parenting Skills
School attendance
Voluntary quit
Any penalty
Sanctions
Sanction for failure to cooperate with child support enforcement
Sanction for failure to participate in Choices

[2]See, however, the section on analysis details for one solution to this problem.

Summary Table of Analysis Variables (continued)

Welfare Dynamics (continued)

Receipt of non-cash benefits

Percent of time on Medicaid of any kind
Percent of time on food stamps
Percent of time on transitional Medicaid
Percent of time receiving transitional child care

Family Self-Sufficiency

Employment

Percent of quarters during which caretaker had wages of any amount

Earnings

Average quarterly caretaker wages

Average quarterly family wages earned
Percent of quarters during which caretaker wages exceeded 155% of poverty
Percent of quarters during which family earnings exceeded 155% of poverty

Combined income sources

Average quarterly family wages earned plus child support collections **retained by family**
Percent of months in which family earnings plus child support collections retained exceeded 155% of poverty

Workforce Development Participation

Percent participating in Choices program
Percent who ever participated in Choices
Average hours of Choices participation per month
Percent ever participating in JTPA, WIA, or WtW
Percent ever participating in post-secondary education
Percent ever receiving post-secondary degree

Family and Child Indicators

Child support case status and paternity establishment

Percent of families with an OAG child support case open
Percent of families with an OAG case open experiencing new paternity establishment(s) monthly
Of families experiencing new paternity establishment(s), average number of children for whom paternity established
Percent of months in which any child support was collected
Average monthly child support collections

Use of subsidized child care

Percent of cases using subsidized child care
Average number of children using subsidized child care monthly, of families receiving child care
Subsidy per child-month using subsidized child care

Children's immunization

Percent of pre-school children with any immunizations reported in ImmTrac
Percent of pre-school children who are fully immunized (age-appropriate) as reported in ImmTrac

Children's education

School attendance rate
School mobility
School dropout rate
TAAS reading: percent passed
TAAS mathematics: percent passed

Child Protective Services

Rate of foster care placements per month
Rate of substantiated reports of abuse or neglect per month

*Race/ethnic categories are mutually exclusive and exhaustive.

Demographic Control Variables

The educational attainment of the caretaker was modeled by two variables, the first being a binary categorical variable (or dummy) which was set to one if the caretaker's TANF records showed that the caretaker never attended high school. The second educational attainment variable was a flag for not having a valid code for educational attainment. This characterization of educational attainment was used because the education code was missing for a non-trivial proportion (about one percent) of the observations. Systematically excluding these observations would have biased the computation of the net effects.

The age of the caretaker was modeled by using the number of years between the date of the observation and the date of the caretaker's birth. The gender of the caretaker was modeled by using a dummy variable that took the value 1 if the caretaker was male, and zero otherwise. This variable was assumed to be time invariant. Similarly, the race and ethnicity of the caretaker were modeled by including three of four dummy variables indicating race and ethnicity. Four mutually exclusive and exhaustive racial and ethnic categories were used: White, Black, Hispanic, and Other, where "Other" means not White, Black, or Hispanic. The caretaker and the caseworker decide the racial and ethnic identity of the caretaker during the application process.

The concept of contamination as it applies to this evaluation is described in detail elsewhere in this document. Contamination was modeled in a time-varying manner using a pair of dummy variables, one each for the experimental and control groups, that took the value 1 if the case had previously been declared contaminated by DHS, and zero otherwise.

Past employment, past earnings, and past TANF history were computed as of the time of random assignment. The caretaker's UI records and TANF records used in this evaluation spanned a period of time long enough that it was possible to count backwards from the date of random assignment for one year, and add up the caretaker's time employed, wages earned, and time on TANF. For the purpose of this analysis, it was necessary to assume that whenever a caretaker was employed in a quarter, she was employed in all three months of that quarter, and the wages were distributed evenly across the three months of the quarter.

The AFDC-UP or two-parent TANF status of the case was modeled by a dummy variable that took the value one if the case was an AFDC-UP or two-parent TANF case and took the value zero otherwise. The AFDC-UP or two-parent TANF status is indicated in the DHS records with a type program code (TP) of 61 or 63.

Welfare Dynamics

The expression "welfare dynamics" refers to changes in the stock of active cases and the flow of cases into and out of active TANF status. Since the stocks and flow of cases are time variant, their analysis is called "dynamic" in the mathematical sense of the word.

Caseload is determined by the following formula:

$$Caseload_t = Caseload_{t-1} + Entries_t - Exits_t \quad (1)$$

Where:

$Caseload_t$=Caseload at the end of month t,
$Entries_t$=New cases arriving during month t, and
$Exits_t$=Active cases leaving the rolls during month t.

Equation (1) shows that effects on caseload are functionally dependent on effects on entries and exits. The design of the experiment, however, precludes the analysis of changes in entry rates from the general population due to the lack of information on every person at risk of entering welfare. Thus, while the complete analysis of caseload would require examination of effects on both entry and exit rates, only the latter can be modeled here. An alternative approach, which measures the total time receiving welfare after random assignment, arguably captures the most interesting variability in both exit and reentry rates in one convenient measure.

TANF Receipt

TANF receipt was measured as the proportion of time receiving benefits after random assignment, defined at the monthly level with a variable taking the value of one in months for which any TANF benefits were received, and zero otherwise. Mean values of this proportion were then compared between the experimental and control groups. If the experiment induced people to exit welfare and stay away, then the proportion would be smaller for the experimental group than it would be for the control group. This approach summarized the entire impact of the experiments on TANF caseload in a single figure for each experiment that measured the combined effects of changes in entries, exits and recidivism. The only real drawback to this approach is that it is relatively insensitive to 'churning,' or frequent exit and reentry, which would be evidenced by elevations in both the exit and reentry rates.

Non-Cash Benefit Receipt

The same approach that was used to summarize TANF receipt was used in the measurement of receipt of non-cash benefits. Summary statistics estimated the proportion of time receiving Medicaid of any kind, and of receiving transitional Medicaid in particular. Receipt of Food Stamp benefits and transitional child care were measured in a similar manner.

TANF Caseheads and Family Aggregation

A number of benefit receipt measures reported here were computed at the family level, although the data sources are typically maintained at the individual level. This presents difficulties for aggregating to the family level when the structure of cases, or families, changed over time, which tends to happen quite frequently in this population. Children are born, for example, or grow up, move out on their own, and perhaps even open their own TANF cases. Children may also go to live with a grandmother or aunt or uncle, or family members can die, or any number of other changes in the structure of families can occur.

For present purposes, due to the centrality of the TANF program to this evaluation, family measures were aggregated according to the structure of the TANF cases to which the family members belonged.[3] However, this aggregation could only be directly accomplished during time periods when the case is on the active TANF caseload. Some method was necessary to aggregate families during off-TANF periods, and to this end the TANF case structure was projected both backward and forward in time, as follows. First, for each client the TANF record was searched to locate the first month, within two years before random assignment, that he or she was associated with an active TANF case. The case number from that first case was then projected backward from that point in time to the beginning of the two-year pre-study period. This case number served to aggregate families in periods before they entered TANF. Case numbers were also projected forward in time for all remaining off-TANF periods, so that when a family left TANF, the family structure was assumed to remain the same until updated information was received when one or more of them re-entered TANF.

The actual aggregation of family information was done at the case-month level, and was achieved by selecting a casehead and then adding information from all associated children on the same case. The casehead was chosen by a formula that examined all members of a case looking first for a primary caretaker, then a secondary caretaker, then a payee, and finally a case-name only.[4] Whoever was highest in this hierarchy on a case in a given month was considered the casehead for that month. For all analyses reported as 'caretaker' measures, only this chosen casehead was included so as to avoid case duplication and the resulting dependence of observations that would violate assumptions behind most statistical tests. Child-level information was then summarized to the family level, for every certified child on a case, by either counting the number of children, or determining whether *any* or *all* of them received the benefit of interest (TANF, Medicaid, etc.) in that month.

[3] Exceptions include the subsidized child care and child support programs, which have their own case structures. For those programs, outcomes are linked to the TANF casehead, but no attempt is made to verify that the children are the same as those on the TANF case.

[4] These classifications are based on 'status-in-group' codes. Although payees and those who are case-name only have different codes in the DHS data system, we do not distinguish between the two in this report, and refer to both as 'payee' since their needs are not counted in eligibility determination.

Retroactive TANF Benefits

Determining whether a client was receiving TANF in a given month using only the information in the monthly snapshot, or client files, could seriously underestimate the rate of TANF receipt.[5] This is because these snapshots were taken at monthly 'cutoff,' which occurs a week or more before the end of the month, and at which point the subsequent month's caseload is presumably fixed. Changes to this information after cutoff are quite common, however, so the best record of who received benefits for a given month can be culled from the 'warrant file.' This data source maintains, for each case, the total dollar amount of benefits paid, including any money that may have been paid retroactively (after cutoff), or subsequently recouped because of overpayment. The warrant data do not, however, contain client-level information, so the TANF case structure in retroactive benefit months is inferred from the family aggregation scheme described above.

Penalty and Sanction Rates

Penalties could be imposed on members of the experimental sample for reasons highlighted in the penalty tables of the main report. Sanctions could be imposed on the control group members as they were under the old AFDC rules. Under these old rules, control group members could be sanctioned for failure to cooperate with child support or Choices requirements. A penalty involved a reduction in the amount of the TANF grant (similar to a fine). A sanction usually involved the removal of the caretaker from the caseload. In addition to reducing the grant amount as a penalty does, a sanction could also involve the revocation of automatic Medicaid eligibility of the casehead. Since penalties and sanctions are two different methods of punishment, it was not strictly valid to compute the net effect of the experiment by simply computing the difference between the number of penalties imposed on the experimental group and the number of sanctions imposed on the control group in RER Choices, RER Non-Choices, and Clint. Since the control group received no penalties, the net effect of the experiment on penalties was

[5] Analysis by RMC researchers indicates that TANF receipt is underestimated by approximately 10 percent when using only the client files.

simply the number of penalties imposed on members of the experimental group. However, many would find it useful to compute an experimental effect by comparing the penalties in the experimental group to the sanctions in the control group. Thus, for all punishment categories for which there was a sanction that corresponded to the penalty, both penalties and sanctions were tabulated and differences and net effects were computed. In the Time Limits experiment, both control and experimental groups received penalties, and they were analyzed in the usual way.

The analysis of penalties and sanctions was done by computing the percent of months of TANF receipt that were spent in penalty or sanction status (only months of TANF receipt were counted because one had to be receiving TANF to be penalized/sanctioned). The average length of the penalty or sanction was computed by dividing the total months in penalty/sanction status by the total occurrences of that penalty or sanction.

Time-Limit Induced Exits (only in areas with Time Limits)

One of the outcomes of the time limits component of the experiments is that some of the clients on whom time limits provisions are imposed were forced to leave the welfare rolls when their time limits were reached. The magnitude of this impact was not strictly estimable because in none of the experiments did time limits apply to both the experimental and control groups. Analysis of those reaching their time limits was thus purely descriptive, tabulating their outcomes for 12 months following their time-limit induced exit.

Family self-sufficiency

Employment

Some limited data on income of TANF recipients is available through the administrative records of the DHS. However, these data cover only current recipients, are reported only at application or recertification, and are based on self-reported income. Previous work in the area of welfare and employment has shown that UI wage data are superior to self-reported data from administrative welfare records, and were therefore used to measure employment. UI wage data cover over 95 percent of all employment in the state of Texas. Some jobs are not covered, including out-of-state employment, self-employment, federal government employment, and most agricultural employment. Any underreporting due to these reasons would fall equally on both the experimental and control groups.

In measuring employment outcomes, RMC researchers created a variable that took the value of one if the recipient earned money and zero otherwise. Taking the mean of this variable for a group of individuals gave the percent employed for that group. The difference in the rate of employment between the experimental and control groups was the employment impact of the experiment.

Earnings

Previous work with UI wage data has shown that a large percentage of the welfare population earns wages. However, the distribution of wages earned is skewed, with a large proportion of the earners at the low end of the wage scale and very few at the high end of the wage scale. Further, there are many participants with zero wages. Earnings were analyzed by comparing the average amount earned by caretakers in the experimental group to the average amount earned by caretakers in the comparison group. This tabulation gives an overall assessment in a single easily understood number of the experiment's effect on changes in the amount of money available to the caretakers.

Many on welfare earn so little in wages that the earnings cannot be reasonably expected to move them into a state of self-sufficiency. The self-sufficiency standard was measured by creating an indicator of whether the client earned 155 percent or more of the official poverty level, a standard that some authors have maintained is necessary to become totally independent of both cash and non-cash welfare benefits.[6] Using the percent of poverty approach has the advantage that family size is a factor in determining whether earnings are large enough to constitute self-sufficiency.

In addition to the analysis of the wages of the caretakers, a second analysis was performed on total family income. The family income analysis is motivated by the idea

[6] Schwarz, J. and T. Volgy. *The Forgotten Americans.* W.W. Norton and Co. New York, 1992.

that the family moves into or out of poverty and welfare dependency as a unit, rather than individually. While it may be impossible for one individual in a family to bring in wages that are 155 percent of poverty, it might be possible to achieve this goal if more than one person in the family works. Family income data were constructed by aggregating UI wage records for all family members, including second parents on TANF-UP cases. The same analytical approaches used for earnings of the primary caretaker were also applied to total family income.

Child Support Collections Retained by Family

Another influence that may help move a family out of welfare dependency is the collection of child support funds from absent parents. Collection of child support by the OAG on behalf of welfare and former welfare recipients was modeled differently for the outcome of self-sufficiency than it was for child support case status progression (see below for the latter). This was due to the way that the OAG disbursed funds collected from non-custodial parents (NCPs). When child support was collected for a *current* welfare recipient, the state kept the proceeds (except for a $50 disregard) unless the amount collected exceeded the amount of the welfare grant. When child support was collected by the OAG for *former* welfare recipients, however, the entire amount of the ongoing support portion of the payment was forwarded to the recipient. Since the portion of child support payments retained by the OAG could not affect a client's self-sufficiency, only the portion of payments that were forwarded to the clients was counted for self-sufficiency purposes.

Participation in Workforce Development Services

RMC researchers analyzed data on both Choices participation and participation in other workforce development services (e.g. JTPA, Workforce Investment Act, Welfare-to-Work) to determine whether the experiments contributed to differential use of workforce development services. Measures were created to indicate whether clients had ever participated in these programs at any point after random assignment. For those who participated in Choices, the actual number of hours spent in activities was also modeled.

Public community, technical, and senior college data were also analyzed for enrollment and completion of both semester-length courses and non-semester length courses. Measures included whether or not individual caretakers were enrolled in any post-secondary program at any point after random assignment, and whether or not a degree was awarded.

Family and Child Indicators

Child Support Case Status and Collections

In addition to the analysis of child support collections retained by the family, which was reported above as a component of self-sufficiency, child support case status and collections were also studied as an end in themselves. This was done because of a provision in the Personal Responsibility Agreement (PRA) that required cooperation with child support enforcement authorities. Failure of TANF caretakers to cooperate could result in either a financial penalty or a sanction, either of which would have entailed the removal of the adult's needs from eligibility determination for the TANF case. Thus, one possible outcome of the experiments would be advancement of the child support case through some of the steps necessary to achieve collection of child support funds, which could ultimately result in increased collections.

Probably the most basic measure of child support case status progression concerns whether or not a child support case has been opened. This required the cooperation of TANF caretakers to identify, if possible, the father[7], and to provide information to help locate him. Another measure of case progression was created by counting instances of new paternity establishments. Due to limitations of the OAG data system, only new establishments occurring on or after September 1997 could be reported with certainty. Finally, total child support collections, regardless of how much the family was allowed to keep, were modeled as a potential indicator of case progression.

[7] In a small percent of cases the TANF caretaker is the father, but for obvious reasons there is never a question of maternity in these cases.

Subsidized Child Care

Subsidized child care (SCC) services were offered to eligible current and former TANF-recipient families under a number of different programs including Choices, Transitional, and At-Risk (later, income-eligible) Child Care. The ACT waiver demonstration, among other things, changed some of the regulations governing the eligibility for and receipt of Transitional Child Care (TCC) services for recipients who exhausted their time-limited benefits. CCMS and DHS' payment data were analyzed to determine whether any of the demonstration treatments affected the patterns of subsidized child care receipt. The experiment was expected to change both the number of children in SCC, and the number of families that availed themselves of this benefit. Thus, experimental effects were computed for the percentage of cases using SCC monthly, the percent using transitional child care in particular, the average number of children using SCC per subsidized family, and the average dollar amount of subsidy per child-month of SCC receipt.

Immunizations

Immunization data for children under age five at any time during the study were made available from ImmTrac, the statewide immunization tracking system, in a deidentified form like that used for education measures. The coverage of the ImmTrac data is incomplete, due to imperfect linkage, unreported immunizations from some providers, and failure of the parents to consent to release of their children's data. Thus, the immunization rates reported are biased downward from their true values, but there is no reason to expect this bias to operate differentially in the control versus experimental groups.

Two measures of immunization were created, both meant to represent clients' level of immunizations as of the last day of the current study period, September 30, 2001. The first measure was a simple indicator of whether the child had any immunizations on record in the ImmTrac system. The second measure indicated whether the child's immunizations were complete and up-to-date as appropriate for his or her age. This measure, designed in consultation with TDH program specialists, was defined to equal one if the client had not missed a scheduled booster or immunization, and set to zero otherwise. This was deemed to be an accurate measure, since children receiving Medicaid immunizations typically had their next scheduled booster entered in the system. For a client to be considered fully immunized, all types of recommended immunizations were required except for Tetanus, since it was deemed relatively unimportant in comparison to the others. Required immunizations included, for example, Hepatitis-B, Polio, Measles, Mumps, Rubella, Diphtheria, Pertussis, and others.

Education Measures

As described previously, primary and secondary education data for families with children between the ages of 5 and 18 at any time during the study were analyzed using a deidentification approach that protected the identities of clients. The attendance data available for individual students included the number of days present and the number of days for which the student was enrolled and instruction was offered, or 'days in membership.' Net effect computations on attendance were conducted using the ratio of these variables: days present divided by days in membership. Individual-level dropout information was modeled using an indicator that was set to one for those who dropped-out in a given academic year and zero for those who did not. Finally, a measure of mobility was constructed by counting the number of different school campuses that individual students attended in a given school year.

TEA also provided individual scores and test passage indicators on the reading and mathematics sections of the TAAS achievement test. For the net effect analyses, only the test passage indicators were used. This gave an indication of the percent of students who passed each test, but no indication of their scores. The test score measures, known as the Texas Learning Indices (TLI), are most appropriate for measuring changes in individual students' learning of the subject matter. They were used as described in the 'Analysis Details' section below for a repeated-measures analysis of changes in students' scores.

Child Protective Services

The rate of placement of one's children into foster care was considered a further measure of child well being.[8] An indicator of foster care placement was defined to equal one in the month that a child began a spell in foster care, but did not count any subsequent placement into a different foster home during the same spell. A spell was considered to have ended when the child was returned to the primary home. Thus, the measure was a rate of *entry* into foster care, and so a given child could have multiple placements over time. This rate was aggregated to the family level for reporting, such that placement of any of one's children into foster care in a given month was counted as a placement incident for that family.[9]

Since child abuse is one of the most dramatic things that can reduce the well-being of children, RMC used data on child abuse investigations collected by the CAPS system maintained by DPRS. The data contain records of investigations showing who was investigated, and the outcomes of the investigations. Records were retrieved based on a match with child SSNs. A measure of abuse or neglect was defined to equal one in the month that a substantiated investigation started (the date of the abuse event itself was unavailable), and zero otherwise. This measure was aggregated to the family level by considering abuse or neglect for any child to represent an incident of abuse or neglect for the family.

Statistical Methods Employed

The fact that the waiver was designed as a series of randomized experiments justified the application of a large body of available statistical methods specifically designed for estimating the net effects of experiments. These standardized, widely accepted techniques were used to estimate unadjusted and adjusted net effects of the experiments on various hypothesized outcomes.

[8] In the intermediate ACT waiver reports, this foster care placement measure was defined based on foster care codes in DHS Medicaid client data, so results from the current measure may differ from those reported previously.
[9] See the section below on Poisson regression for low-frequency outcomes for a different treatment of the child protective services measures.

ANOVA

The simplest and most obvious approach to estimating the unadjusted net effect of an experiment is to ignore all pre-experiment data, assume that the randomization has resulted in equivalent experimental and control groups, and simply perform a test on the null hypothesis that the post-experimental means of the outcome variables are equal. This post-treatment difference between means is called one-way analysis of variance (ANOVA). If the null hypothesis of no experimental-versus-control difference between post-treatment means is rejected, then the observed difference in means is declared to be statistically significant. This simple test was performed and reported for all outcome measures under the title 'Difference.'

Adjusted Net Effect Regression

In addition to the ANOVA tests, adjusted net effect regression analysis was performed when data permitted. This analysis was used to adjust the estimated net impact by performing a regression in which the dependent variable represented the outcome being analyzed, and the independent variables include descriptors of the characteristics of the case participants. The purpose of computing the adjusted net effect is twofold: 1) to adjust the impact measure for the slight differences in the attributes of the experimental and control groups that inevitably occur in randomly assigned subgroups, and, 2) to provide an impact estimate with a smaller standard error than the simple post-experimental difference in means. In the case of the Clint office, in which the existence of entry effects made the control and experimental groups not strictly comparable, the adjustment procedure provided an estimate of the experimental effect minus the entry effects.

The adjusted net effect regressions contained time-invariant variables measuring high school graduation and race, and contemporaneous measures of age and experimental contamination. Additional variables were included in the regression using their values as of the time of random assignment, including the age of the oldest child and whether it was a two-parent case, as well as TANF receipt, employment, and wage histories over the prior year.

The inclusion of the contamination variables was done to counteract the effects of contamination of the subjects by exposure to the incorrect treatments.[10] Various methods of adjustment for this contamination were tested, and retaining the contaminated observations in the regression was deemed superior to dropping them. The dummy variable procedure produced approximately the same magnitudes for the regression coefficients as dropping the affected cases, but the standard errors were smaller (that is the estimates were more precise) using the dummy variable method.

How the Net Effects were Expressed

Throughout this report most of the outcomes were expressed in percents. For example, one might say that experimental group members spent 42.3 percent of the time after random assignment on TANF, versus 42.9 percent for the control group. This gives a difference of 0.6 percent. However, another way to express the difference would be to say that the experimental group's figure is 98.6 percent as large as the control group figure. (That is, 42.3/42.9=0.986) Expressing the difference this way, one could say the difference is 1.4 percent. (That is 0.6/42.9=0.014) These two views can give quite a different impression of the magnitude of change when comparing rather small differences. For example, if the control group's figure were 1 percent and the experimental group's figure were 2.5 percent, the difference may be expressed as 1.5 percentage points, or a 150 percent change. Since the language used in expressing percents of percents can be ambiguous and confusing, a standard convention was followed. According to this convention, the term 'percentage points' was used when describing the difference between two percents, and 'percent change' when describing the size of the change relative to the value for the control group.

[10] See section below on contamination of treatments.

Potential Limitations of Analysis

The following section presents short descriptions of some of the difficulties in executing this research.

Disparate Sources for Variables

For some measures, the primary data source changed over time. One example of this occurred with SCC variables, in which current SCC data came from TWC and historical data from DHS. This dual provider situation raises the danger that observed changes in the variables may not have been due to an actual change in the amount of SCC, but merely a change in the way these variables were tracked under the two different data systems.

RMC took steps to help ensure that data from disparate sources were as comparable as possible, but in these situations, it was never possible to be absolutely sure that all observed differences were due to the experimental treatments rather than disparate data collection procedures. The only mitigation to this difficulty was the knowledge that any biases in the data should fall equally on the experimental and control groups.

Data Censoring

Because the ACT evaluation was of finite duration, it was impossible to follow all of the experimental participants to the end of time to see how the ACT provisions affected their lives. Since the follow-up period ended for everyone at a single point in time, people who had been in the evaluation from the beginning had a longer follow-up period than those who entered near the end of the period. RMC researchers adopted a number of statistical procedures to handle this data-censoring problem. In the measurement of caseloads, data were tabulated as a proportion of time spent in a particular status out of all possible opportunities to be in that status. In most instances, this was done at the case-month or client-month level, so that clients would contribute to the analysis in direct proportion to how long they received the experimental treatment. In measuring penalty spells, censored spells (ongoing at the end of the study period) were not counted. This should have little effect on the mean penalty spell length, however, since most were uncensored, and the mean spell was only a few months.

Contamination of Treatments

In the implementation of the experiment, contamination of treatments may have taken place for any of several reasons:

- faulty information from caseworkers
- word of mouth and news reports
- other agencies' non-experimental treatments
- movement of TANF recipients between demonstration and non-demonstration sites, and
- changes in family composition of a TANF caseload.

Since the primary point of contact between DHS and the client was the caseworker, the success of the evaluation depended on the ability of the caseworkers to accurately administer the treatments to the clients. If a caseworker incorrectly interacted with a client while unaware of the client's special status as a control group member, there was a possibility that the caseworker may have inadvertently exposed the client to the experimental protocol, and this exposure could have affected the client's behavior.[11] In terms of contamination by word of mouth, control group members would likely behave in a similar manner to their experimental group counterparts because they would have been made aware by word of mouth and news reports that welfare rules had changed statewide. It was a necessary condition for the internal validity of this evaluation that control group members clearly understood that they were not subject to the new rules that were imposed statewide. Also, the various agencies charged with providing social services were continually changing the way they provided services, and many of these changes (i.e., welfare to work grants and workforce development reform) could have impacted the treatments in this evaluation.

[11] Unintentional contamination by DHS staff was evidently not uncommon in the early years of the demonstration. Anecdotal evidence collected by the process analysis team included the description of instances in which both experimental and control group members were present in the same room at the same time for 'group informing' sessions at which time limits were discussed.

To minimize these potential contamination problems, DHS conducted a process evaluation to ensure that procedures used in the ACT demonstration were carried out in a consistent manner. Case-level contamination that occurred due to the changing composition of cases (e.g., an experimental client joining a control case) and client movement was recorded by the DHS data system beginning in January 1997. As can be seen in Table A-2, from 19 to 26 percent of the cases had become contaminated by the end of the study. While this may seem like a large share, one must consider that the bulk of these cases had been in the experiment for more than five years. A more equitable method for examining the risk, or hazard, of becoming contaminated over time would involve event history, or survival analysis techniques. RMC researchers used such methods to estimate how long it would take before half of the participants in each experiment would be contaminated.[12] The results of this analysis, shown in the bottom row of Table A-2, indicate that it would take between 10 and 14 years for half of the participants to become contaminated. From this perspective, the degree of contamination in these experiments does not seem so great.

Table A-2: Case-level treatment contamination

	Experiments			
	Time Limits	**RER Choices**	**RER Clint**	**RER Non-Choices**
Percent contaminated by end of study	19.0%	25.1%	24.1%	26.0%
Time until one-half of participants will be contaminated, projected	13.5 yrs	10.6 yrs	11.8 yrs	9.9 yrs

This type of case contamination was accounted for in the computation of adjusted net effects by the creation of a pair of contamination indicators, one each for members of the experimental and control groups. These indicators took the value of one for client-months in which case-level contamination occurred and persisted at that value into the future, and zero otherwise. This pair of indicators was included in all adjusted net effect regressions, thus controlling for the effects of contamination without compromising the prior equivalence of the experimental and control groups. For the few measures that

[12] For each experiment, RMC researchers fitted an accelerated failure time model using maximum likelihood estimation to estimate the median length of spells in an uncontaminated state, beginning at the time of random assignment.

spanned longer time periods, such as 'ever participated in Choices,' a similar effect was achieved by only counting uncontaminated case-months.

Despite these efforts, not all sources of experimental contamination could be eliminated. Those that remained likely served to reduce the behavioral differences between the experimental and control group clients. As a result, the outcomes reported here should be considered as slightly conservative estimates of the impact of these experimental interventions.

Analysis details

Tests of Random Assignment

The waiver was evaluated by comparing differences in outcomes for randomly assigned experimental and control group members. RMC researchers expected to observe, in a well-done random assignment, that the measurable characteristics of the two groups should differ only by chance. In order to test the hypothesis that the measurable characteristics of the two groups differed only due to chance, researchers performed tests on the means of continuous variables and proportions of qualitative variables that described the two groups.[13]

One of the first characteristics examined was the proportion of participants in the experimental and control groups, with the expectation that equal numbers of participants would be found in each group. Therefore, 95 percent confidence intervals were constructed for each of the four experiments (Time Limits, RER Choices, RER-Clint, and RER Non-Choices). Results are shown in Figure A-1 through Figure A-4. While all experiments exceeded the confidence interval during at least one month for the period covering August 1996 to September 2000, the total number of persons assigned to experimental and control groups for the TL, RER Choices and RER Non-Choices experiments generally fell within acceptable ranges. However, the results for the RER-

[13] Actually, due to the nature of statistical inference, one can expect to find approximately one spurious difference for every twenty comparisons made. This is because the probability of a type I error (wrongly rejecting the null hypothesis) is 0.05 when using a 95 percent confidence level. Due to the large number of comparisons involved in the tests of random assignment, one should only be concerned if the number of significant differences exceeds that which could be expected due to chance alone.

Clint experiment were unexpected. In this experiment, the proportion of participants in the experimental and control groups were significantly different for an unacceptably large number of months, thus increasing the likelihood of significant differences in subsequent analyses.

Other tests of random assignment are reported in the main text of the report. Taken together, all of the tests of random assignment indicate that entry effects are probably the reason the tests failed in Clint. In all RER sites, the experimental group was permitted to have more assets than the control group. In addition, rules governing documentation of recent work experience were eliminated for two-parent recipient families in the experimental groups. Because the demographics of TANF families in Clint included a disproportionate share of families meeting these criteria, the experimental group is larger in Clint. As was reported in the main body of the report, their demographic attributes were also different.

Figure A-1: Bounds for Time Limits

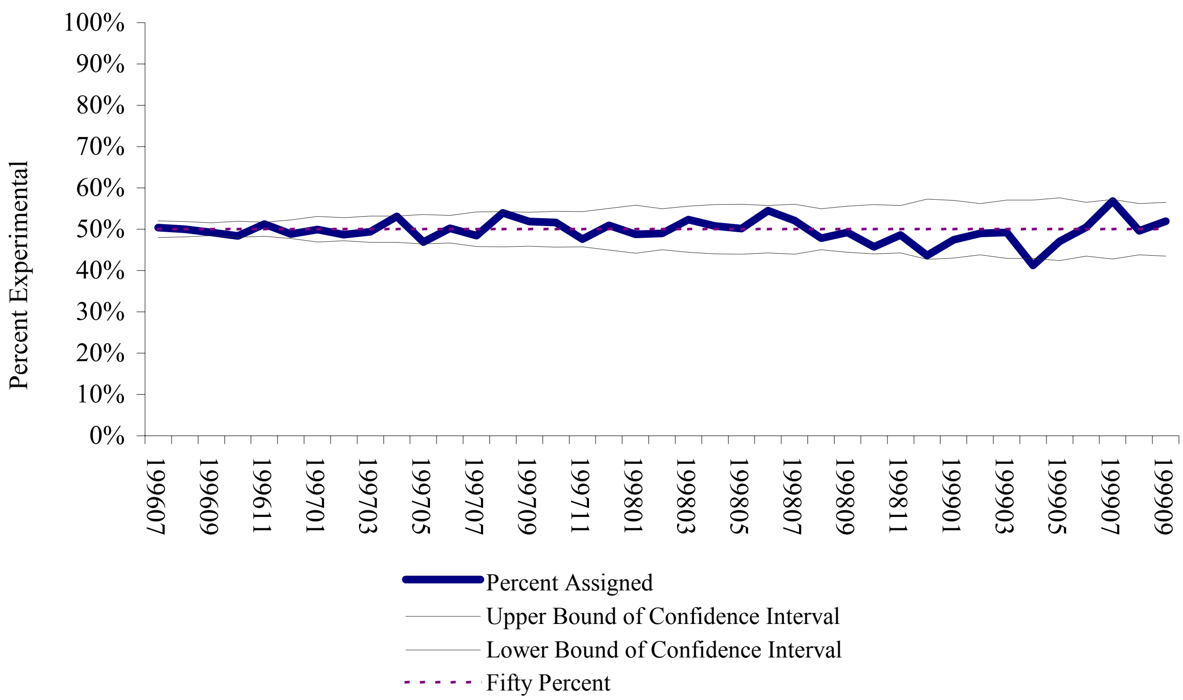


Figure A-2: Bounds for RER Choices

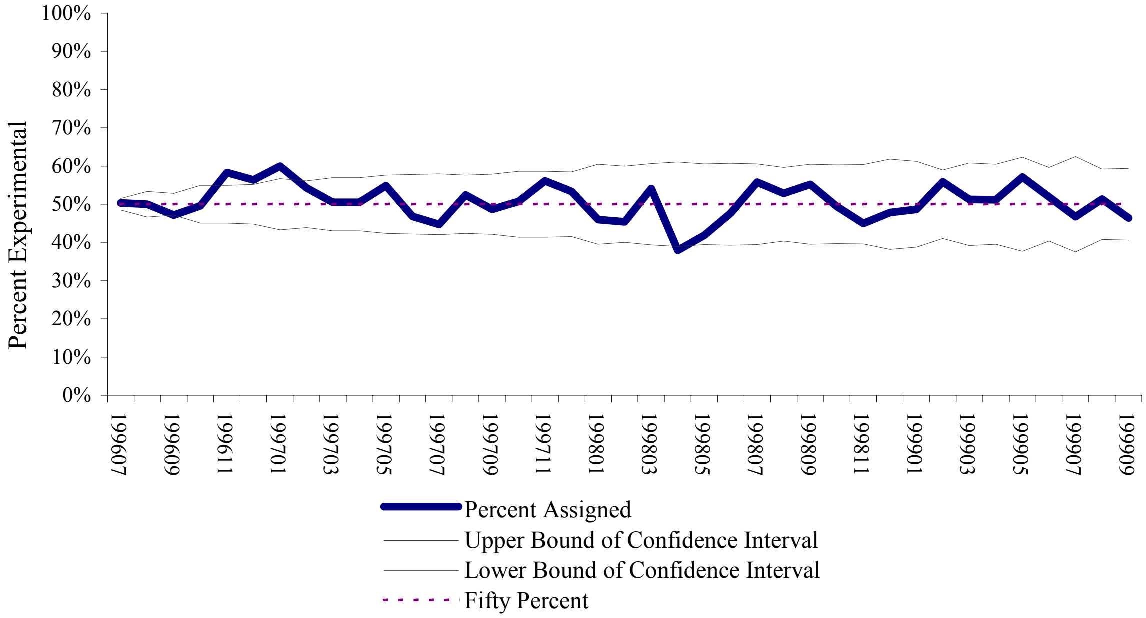


Figure A-3: Bounds for RER Clint

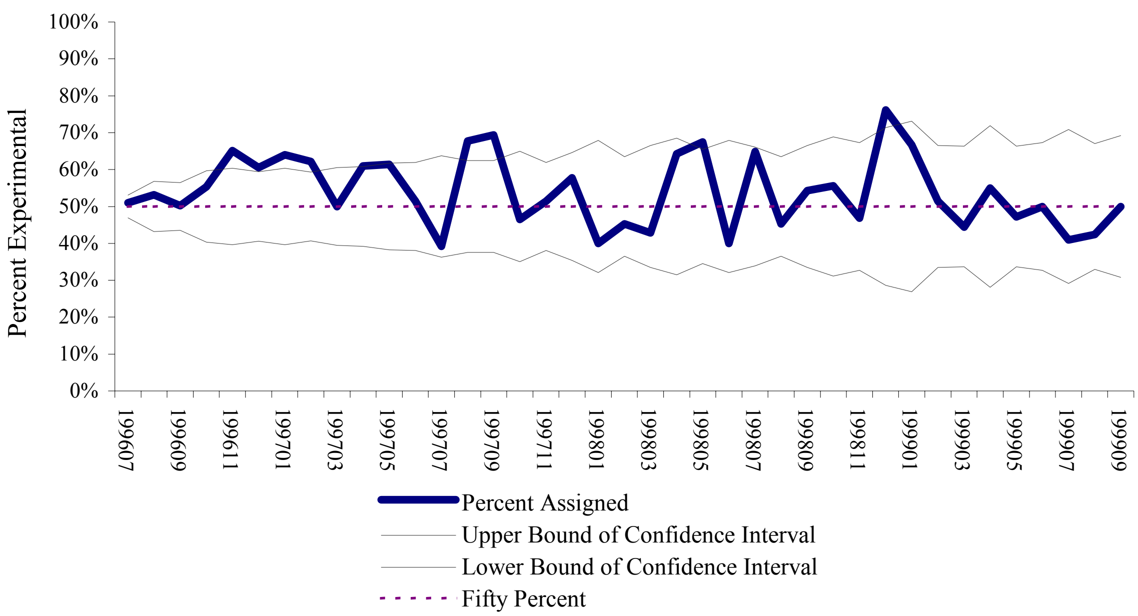
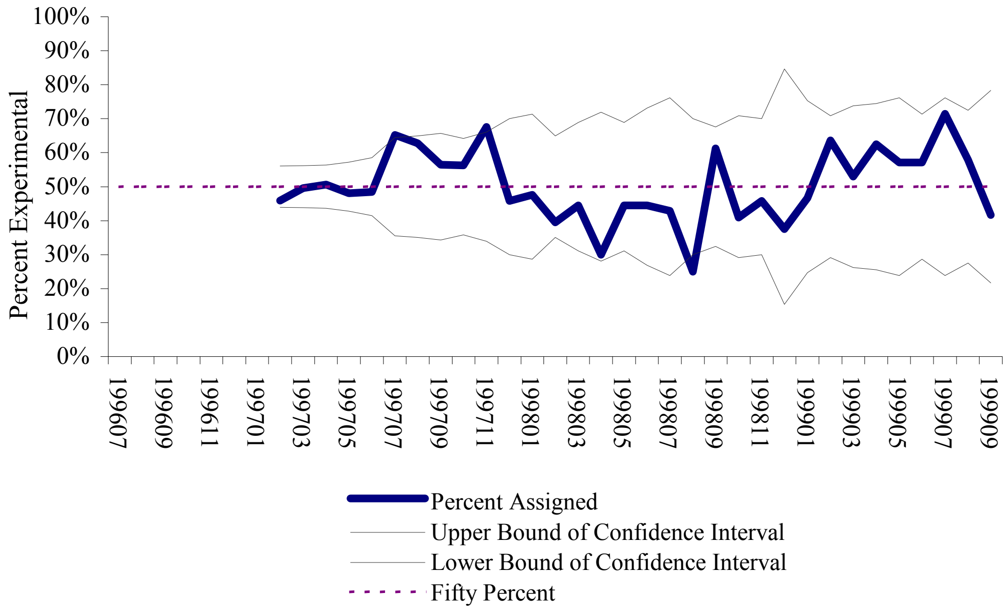


Figure A-4: Bounds for RER Non-Choices



Overall Effects and Subgroup Analysis

Overall net impacts and adjusted net impacts for all variables were first computed for all months following random assignment for all persons assigned to each experiment. These values constitute the overall effects of the evaluation and are reported first for each measured variable. This overall effect has drawbacks, however, in that it has the potential to conceal a great deal of heterogeneity of effects. It is possible that the results of these experiments could vary depending on the time period during which someone was assigned to the experiment, key personal characteristics of individuals at random assignment, or the total length of time available after random assignment available for observation. To test for these differential effects, analyses of twelve key variables from each experiment were conducted for several subgroups.

Subgroup analyses were conducted for the following variables: percent of time on TANF for any family member, average monthly TANF benefit, percent of months in child support penalty status, percent of months in Choices penalty status, percent of time on Medicaid both for caretaker and for any child, percent of time on food stamps, percent of quarters of employment, average quarterly wages, percent of months of Choices participation, percent of months in which child support was collected, and percent of cases using subsidized child care each month. If differential effects were observed for any of these subgroups, they are presented and discussed following the discussion of overall impacts for that variable.

Four-year outcomes

This analysis included all persons randomly assigned early enough in the experiment so that four-year outcomes could be computed. This test was performed for all caretakers assigned within six months of the beginning of each experiment. To test whether the experiments had different effects for those with longer histories with the welfare system, measures were computed separately for short-term and long-term TANF recipients (those with less than 30, and 30 or more months of prior welfare receipt as of random assignment).

Before/after policy changes

This analysis included persons assigned before and after several key policy changes implemented between October 1999 and March 2000. These include the imposition of federal time limits, an expansion of the earned income disregard, and a tightening of the 'age of child' exemption, and are referred to as '1999 Policy Changes'. One-year outcomes were computed for recipients assigned prior to and following these key policy changes to judge whether the experiments had different effects under the new policy environment.

Tier group

This analysis included separate calculations of impacts for members of Tiers 1, 2, and 3 and a test of whether the overall impact for each measure varied by tier.

Poisson regression for low-frequency outcomes

For several of the outcome variables that were modeled as percent-of-time measures- including instances of foster care placement, child abuse or neglect, and receipt of a post-secondary degree- the outcomes of interest occurred very infrequently, often substantially less than one percent of the time on average. While ordinary least squares (OLS) regression is commonly known to provide robust effect estimates even in the presence of non-normality in the distributions of the dependent variables, given sufficient sample size, these low-frequency outcomes present some rather severe departures from normality. Thus, in order to confirm the OLS effect estimates that are reported in the main text, Poisson regressions were also conducted for these outcomes.[14]

The Poisson distribution is most appropriate for describing data that consist of counts of low-frequency events. Thus, measures were constructed that counted the number of events (e.g., foster care placements) that occurred within a given number of units of time or trials (e.g., months available for observation following random assignment). These event and trial variables were subjected to regression modeling using maximum likelihood (ML) estimation, and the results were corrected for the common problem of overdispersion by setting the scale factor equal to the ratio of the deviance to the degrees of freedom.

Results of these regressions, run separately by experiment on the outcomes of foster care placement, child abuse or neglect, and receiving a post-secondary degree confirmed the OLS results reported in the main text. Only one effect was statistically significant when controlling for the same baseline characteristics as the adjusted net effect regressions. Those subject to time limits were found to have significantly reduced odds of receiving a post-secondary degree, relative to controls (b=-.34, Wald Chi-square(1, 32767)=7.70, $\underline{p}$=.006). This parameter estimate can be interpreted as a 28.6 percent reduction in the odds of getting a post-secondary degree for those subject to time limits.

[14] Poisson regression was also used to model school mobility, defined as the number of times one *changed* schools within an academic year. See details in the following section.

Repeated-measures analysis of deidentified measures

A major limitation of the deidentification approach was the impossibility of computing adjusted net effects, which might have allowed for statistical control of family characteristics. As discussed previously, this prevented the separation of entry effects due to changes in eligibility requirements (for the RER experiments, and Clint in particular) from the other effects of interest due to changes in the TANF program. A solution to this problem was achieved for primary and secondary education measures (excluding dropout rate) through implementation of a repeated-measures, or fixed-effects, type of design. This allowed for statistical control of all pre-existing differences (even unmeasured differences) in the control and experimental client populations. Thus, any remaining effects could be attributed solely to the experimental intervention, but not including entry effects, which would have been controlled for as well. Repeated measurements were not available, however, for the immunization statistics.

Although the education data were deidentified as described above, they were provided in such a way that repeated measurements across the four academic years for which data were received (1997-98, 1998-99, 1999-2000, 2000-01) could be linked for individual students. In order to maximize comparability with the 'overall' results that form the core of this report, as well as the subgroup analysis by tier, one or more repeated measurements were made for each client as follows. The baseline measurement for any given client was defined as the earliest valid measurement that occurred in the academic year ending in either the year before or the year of random assignment for that client. Follow-up measurements were defined to include all valid measures reported for that client in academic years ending in the year after random assignment or later. Any client who did not have both a baseline and one or more follow-up measures, as defined here, was dropped from this analysis.

Overall

Regressions were conducted for each of the four experiments (treating Clint separately from the remainder of RER Choices) to separately predict follow-up measurements of TAAS reading and math (using Texas Learning Index scores), attendance, and mobility. These regressions contained only two predictors: an indicator of experimental group membership and the baseline measure. Inclusion of the baseline measure as a covariate has been shown to be the statistical equivalent of repeated-measures ANOVA, and also of a simple ANOVA on difference scores.

The results of these analyses can be summarized as follows. For the attendance and TAAS reading and math score measures, the inclusion of the baseline measure as a covariate reduced the effect of all experimental interventions to nonsignificance (all ps>.01). Thus, it appears that the experimental impacts reported for TAAS reading and school attendance in RER Clint were actually due to pre-existing differences between the experimental and control groups on these measures, most likely caused by entry effects.

A similar repeated measures approach was used to test for effects on school mobility, but a Poisson regression (described above) was used due to the fact that the mobility measure consists of *counts* of numbers of schools attended, which has a distinctly non-normal distribution. This analysis produced two interesting reversals of the overall effects cited in the main text. First, similar to what happened with the other education effects cited above, the effect of RER Non-Choices on school mobility was eliminated when baseline mobility was included as a covariate (Likelihood-Ratio (LR) Chi-square(1, 1123)=1.80, p=.18). Second, the lack of an effect on school mobility that was reported for the RER Choices experiment was revealed to actually be an effect when baseline mobility was held constant (b=-0.17, LR Chi-square(1, 3877)=9.23, p=.002). In this case, the experimental parameter estimate can be interpreted to mean that for the average child the RER Choices experiment caused a 15 percent reduction in the odds of changing schools in a given year.

Subgroups

Subgroup analysis was done to determine whether the experimental effects on changes in, or repeated measures of, education outcomes depended on the tier level of the caretaker.[15] Results indicated that the experimental effect on increases in TAAS reading scores varied by tier only in RER Clint (F(2, 487)=4.86, p=.008). This effect is difficult

[15] The limitations of the deidentification approach prevented analysis of the before-after policy change or the short-term/long-term recipient impacts on repeated education measures.

to interpret, however, since the experimental effect did not differ from zero in any of the three tiers (all ps>.20). The effects on increased TAAS math and attendance varied by tier only in RER Non-Choices sites ($F(2, 432)=6.67$, $p=.001$; and $F(2, 1119)=6.05$, $p=.002$). These effects are easier to describe, since for both measures the effect of RER was significant only in Tier 3, and was quite a large effect in both instances. Parameter estimates indicate that TAAS math scores for children of caretakers subject to RER Non-Choices in Tier 3 increased by approximately 9.5 points (the average was 76.4 points; $F(2, 432)=6.67$, $p=.001$), while attendance increased by approximately 4.8 percentage points (the average attendance was 92.8 percent; $F(2, 1119)=6.05$, $p=.002$).

Poisson regression on the changes in student mobility revealed that the experimental effects varied by tier only in the RER Choices experiment (LR Chi-square(1, 3873)=12.32, $p=.002$). The parameters suggest that there was no effect of RER on school mobility in Tier 1, but those subject to RER provisions in Tier 2 had increased mobility, and those in Tier 3 had decreased mobility.