RAND Statewide CalWORKs Evaluation

Welfare Reform in California: Design of the Impact Analysis

Preliminary Investigations of Caseload Data

Steven Haider, Jacob Alex Klerman, Jan M. Hanley, Laurie McDonald, Elizabeth A. Roth, Liisa Hiatt, Marika Suttorp

MR-1266.0/1-CDSS

July 2000

Prepared for the California Department of Social Services

LABOR AND POPULATION

PREFACE

In response to the national welfare reform legislation of 1996, the Personal Responsibility and Work Opportunity Act (PRWORA), California passed its own welfare legislation on August 11, 1997. The legislation replaced the existing Aid to Families with Dependent Children (AFDC) and Greater Avenues to Independence (GAIN) programs with the California Work Opportunity and Responsibility to Kids (CalWORKs) program. The California Department of Social Services (CDSS) administers the CalWORKs program. Following an open and competitive bidding process, the CDSS awarded a contract to RAND to conduct a statewide evaluation of the CalWORKs program.

This report serves as an appendix to the impact analysis report, *Welfare Reform in California: Design of the Impact Analysis* (MR-1266.0-CDSS; Jacob Alex Klerman, et al, 1999), and provides detailed information on the welfare caseloads in California as of September 1999.

For more information about the evaluation, see: http://www.rand.org/CalWORKs or contact:

Jacob Alex Klerman
RAND
1700 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
(310) 393-0411 x6289
klerman@rand.org

Aris St. James
CDSS
744 P Street, MS 12-56
Sacramento, CA 98514

(916) 657-1959
astjames@dss.ca.gov

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Many people provided their expertise to the acquisition and processing of the data used in this report, and we are grateful to each of them. Unfortunately, we can only explicitly acknowledge a few of these people here.

Many state and county officials gave generously of their time and expertise to transfer the data sets to us and help us understand what the data sets contained. We specifically thank Tom Burke, Aris St. James, Levi St. Mary, Paul Smilinack and Virginia Uchida.

Within RAND, Rodger Madison, Deborah Wesley and Shaoling Zhu provided much programming expertise and Natasha Kostan, Joan Verdon, Christopher Dirks and Paul Steinberg provided invaluable assistance in preparing the document for publication. Robert Schoeni provided many useful substantive comments. Finally, the publications department, particularly Phyllis Gilmore, Betty Amo and Benson Wong, produced the final document under significant time constraints.

1. INTRODUCTION

BACKGROUND

The Personal Responsibility and Work Opportunities Reconciliation Act of 1996 (PRWORA) fundamentally changed the American welfare system, replacing the Aid to Families with Dependent Children (AFDC) program with the Temporary Assistance for Needy Families (TANF) program. In addition, PRWORA deliberately and decisively shifted the authority to shape welfare programs from the federal government to the individual states. California's response to PRWORA was the California Work Opportunity and Responsibility to Kids (CalWORKs) program. CalWORKs is a work-first program that provides support services to help recipients move from welfare to work and toward self-sufficiency. In addition, CalWORKs imposes lifetime time limits to motivate recipients to make these transitions. Finally, CalWORKs further devolves much of the responsibility and authority for implementation to California's 58 counties, increasing counties' flexibility and financial accountability in designing their welfare programs.

The California Department of Social Services (CDSS)—the state agency in charge of welfare—contracted with RAND for an independent evaluation of CalWORKs to assess both the process (or implementation) and its impact (or outcomes), at both the state and county levels. RAND has released the findings of the first phase of the process analysis in a series of documents[1]; two follow-on process analysis reports for the subsequent two phases are due to be released in 2001.

---

[1] See Zellman, G., J. Klerman, E. Reardon, D. Farley, N. Humphrey, T. Chun, and P. Steinberg, *Welfare Reform in California: State and County Implementation in the First Year*, MR-1051-CDSS, Santa Monica, CA: RAND, 1999; Zellman, G. , J. Klerman, E. Reardon, and P. Steinberg, *Welfare Reform in California: Results of the 1998 All-County Implementation Survey, Executive Summary*, MR-1051/1-CDSS, Santa Monica, CA: RAND, 1999; Ebener, P., and J. Klerman, *Welfare Reform in California: Results of the 1998 All-County Implementation Survey*, MR-1052-CDSS, Santa Monica, CA: RAND, 1999; and Ebener, P., E. Roth, and J. Klerman, *Welfare Reform in California: Results of the 1998 All-*

RAND is now working on the first phase of the impact analysis component of the evaluation, the results of which are scheduled for release in October 2000 and October 2001. As part of conducting the impact analysis, RAND is working with various administrative data sets to gather data about welfare caseloads.

OBJECTIVE

The primary objective of this report is to present background information regarding administrative data sources that contain welfare caseload information. Toward this end, we discuss the underlying structure of the data sources and present some preliminary tabulations to demonstrate their usefulness.

We examine three administrative data sets in this report. Two of the data sets are official county reports to CDSS: the CA237 and GAIN25. The CA237 contains information on AFDC/TANF application and caseload activity. The GAIN25 contains information on welfare-to-work (WTW) caseload activity.

Although both of these data sets provide up-to-date official caseload counts, they provide only aggregate caseload counts (for the Family Group and Unemployed Parent programs) for each county. Thus, these data do not allow us to answer many interesting questions, such as whether there have been substantial changes in the demographic composition of the welfare caseload. To answer these types of questions, we use the MediCal Eligibility Determination System (MEDS). The MEDS is an individual-level database that is primarily used to verify MediCal eligibility but also contains information about who is on AFDC/TANF. This database contains information about the age, race/ethnicity, county of residence, and month-to-month status changes for individuals on welfare.

To explore the quality of these data, we present four types of tabulations in this report. First, we examine the level and trend of the AFDC/TANF caseload during the last 15 years. Second, we examine the level and trend of AFDC/TANF application information. Third, we provide

---

County Implementation Survey, Appendix, MR-1052/1-CDSS, Santa Monica, CA: RAND, 1999.

detailed information about the characteristics of the AFDC/TANF caseload, including race, family size, and duration on aid. Finally, we examine the level and trend of the WTW program participation.

Preliminary findings include the following:

- After a substantial increase in the California welfare caseload between 1988 and 1994, the caseload remained fairly constant until 1996 and then began to decline dramatically.
- Most of the change in new welfare cases has come from changes in the number of applicants rather than from the approval/denial rate of applications.
- Although the trends for geographic regions moved disparately during the late 1980s, regional trends have been quite similar during the 1990s.
- The proportion of individuals who were short-term welfare users decreased during the early 1990's and then returned to previous levels by the late 1990's.

Overall, we find that the use of administrative data provides a rich and timely method to evaluate changes in the California welfare caseload.

ORGANIZATION OF THIS REPORT

Because this appendix is intended to outline the administrative data sources we rely on for the larger CalWORKs evaluation, we organize this report by data source. Sections 2, 3, and 4 provide an overview of the three data sources—the CA237, the GAIN25, and the MEDS, respectively—the issues that are faced in their use, and descriptive tabulations from each data source. We provide a brief discussion and conclusions in Section 5. Appendix A provides technical details on our analysis procedures, while Appendix B provides supplemental results.

2. CA237: CASELOAD ADMINISTRATIVE DATA

The CA237 Family Group/Unemployed Parent (FG/UP) Statistical Report is completed by every county in every month to report caseload information to the state. The form contains information on the following four characteristics of the welfare caseload:

- The number of applications received, approved, and denied during the month;
- The level of the total caseload and the number of cases commenced and terminated during the month;
- Information disaggregated by FG versus UP cases; and
- The total expenditure on welfare benefits and total collection of revenues from child support payments.

The form is currently being revised to collect information on additional aspects of the TANF caseload that are relevant because of the changes in California welfare policy.

An advantage of the CA237 is that it contains the official county welfare caseload statistics. With these data, we will be able to answer detailed questions about aggregate (at the county level) application and caseload movements. For example, we will be able to determine whether the recent decline in the California caseload is associated with increased exit rates, with a decline in application levels, or with the increased rejection of applications. Such answers will provide important information about the underlying causes of the change in AFDC/TANF caseload and expenditures.

DATA ISSUES

The data file provided by CDSS contains information for each of the fields listed in the form. For most of the tabulations we present, it is clear what items from the form we use. However, two tabulations require clarification. First, we use the total number of cases that received cash grants during the month (section B, item 8a) to report

caseload totals, rather than the item entitled caseload totals (section B, item 8). This decision accords with how the CDSS reports caseload totals. Second, to report expenditures per county, we use gross expenditures (section C, the sum of items 1 and 1a) rather than net expenditures (section C, item 1), where net expenditures subtracts the child support that the county was able to collect. We use gross expenditures because we are interested in program generosity.

A few additional fields exist in the underlying data set that are called "adjustment" fields. These fields are intended to reconcile month-to-month inconsistencies in the data. For example, the number of applications left pending at the end of one month should equal the number of applications pending at the beginning of the next month. However, because these fields are not consistently used and the month-to-month discrepancies are quite small in size, we do not use the adjustment fields in our analysis. Furthermore, when we compare tabulations to CDSS published results, it does not appear that CDSS uses the adjustment fields either.

In this report, we analyze CA237 data for the months July 1985 to March 1999. Monthly updates are available with a four-month lag; for example, we should receive January 2000 data in April 2000.

Overall, the quality of the CA237 appears to be quite high. We can match the CA237 numbers to those published by the CDSS and the CDSS uses these numbers for their own planning.[2]

TABULATIONS

We first present statewide totals, followed by results disaggregated by region and level of urbanization.

Statewide

In Table 2.1, we present information on the quarterly TANF caseload for the state of California.[3] For the first quarter of 1999, there were 632,796 cases per month receiving cash grants in California, of which 83

[2] CDSS (1997) tables match the numbers we present here, and CDSS (1998) uses the CA237 for forecasting purposes.

[3] All dollar figures are deflated with the monthly CPI-U to January 1998 dollars.

percent were FG cases and 17 percent were UP cases. There were 1,825,811 recipients per month (i.e., persons per month) associated with these cases, implying that 4.40% of the population in California were receiving cash grants. In total, California spent $316 million per month on these cases, representing $499 per case.[4]

Table 2.1

Average Monthly TANF Caseload for California

	Level 1Q 1999	% Change from 4Q 1998	% Change from 1Q 1998
Average Monthly Caseload	632,796	-2.18	-12.1
FG	525,060	-2.63	-12.1
UP	107,735	0.03	-12.1
Total Recipients	1,825,811	-2.49	-13.9
Percent of Population Receiving Aid	4.40	-2.49	-13.9
Total Expenditures	$315,646,861	-0.27	-8.7
Expenditures per Case	$499	1.95	-3.9

These levels represent substantial changes from the previous quarter and previous year. Specifically, there was a decline in caseload of 12 percent as compared to the same quarter one year ago, with similar declines for the FG and UP groups separately. Total expenditures declined less rapidly, with a 9 percent decline as compared to the first quarter of 1998. The decline in expenditures per case was less (3.9 percent). Changes with respect to the previous quarter should be interpreted with caution because many of the figures have strong seasonal components.

We present application information in Table 2.2. There were 52,514 open applications per month during the first quarter, 34,628 of which were new applications with the rest pending from the previous month. These levels imply that there were 0.13 percent of the population had an

[4] Actual benefit levels for a case are set at the state level. These benefit levels vary according to family size, whether the family lives in a high-cost or low-cost county, and family income. The level reported here should be interpreted as the average benefit levels across all families.

application for aid pending.[5] Of these applications, 33 percent were approved, 20 percent were denied, 34 percent were left pending, and the rest experienced "other action," where other action includes cancellations and withdrawals. The application pool was 8.7 percent smaller than it was one year ago and the approval rate was 0.7 percent higher.

Table 2.2

Average Monthly TANF Applications for California

	Level 1Q 1999	% Change from 4Q 1999	% Change from 1Q 1998
Total Applicant Pool	52,514	-3.47	-8.7
Pending Applications	17,886	-3.35	-12.8
New Applications	34,628	-3.54	-6.4
Percent of Population with Pending Application	0.13	-3.47	-8.7
Action on Application			
Approved	33%	-1.02	0.7
Denied	20%	-0.59	-0.6
Other Action	14%	0.48	0.8
Left Pending	34%	1.13	-0.9

Longer trends for caseload levels and expenditures are presented in Figures 2.1, 2.2 and 2.3. These figures, as well as many that follow, have been normalized so that the level for March 1995, the statewide caseload peak, is equal to 100; a vertical line marks this peak in this and all subsequent figures that have been normalized. It is clear from Figure 2.1 that the state's caseload increased fairly slowly from July 1985 until July 1990 and then increased more rapidly to the peak in March of 1995. After remaining relatively constant for a year, the caseload began to decline quite rapidly from July 1996 until the final month of our data. In Figure 2.1, we also present the trend for the United States overall.[6] Declines in California have not been as large

[5] The population data are Intercensal Estimates from the U.S. Bureau of the Census.

[6] The U.S. caseload data were obtained from the Office of Planning, Research and Evaluation at the Administration for Children and Families, the United States Department of Health and Human Services.

as in the United States. Such an outcome is consistent with the recession in the early 1990s being more severe in California than in the United States overall and that CalWORKS was implemented later than similar PRWORA legislation in other states (see Zellman, et al, 1999); however, a formal analysis of this possibility is beyond the scope of this report.

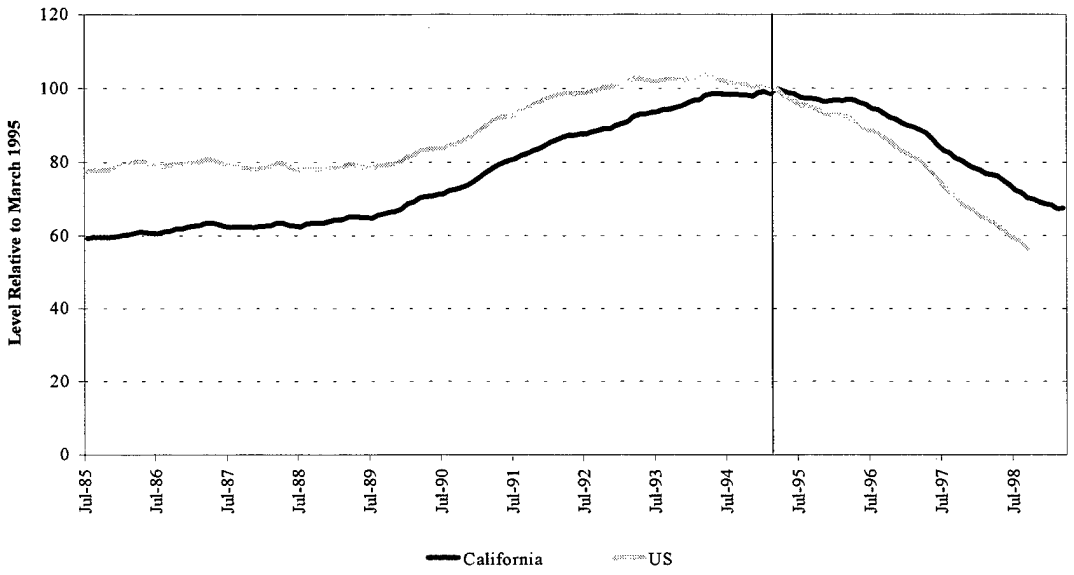


Figure 2.1--AFDC/TANF Total Recipients - U.S. and California

Figure 2.2 demonstrates that total expenditures follow a pattern similar to that of the total caseload, except that total expenditures declined more rapidly during the 1995-1999 decline (the caseload declined by 32.4% and expenditures declined by 47.2%). Together, these trends imply that expenditures per case tended to decline over the entire period, directly observable in Figure 2.3. These figures demonstrate that the increase in welfare expenditures in California during the early 1990s resulted entirely from increasing caseload levels rather than from increasing "generosity."[7] In fact, "generosity" declined by 39.4% between 1989 and 1999.

[7] We use generosity to refer to changes in the average payment received, which is affected by changes in the payment schedule and in family characteristics. We discuss the interpretation of these changes below.

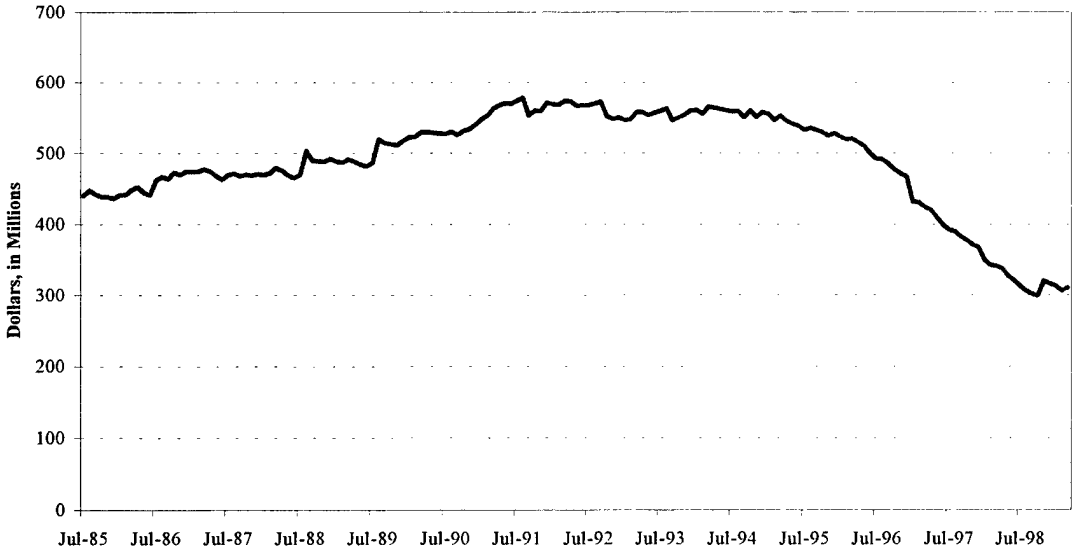


Figure 2.2--AFDC/TANF Total Expenditures - California

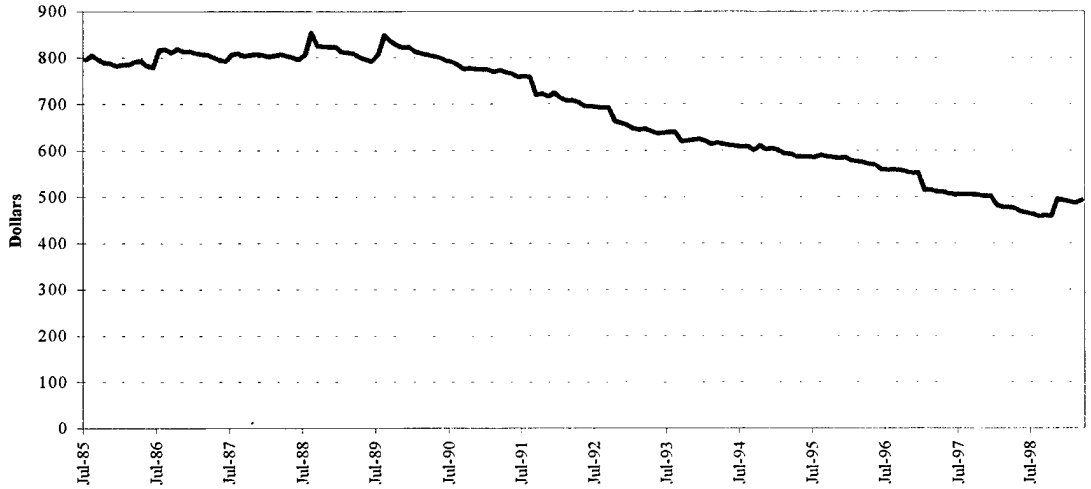


Figure 2.3--AFDC/TANF Expenditures per Case - California

We present application trends in Figures 2.4 and 2.5. In Figure 2.4, we see that new applications increased during the early 1990s, peaking earlier than the total caseload level and that the number

declined a few years earlier.[8] In Figure 2.4, we also present application information that is seasonally adjusted.[9] The seasonal adjustment reduces the monthly variation in the application trend. In Figure 2.5, we present the percentage of applications approved each month (relative to the number of applications processed). The approval rate is relatively constant during the 11-year period, with a slight downward trend. Taken together, these figures suggest that the changes in caseload during the 1990s are largely driven by changes in the number of new applications rather than by changes in the approval rate.

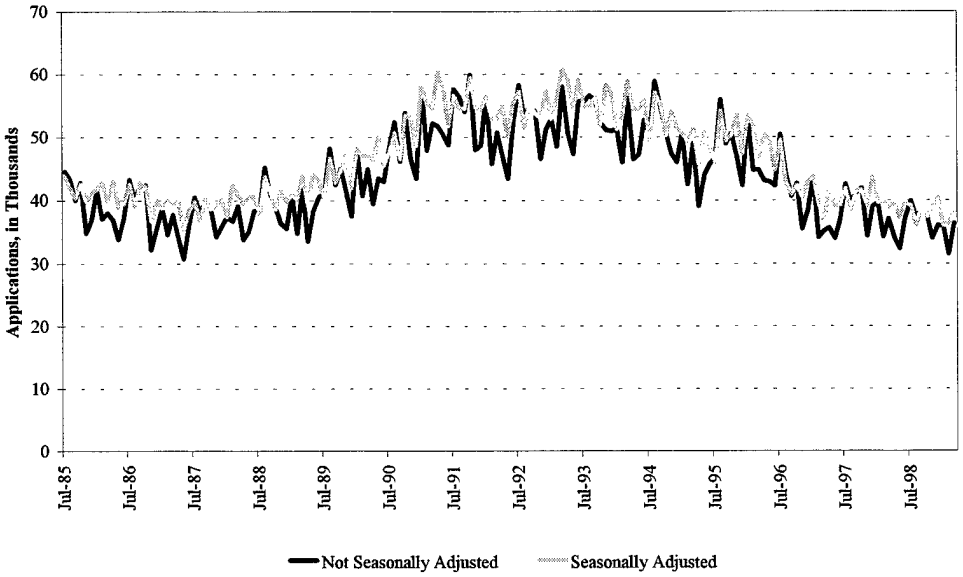


Figure 2.4--AFDC/TANF Total New Applications - California

[8] This finding is expected given the "stock-flow" relationship between applications and the caseloads. Quite simply, an increased flow onto welfare (new applications/new cases) will cause the stock (the caseload) to increase all else equal. Furthermore, because of the application processing time lag, we expect the new application flow to increase before the stock (the total caseload).

[9] We plot a "seasonally adjusted" time series in order to focus on underlying secular trends. To adjust the data, we use a regression-based method to remove the predictable monthly variation. Further details on the adjustment technique are provided in Appendix A.

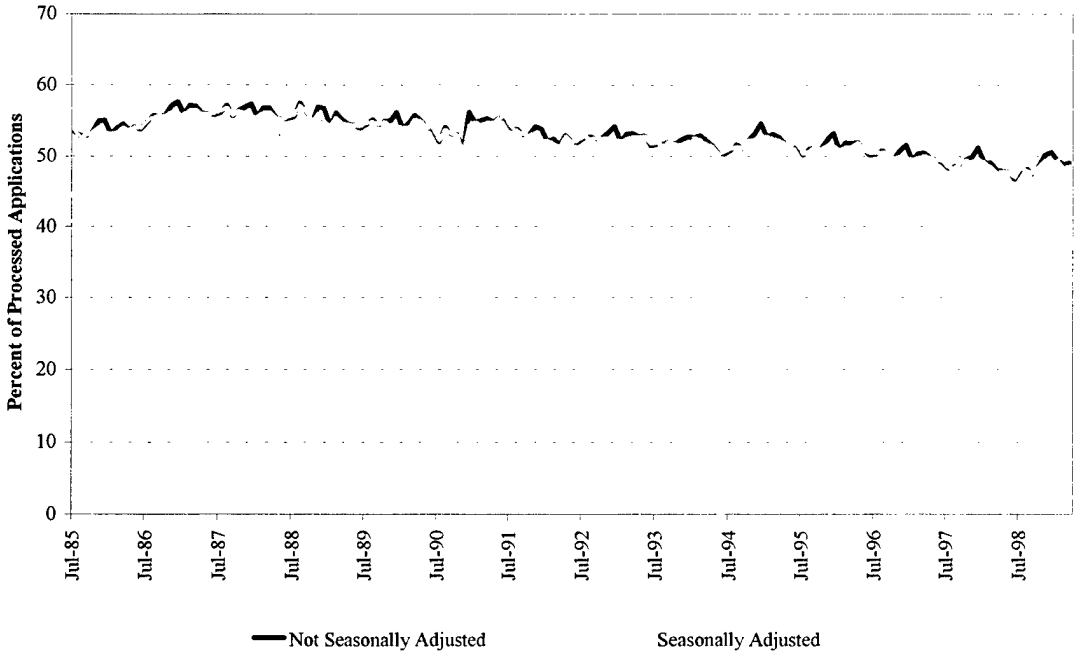


Figure 2.5--AFDC/TANF Approved Applications - California

Region and Urbanization

Thus far, we have presented results at the state level. The ability to examine disaggregated results will also be very important to analyzing the California caseload. Because counties gained considerable discretion in designing welfare programs and because the counties are very heterogeneous, a particular reform could have differential impacts across counties. The administrative data allow us to examine caseload results disaggregated by county. In this section, we present caseload results in which counties are grouped by geographic region (Northern, Central, Coastal, and Southern) and level of urbanization (Urban, Mixed, Rural, Los Angeles County) to demonstrate the usefulness of disaggregated results. The assignment of counties to groups is given in Appendix A.

Table 2.3

Average Monthly TANF Caseload by Region

Region	Level 1Q 1999	% Change from 4Q 1998	% Change from 1Q 1998
Northern Region			
Average Monthly Caseload	27,324	-0.29	-12.0
FG	22,305	-1.17	-12.2
UP	5,019	3.82	-10.9
Total Recipients	79,032	-0.76	-13.7
Percent of Population Receiving Aid	6.55	-0.76	-13.7
Total Expenditures	$12,900,198	1.04	-10.5
Expenditures per Case	$472	1.33	1.7
Central Region			
Average Monthly Caseload	144,917	-0.77	-11.4
FG	115,224	-1.82	-11.6
UP	29,694	3.54	-10.6
Total Recipients	450,884	-0.91	-12.4
Percent of Population Receiving Aid	6.93	-0.91	-12.4
Total Expenditures	$71,442,252	1.54	-8.2
Expenditures per Case	$493	2.33	3.6
Southern Region			
Average Monthly Caseload	379,323	-2.72	-11.7
FG	317,866	-2.86	-11.6
UP	61,457	-1.97	-12.2
Total Recipients	1,073,249	-3.12	-13.7
Percent of Population Receiving Aid	4.19	-3.12	-13.7
Total Expenditures	$186,620,720	-1.31	-9.9
Expenditures per Case	$492	1.45	2.0
Coastal Region			
Average Monthly Caseload	81,232	-2.78	-15.1
FG	69,666	-3.32	-15.1
UP	11,566	0.64	-15.2
Total Recipients	222,646	-3.23	-17.6
Percent of Population Receiving Aid	2.73	-3.23	-17.6
Total Expenditures	$38,931.510	-1.99	-14.7
Expenditures per Case	$479	0.81	0.5

We present quarterly average caseload information for the four geographic regions in Table 2.3. The Southern region has a significantly larger caseload (379,323) than the other regions because in part it contains the populous counties of Los Angeles and San Diego (the Northern region has 27,324, the Central region has 144,917, and the Coastal region has 81,232). The Northern region's caseload is

approximately one-fourteenth the size of the Southern region's caseload. However, from examining the percent of the population receiving aid, the Northern (6.55) and Central (6.93) regions have significantly higher caseloads as compared to the Coastal (2.73) and Southern (4.19) regions. Expenditure per case varies little across the four regions.

Examining changes from the previous year, the Coastal region experienced the largest decline in total caseload—15.1 percent for Coastal, 12.0 percent for Northern, 11.4 percent for Central, and 11.7 percent for Southern. This finding is likely due to the coastal economy recovering more quickly than the rest of the state following the 1990 recession.

Variation among the regions with respect to application information is shown in Table 2.4. The Northern and Central regions had more applications per thousand individuals than the Southern and Coastal regions (2.6 and 2.0 versus 1.1 and 0.9, respectively). The Central and Coastal regions had the highest approval rate of applications (35 percent) and the Northern region had the lowest denial rate (13 percent).

Table 2.4

Average Monthly TANF Applications by Region

Region	Level 1Q 1999	% Change from 4Q 1998	% Change from 1Q 1998
Northern Region			
Total Applicant Pool	3,179	6.46	-27.4
Pending Applications	1,385	11.37	-40.2
New Applications	1,794	2.96	-13.1
Percent of Population with Pending Application	0.26	6.46	-27.4
Action on Applications			
Approved	32%	-4.98	3.9
Denied	13%	-1.20	0.8
Other Action	7%	-0.01	1.7
Left Pending	48%	6.19	-6.4
Central Region			
Total Applicant Pool	13,275	-1.95	-8.3
Pending Applications	5,132	0.83	-15.7
New Applications	8,143	-3.62	-2.9
Percent of Population with Pending Application	0.20	-1.95	-8.3
Action on Applications			
Approved	35%	-0.21	-0.1
Denied	19%	-0.22	0.6
Other Action	9%	0.48	1.4
Left Pending	38%	-0.05	-1.9
Southern Region			
Total Applicant Pool	28,936	-5.92	-6.9
Pending Applications	9,050	-8.69	-5.0
New Applications	19,886	-4.61	-7.7
Percent of Population with Pending Application	0.11	-5.92	-6.9
Action on Applications			
Approved	31%	-0.77	0.4
Denied	20%	-0.49	-1.1
Other Action	19%	1.05	0.4
Left Pending	30%	0.21	0.4
Coastal Region			
Total Applicant Pool	7,125	0.05	-6.2
Pending Applications	2,3230	2.53	-10.5
New Applications	4,805	-1.11	-3.9
Percent of Population with Pending Application	0.09	0.05	-6.2
Action on Applications			
Approved	35%	-2.22	1.6
Denied	24%	-1.31	-2.2
Other Action	9%	-0.60	-0.4
Left Pending	33%	4.13	1.0

In Figures 2.6 through 2.10, we show the time trends of caseload characteristics for the four regions. Figure 2.6 presents the total caseload by region. The Northern region had a much more gradual increase in caseload to its peak in 1995 then the increase in the Southern region; all the regions experienced similar declines in the caseload after 1995, with the exception of the Coastal region, which declined moderately faster.

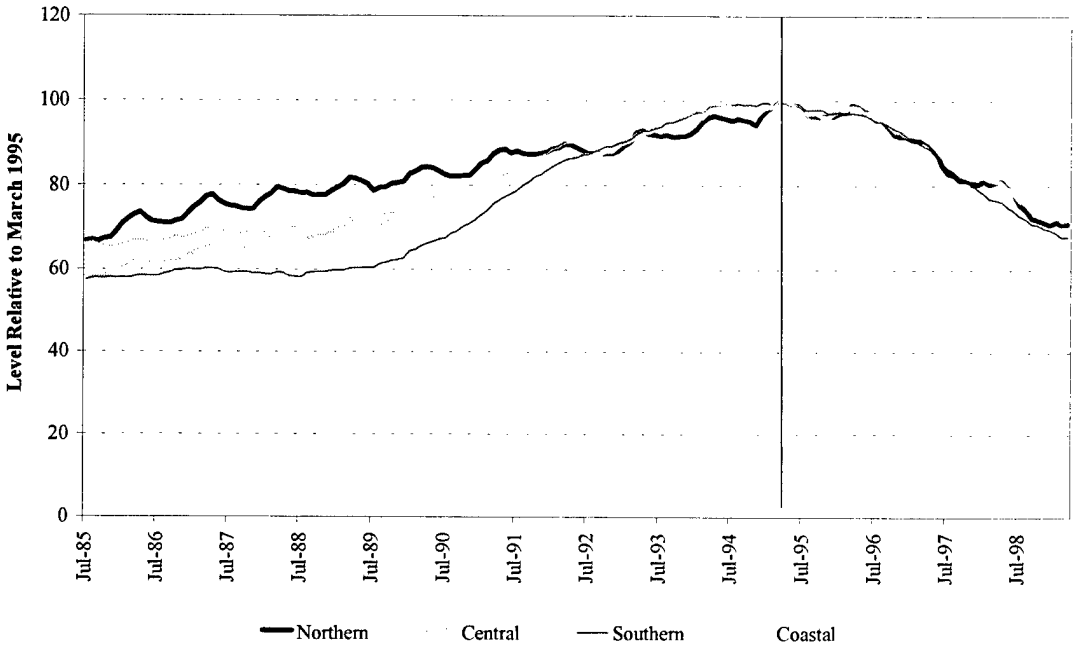


Figure 2.6--AFDC/TANF Total Caseload by Region

Similar trends are apparent in the total expenditure time series presented in Figure 2.7. The time trends of expenditures per case are remarkably similar across the four regions (see Figure 2.8).

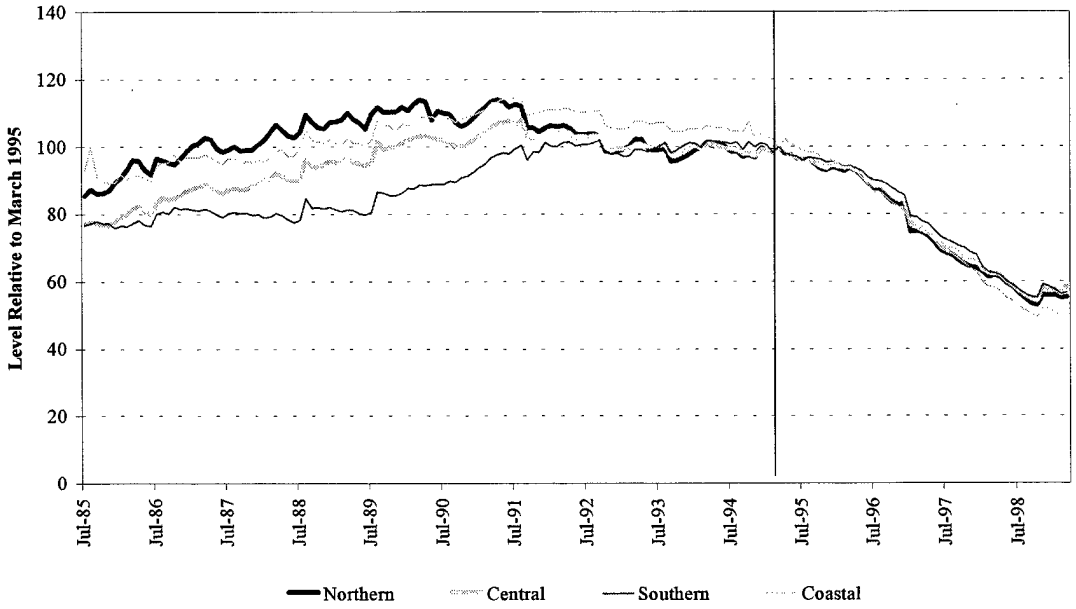


Figure 2.7--AFDC/TANF Total Expenditure by Region

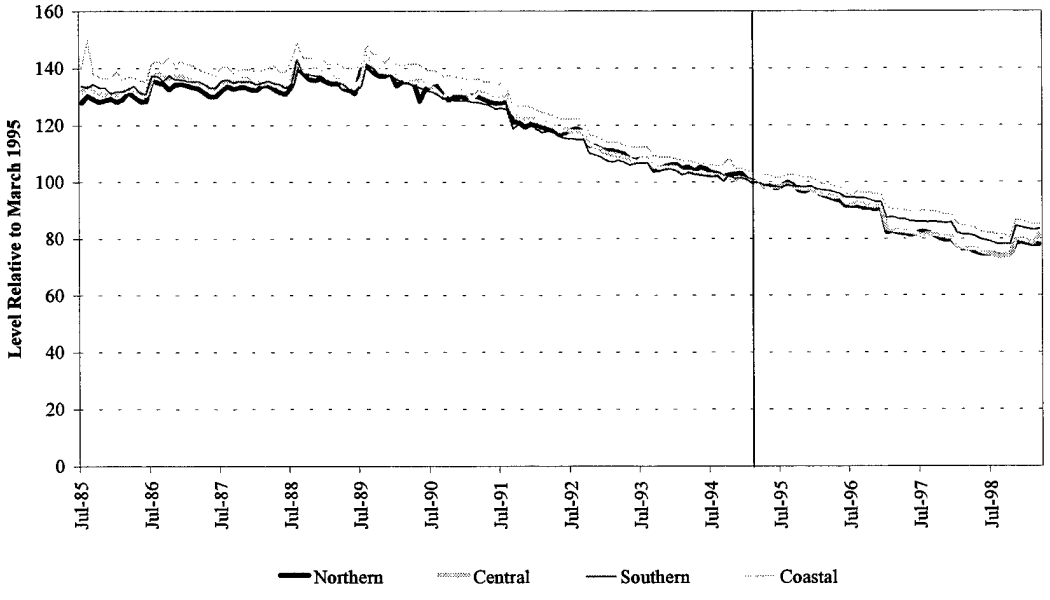


Figure 2.8--AFDC/TANF Expenditure per Case by Region

In Figure 2.9, we present the time trend of new applications. The regions appear to move separately in the years before 1995 but show little variation after 1995.

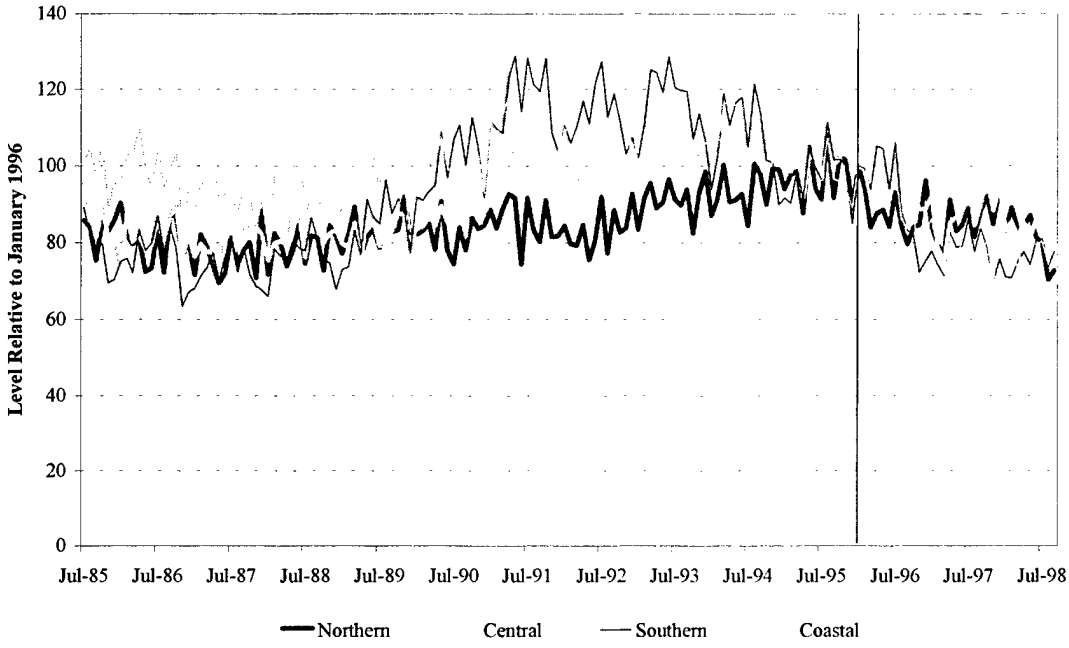


Figure 2.9--AFDC/TANF Total New Applications by Region

Finally, we present approval rates in Figure 2.10. The approval rate for each region declines similarly across the four regions, but there is a distinct level difference; in particular, the Northern and Central regions consistently have higher approval rates compared to the Southern and Coastal regions.

Results by urbanization vary much less than they do by region, so we do not present them here.

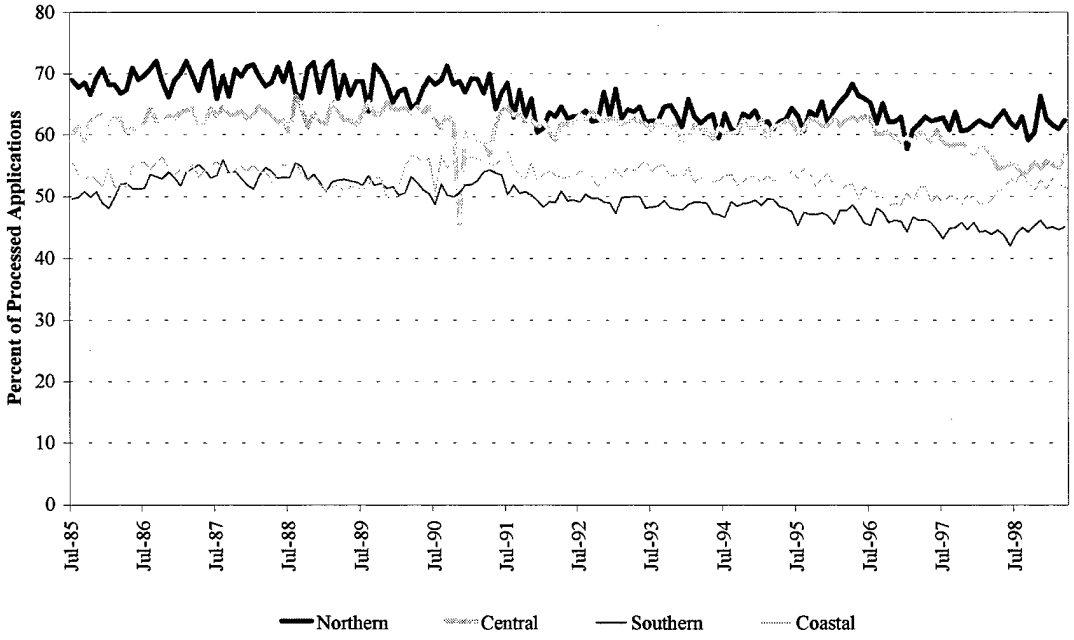


Figure 2.10--AFDC/TANF Approved Applications by Region

3. GAIN25: WELFARE-TO-WORK ADMINISTRATIVE DATA

The GAIN25 Statistical Report is completed by every county in every month to report participation information for California's WTW program, formerly known as GAIN.[10] California's GAIN program preceded the federal Job Opportunities and Basic Skills (JOBS) program of 1992 that mandated work requirements for certain welfare recipients. After the introduction of JOBS, GAIN was adjusted to accord with the federal program, and the GAIN25 form was used to satisfy federal reporting requirements. CalWORKs has replaced GAIN, but the GAIN25 form continued to be used at the start of CalWORKs to report program participation to the CDSS.

The GAIN25 form has been used since the late 1980s and was replaced mid-summer 1999. GAIN25 collects information on the participation of welfare recipients in the WTW program, including the following:

- Overall program participation;
- Educational activity and job search participation;
- No-show and sanction levels.

GAIN25 provides important information about the character of previous welfare-to-work programs in the counties. For example, it will help us determine which counties had well-developed WTW programs prior to CalWORKs and which counties previously had high sanction rates. Answers to these questions will provide information about which counties should be expected to change their policies the most with the introduction of CalWORKs.

DATA ISSUES

We only present very preliminary results from the GAIN25 in this report because we are still in the process of evaluating the quality of the data. Various county and state officials have questioned the

[10] Further information about the history and characteristics of WTW programs in California is available in Zellman et al.

accuracy of GAIN25 information, particularly after the implementation of CalWORKs in January 1998. Even if the information is only accurate for the period before January 1998, the data will still be quite valuable. GAIN25 represents some of the only information available on counties' implementation of WTW programs before CalWORKs. For counties that had implemented substantial WTW programs before January 1998, the WTW activities of CalWORKs could represent less of a change as compared to counties who did not have substantial WTW programs.

We analyze data for the months July 1997 through January 1998. We are exploring the possibility of obtaining additional historical data.

TABULATIONS

We present initial results from the GAIN25 data in Table 3.1.[11] There were 265,282 total registrants per month during the quarter, with close to 95 percent of the registrants being mandatory. Of these registrants, 4,690 were sanctioned, implying a sanctioning rate of under 2 percent.

Table 3.1

GAIN Data: Average Monthly Registrants for California

	Level 4Q 1997	% Change from 3Q 1997	% Change from 4Q 1996
Total Registrants	265,282	-0.03	
Mandatory FG	173,771	0.84	
Mandatory UP	76,403	1.87	
Voluntary	15,108	-16.3	
Percent of Population Registered	0.21	-0.03	
Total Sanctions	4,690	-2.65	
Sanctions per Registrant	0.0177	-2.62	
Registrants per TANF case	0.36	3.52	

Note: 4Q 1996 data are not available at this time.

In Table 3.2, we present similar information for each region. The Southern Region has the largest number of participants in its GAIN program (135,974 per month) and the smallest number of participants in the Northern Region (16,280 per month). However, as is clear by now,

[11] It should be noted that this data is for the fourth quarter of 1997, whereas previous data were for the first quarter 1999.

the regions are of widely disparate size. Comparing registrants per caseload, the Northern Region has the highest participation rate (0.52 participants per TANF case) and the Southern Region has the smallest (0.30 participants per TANF case). Examining sanction rates across the regions, the Southern Region sanctions the most (2.0 percent) and the Northern Region sanctions the least (1.5 percent).

We present a similar series of results by urbanization in Table 3.3. There is less variation in registrants per population across levels of urbanization overall. However, the sanction rate for Los Angeles is dramatically different from the other three designations.[12] Los Angeles County sanctions over 3 percent of its participants, a value almost twice as large as the next highest level of urbanization (the rural group with 1.6 percent).

[12] This finding is consistent with other researchers. See Manpower Demonstration Research Corp.

Table 3.2

GAIN Data: Average Monthly Registrants by Region

	Level 4Q 1997	% Change from 3Q 1997	% Change from 4Q 1996
Northern Region			
Total Registrants	16,280	-4.59	
Mandatory FG	9,838	-3.24	
Mandatory U	5,157	-8.06	
Voluntary	1,285	-0.1	
Percent of Population Registered	0.44	-4.59	
Total Sanctions	241	8.23	
Sanctions per Registrant	0.0148	13.44	
Recipients per TANF case	0.52	-2.93	
Central Region			
Total Registrants	78,973	2.08	
Mandatory FG	48,684	4.06	
Mandatory U	26,516	-0.54	
Voluntary	3,773	-3.73	
Percent of Population Registered	0.40	2.08	
Total Sanctions	1,280	6.58	
Sanctions per Registrant	0.0162	4.41	
Recipients per TANF case	0.48	5.58	
Southern Region			
Total Registrants	135,974	-1.95	
Mandatory FG	91,924	-1.61	
Mandatory U	37,073	4.94	
Voluntary	6,977	-29.7	
Percent of Population Registered	0.17	-1.95	
Total Sanctions	2,654	-9.34	
Sanctions per Registrant	0.0195	-7.53	
Recipients per TANF case	0.30	1.41	
Coastal Region			
Total Registrants	34,055	5.55	
Mandatory FG	23,325	6.3	
Mandatory U	7,657	3.41	
Voluntary	3,073	5.31	
Percent of Population Registered	0.14	5.55	
Total Sanctions	515	10.36	
Sanctions per Registrant	0.0151	4.56	
Recipients per TANF case	0.34	10.74	

Table 3.3

GAIN Data: Average Monthly Registrants by Urbanization

	Level 4Q 1997	% Change from 3Q 1997	% Change from 4Q 1996
Rural			
Total Registrants	61,492	2.56	
Mandatory FG	39,352	2.79	
Mandatory U	19,526	2.29	
Voluntary	2,614	1.15	
Percent of Population Registered	0.36	2.56	
Total Sanctions	992	-1.26	
Sanctions per Registrant	0.0161	-3.73	
GAIN recipients per TANF case	0.48	5.11	
Mixed Urban/Rural			
Total Registrants	78,144	-0.78	
Mandatory FG	51,789	1.55	
Mandatory U	22,762	-5.87	
Voluntary	3,593	0.33	
Percent of Population Registered	0.29	-0.78	
Total Sanctions	1,022	12.51	
Sanctions per Registrant	0.0131	13.4	
GAIN recipients per TANF case	0.51	2.95	
Urban			
Total Registrants	71,333	-0.55	
Mandatory FG	44,307	4.07	
Mandatory U	19,788	0.96	
Voluntary	7,238	-24.2	
Percent of Population Registered	0.16	-0.55	
Total Sanctions	971	-1.05	
Sanctions per Registrant	0.0136	-0.51	
Recipients per TANF case	0.36	4.15	
Los Angeles			
Total Registrants	54,313	-1.13	
Mandatory FG	38,323	-5.31	
Mandatory U	14,327	18.11	
Voluntary	1,663	-28.6	
Percent of Population Registered	0.14	-1.13	
Total Sanctions	1,705	-11.4	
Sanctions per Registrant	0.0314	-10.3	
Recipients per TANF case	0.21	1.93	

4. MEDS: MEDICAL ELIGIBILITY DETERMINATION SYSTEM DATA

Although the CA237 form provides important information about aggregate AFDC/TANF caseload movements, it lacks information about the characteristics of individuals on AFDC/TANF. For example, it is not possible to determine the racial distribution or the average family size of the welfare population. However, another administrative data source, the MediCal Eligibility Determination System (MEDS), can be used to formulate caseload characteristic information. The MEDS is a statewide roster of all individuals who are receiving MediCal and is used to verify eligibility of health services by service providers. Because AFDC/TANF individuals automatically qualify for MediCal and the roster indicates whether individuals qualify for MediCal because of receiving AFDC/TANF, caseload characteristic information can be extracted from the MEDS.[13]

The MEDS data contain the following information for each person who is receiving MediCal benefits: county of residence, reason for MediCal qualification for each month (including AFDC-FG and AFDC-UP), date of birth, race/ethnicity, and primary language spoken (after 1990). Thus, MEDS is an individual-level, statewide data set on welfare recipients and includes such demographic information as age, race/ethnicity, and language ability.

The MEDS database represents an unparalleled source of caseload information, both because of its sample length and size. With these data, we will be able to examine important questions about the characteristics of the welfare caseload. For example, we can determine whether the significant decline in the welfare caseload is associated with the short-duration cases leaving, whether there are important racial/ethnic differences in the caseload, and/or whether the distribution of family size has changed. Answers to these questions

[13] It should be noted that the vast majority of individuals on AFDC/TANF qualify for MediCal coverage. However, it is possible that individuals who qualify for MediCal and TANF are only enrolled for TANF.

will be important to assess the impact of specific program changes and for forecasting future changes in the welfare caseload.

DATA ISSUES

There are three key data issues in analyzing the MEDS. The first issue arises because of the number of records in the MEDS. The MEDS is an individual-level data set that contains one record for every person who qualifies for MediCal. For December 1998, this amounted to over 6.1 million records. Even after extracting the AFDC recipients, we have a data set that contains over 2.7 million records. Moreover, we analyze data not just for December 1998, but for the period January 1987 through December 1998.

We rely on two strategies to make the processing feasible. First, we collapse the MEDS into a summary data set of caseload counts. For example, we calculate the number of cases in a county in a month that has a particular combination of demographic characteristics. To collapse the data, we use the following categories for every county and every month:

- Race groups: Latino, black, white, other;
- Family size: 1 child, 2 children, 3+ children;
- Type of aid: FG, UP
- Language: English, Other (after 1990);
- Age of Oldest Adult: 0-18, 19-28, 29-38, 39+;
- Age of Youngest Child: 0-3, 4-6, 7-19.

Thus, the summary data set stores the total number of cases each month in each county that, for example, had the following characteristics: the head of the case was Latino, there was a child, the case qualified under FG, the head spoke English, the oldest adult was 0-18, and the youngest child was 0-3. We repeat this type of tabulation for every possible combination of the categories listed above. We then rely on the summary data file to examine trends in the caseload for specific groups.

To keep the summary data file to a manageable size, the categories for the data must remain relatively broad. For example, even with the relatively broad categories chosen above, we were left with a summary file of 44,544 observations per month (multiplying 4 race groups, by 3 family size groups, ..., by 58 counties).

Our second strategy to process the MEDS is to draw a random sample from the underlying data set and then analyze the individual-level data directly. For this report, we draw a 1 percent random sample.[14] Although this strategy ignores much of the information in the underlying database, it allows for much more flexibility in analyzing the data. In future analysis, it may be necessary to draw a larger random sample to present duration results for subgroups of the California population.

A second issue we face in using the MEDS data is that there is a processing lag in designating MediCal claimants as AFDC/TANF recipients.[15] This lag has two effects on the data sets we analyze. Results for the most recent months will be subject to updating as additional claims are entered into the MEDS. In addition, given the construction of our data set, the lag causes a false periodicity in the data. In particular, the data sets were extracted from the underlying MEDS database in the same months in most years (June or December). Months that are closer to the extraction month will have lower caseloads because the database has yet to be updated with the new cases. Because of the regularity in the processing lag, we are able to construct a data set that minimizes the periodicity and statistically adjusts the data set to account for the remaining periodicity. For the summary file analysis, we use a weighting scheme to statistically adjust for the

---

[14] In future work, we will use larger random samples. For example, we are currently constructing a stratified (by county) random sample. In this scheme, the county sampling probabilities are chosen so that approximately equal samples are chosen from the large counties and all cases are chosen from the small counties.

[15] The processing lag occurs because MediCal eligibility can be determined more easily than AFDC/TANF eligibility. New claimants are initially designated as "MediCal only" recipients in MEDS while the AFDC/TANF application is being processed. Then, if the AFDC/TANF application is successful, MEDS must be updated to reflect the new eligibility classification. Updating the MEDS record can be delayed a few months after the actual AFDC/TANF approval.

remaining periodicity. Specific details on the weighting scheme are provided in Appendix A. All population tabulations from the summary file use this adjustment. For the analysis of the random sample, we account for the processing lag in the models directly; further details are provided below.[16]

A third issue we face is that because of the structure of the data, only one case serial number per person per data extraction exists.[17] To obtain a person-month data set, we assign the available case serial number to an individual for every month in the extraction. Thus, individuals can only switch cases (e.g., splitting off to a new case) between extractions.

The MEDS data has several drawbacks for examining the welfare caseload. First, it contains the population of MediCal recipients, not AFDC recipients. Thus, an individual who receives AFDC but not MediCal will not be in the MEDS database. Other research and results we present below suggest the size of the population that receives AFDC but not MediCal is small, but the possibility does exist.[18] Furthermore, the MEDS database only follows individuals in California. We cannot distinguish between individuals leaving California and individuals exiting the welfare rolls.

We currently have MEDS data for January 1987 through December 1998. We expect to receive regular updates of the MEDS database, receiving the updates with a two-month lag.

TABULATIONS

In this section, we look first at the summary file and then at the random sample file.

---

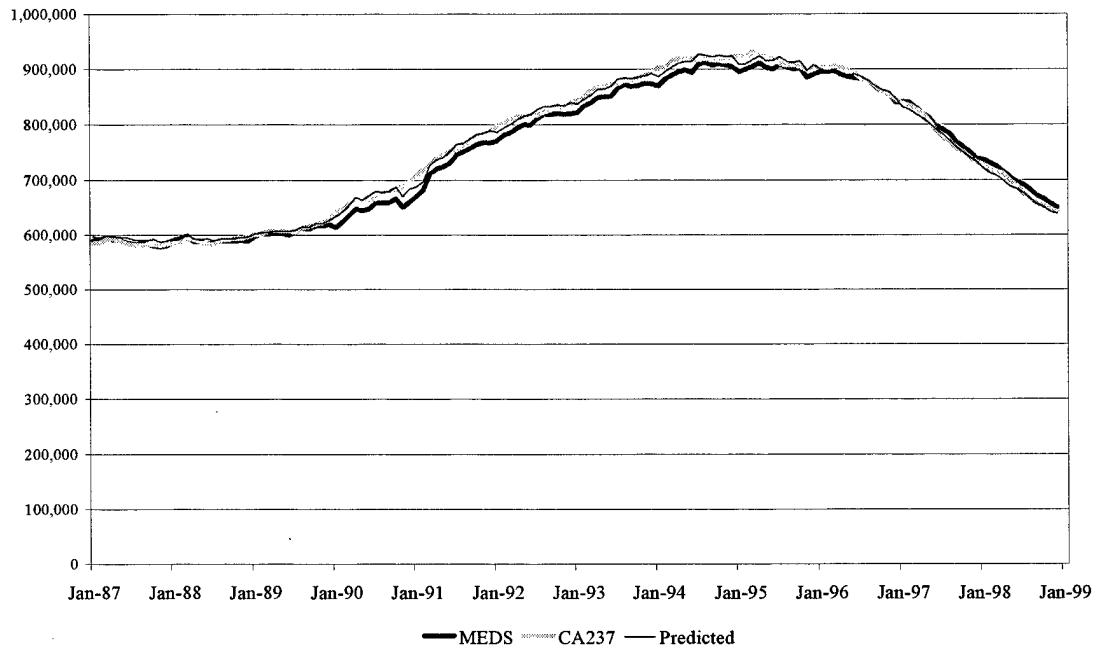
[16] An additional dip appears in December due to processing around the winter holidays. Our weighting procedure will adjust for this dip also.

[17] Again, we use some six-month and some twelve-month extractions. See the discussion on weighting in Appendix A for a complete description.

[18] Hoynes (1997) reports that 97 percent of AFDC recipients participate in Medicaid nationwide.

Summary File

As a check on the quality of the MEDS data, we first compare the total caseload calculated from the MEDS, both unweighted and weighted, to the CA237 and present the results in Figure 4.1. First, it is clear from the figure that both the MEDS series match the CA237 very well. Specifically, the MEDS replicates both the caseload level and trend in the CA237. Second, the unweighted MEDS series has the seasonal pattern expected because of the processing lag; in particular, there is a spike toward the end of most years. Finally, we see that the weighted MEDS, where we use the weighting procedure described in the appendix, removes some but not all of the spike. All tables and figures in this subsection rely on the weighted MEDS.



Note: The line marked "MEDS" refers to tabulations from the MEDS that are not weighted. The line marked "predicted" refers to tabulations from the MEDS that are weighted as explained in Appendix A. The line marked CA237 refers to the official caseload tabulations from the CA237.

Figure 4.1--Actual v. Predicted Caseload Totals

We present tabulations from the MEDS summary file in Table 4.1. This table presents the demographic distribution of the caseload for 1998. From the first part of the table, we see that the largest segment

of the welfare caseload is Latino (42 percent) followed by whites (28 percent). In the second part of the table, we see that 31 percent of the caseload has the oldest adult on the case being less than 18 years old; this percentage does not include foster care cases. From the last two parts of the table, 37 percent of the cases have the youngest child between the ages of 7 and 19 and 40 percent of the cases have only one child on the case.

Table 4.1

1998 MEDS Caseload for California

Demographic Information	Percent
Race/Ethnicity	
White (non-Latino)	27.6
Black (non-Latino)	20.2
Latino/a	42.0
Other Race (non-Latino)	10.2
Age of Oldest Individual on Case	
0-18 years	30.5
19-28 years	23.3
29-38 years	25.8
39 years and up	20.4
Age of Youngest on Case	
0-3 years	41.7
4-6 years	21.5
7-19 years	36.8
Family Size	
1 Child	39.5
2 Children	31.2
3 or more Children	29.3

We present results for how the distribution by race and family size changed over time in Figures 4.2 and 4.3. First, examining the race figure for the state of California, it is clear that the Latino welfare population increases significantly more quickly than the other race/ethnic populations. Second, although the caseload for different family sizes moved together from 1987 to 1996, the changes after 1996 are ranked by family size. Specifically, the caseload of families with one child declined most quickly, followed by the caseload of families with two children; caseloads with three children declined most slowly.

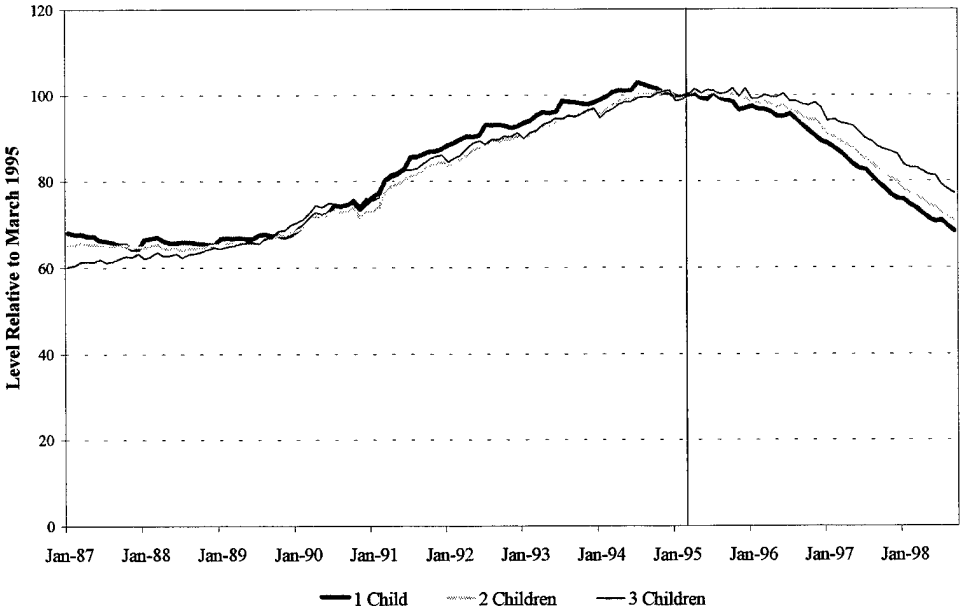


Figure 4.2--AFDC Total Recipients by Number of Children in Household

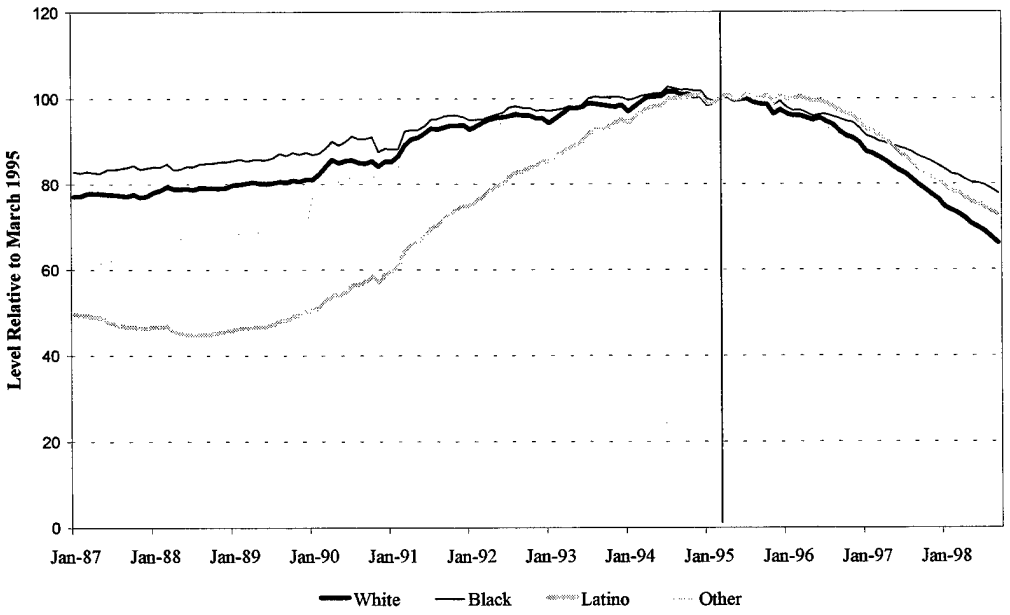


Figure 4.3--AFDC Total Recipients by Race

Random Sample File

With the random sample file, we examine the duration of time that individuals stay on welfare by following individuals over time.[19] To examine spell lengths, we ignore all single month spells and all single month exits off welfare. We do this because one month changes usually occur for administrative reasons. For example, a single month exit could occur because of individuals not renewing their MediCal coverage in a timely manner rather than because of individuals not qualifying for AFDC/TANF for one month.[20] In addition, we ignore all spells that are in progress in the first month of our data because we do not know how long the spell has lasted. Such spells are often referred to as "left censored" spells in the analytic literature.

In Table 4.2, we present the "population at risk" and the exit probabilities (also known has hazard rates) for different spell lengths. The population at risk is the number of spells that last for at least a given duration and that are observed for the whole duration. For example, the numbers in the table indicate that 41,362 spells were at risk to end during the first year, and 17,889 spells were at risk to end during the second year. Exit probabilities are the proportion of the population at risk that ended in a given period. For example, the table suggests that 54 percent of the population at risk left welfare in the first period and 39 percent left welfare in the second period.

It should be noted that the size of the population at risk declines more rapidly than the exit probabilities would suggest. This is because some spells are no longer at risk due to being "right-censored" rather than the spell ending.[21] Looking down the columns, it is apparent that

---

[19] No corrections are made for the processing lag for the random sample file analysis at this time.

[20] Hoynes (1997), using the same data, makes the same assumption because of a suggestion by state officials.

[21] If no spells were "right-censored" (explained below), then the population at risk in a period would simply be the product of the exit probability and the population at risk in the previous period. However, consider a spell that starts July 1997 and that is still in progress in December 1998. Such a spell is considered to be right-censored (in the 18th month) because we do not observe the spell ending before the end of the sample period. This spell is at risk for ending during the first year because it is possible to observe the spell throughout the first

the exit probabilities decline as the duration increases. Such a pattern is referred to as "negative duration dependence" and is consistent with the findings of other researchers. The exit probabilities are larger than those calculated by previous authors.[22] One reason may be that because we use administrative data, we capture every exit from the caseload, even if it is for only a couple months.

Table 4.2

AFDC/TANF Exit Probabilities

Spell Length in Years	Population at Risk	Exit Probability	Survivor Probability
1	41,362	0.54	0.46
2	17,889	0.39	0.28
3	10,043	0.33	0.19
4	6,186	0.29	0.13
5	3,944	0.28	0.10
6	2,475	0.27	0.07
7	1,507	0.26	0.05
8	899	0.26	0.04

We present survivor probabilities in the final column. These numbers represent the percentage of spells that "survive" (i.e., do not end) through a given period. For example, 0.54 of the spells survive through the first year and 0.39 of the spells survive through the first two years. Notably, only 0.10 percent of the spells survive for over five years.

The calculations in Table 4.2 are useful to look at average stays on welfare, but they ignore potential changes in the rates over time (i.e., calendar month/year). As an initial attempt to look at this issue, we present in Table 4.3 cumulative exit probabilities for individuals who start begin aid in each two-year interval between 1987 and 1997. For each time period, we calculate the probability that a spell lasts less than or equal to 6, 12, 18, and 24 months. The pattern for each length of time is remarkably similar, so we will focus on

---

year. However, the spell is not at risk for the second year because the spell could have ended in the 19th through the 24th month but we do not observe it.

[22] Bane and Ellwood (1994)present a similar table on page 32.

spells that last for 6 months or less. Looking down that column, the exit rate is 0.34 for 1987/88, declines to 0.29 for 1993/94, and then increases to 0.38 in 1997. Thus, the increase in the exit probabilities associated with the recent decline in the welfare caseload seems to be a return to the exit probabilities experienced before the early 1990's recession. One interpretation of this finding is that the change in exit probabilities results from business cycle changes.

Table 4.3

AFDC/TANF Predicted Spell Durations

Spell Starting	# of Spells	Probability That a Completed Spell Lasts ≤ 6 months	≤ 12 months	≤ 18 months	≤ 24 months
2/87 - 12/88	6,550	0.34	0.55	0.65	0.72
1/89 - 12/90	7,035	0.35	0.56	0.66	0.72
1/91 - 12/92	8,209	0.33	0.53	0.63	0.69
1/93 - 12/94	8,477	0.29	0.49	0.62	0.70
1/95 - 12/96	7,636	0.33	0.55	0.67	0.76
1/97 - 12/97	3,126	0.38	0.61	X	X

5. CONCLUSIONS

The goal of this report is to provide detailed information on the various administrative data sets that we will use to analyze changes in the welfare caseload in California. In addition, we provide initial tabulations from these data to describe the general changes in the caseload.

Overall, the administrative data sets provide a rich picture of the changes in the welfare caseload. Two data sets, the CA237 and GAIN25, provide county-level information on the caseload, as reported directly to the CDSS. A third data set, the MEDS, provides a comprehensive, individual-level information as maintained for MediCal purposes. Where comparisons are possible between data sets (MEDS and CA237) and with published data sources, these data sources match up very well.

Our initial tabulations in this report suggest the following:

- After a substantial increase in the California welfare caseload between 1988 and 1994, the caseload remained fairly constant until 1996 and then began to decline dramatically.
- The majority of the change in California caseload has come from changes in the number of applicants rather than from the approval/denial rate of applications.
- Although the trends for geographic regions moved disparately during the late 1980s, regional trends have been quite similar during the 1990s.
- The proportion of cases that were of short duration declined during the early 1990's and returned to previous levels.

The tabular evidence presented in this report will be extremely important to understanding the basic trends that need to be explained and is suggestive of the mechanisms that could be important in explaining changes in the welfare caseload. However, it is also clear that more sophisticated analyses will be needed to evaluate the relative importance of various programmatic changes.

APPENDIX A. TECHNICAL DETAILS

In this appendix, we provide further details of our analyses procedures.

CONTROLLING FOR SEASONALITY

There may be a seasonal component for many of the time series we report, particularly for regions that depend on agriculture. For example, applications for AFDC tend to increase during the winter months. We graph adjusted data for the predictable seasonal variation in some of the figures so that long-term trends are more readily discernible.

To adjust an arbitrary monthly time series x_t for predictable seasonal variation, we first run the regression,

$$\log x_t = \alpha' M + \varepsilon_t ,$$

where M is a vector of monthly dummies. We then calculate the seasonally adjusted time series as:

$$\hat{x}_t = \exp[\alpha_1 + e_t], \qquad \text{(A1)}$$

where α_1 is the estimated coefficient for January and e_t is the residual for month t. This procedure controls for predictable percentage changes in the monthly caseload and then adjusts all data to "look like" January.

SCALING THE MEDS SUMMARY DATA SET

As explained in Section 2, one drawback to the MEDS is that there is a processing lag in the underlying database. Consequently, there is a systematic pattern in which the number of new cases is lower when the reporting month is closer to the extraction month. To correct for this lag, we calculate a set of scale factors to apply to demographic-specific counts. To calculate these scale factors, we use the county-level caseload levels from the CA237 as the "true" monthly caseload report.

We currently have 12-month extractions for every June and December for the years 1987 through 1998. First, consider the case where we only had data from December extractions. The processing lag affects a given

month similarly from year to year, but the months closer to December will be affected more and December will be affected the most. Therefore, we calculate scale factors that are allowed to vary from month to month. In our implementation, we also allow the processing lag to vary systematically across years and counties.

Formally, let y^*_{tk} be the "true" monthly caseload level reported in the CA237 data for the tth time period and the kth county and let y_{tk} be the MEDS monthly caseload for tth time period and the kth county. In addition, let $\boldsymbol{M}$ be a vector of monthly dummies (January through December), let $\boldsymbol{Y}$ be a vector of year dummies (1987 to 1997, excluding 1998), and $\boldsymbol{C}$ be a vector of county dummies (57 counties, excluding Alameda). We estimate the model,

$$y^*_{tk} = y_{tk}(\alpha' M)(1+\beta' Y)(1+\delta' C)+\varepsilon_{tk} \quad , \qquad \text{(A2)}$$

where $\boldsymbol{\varepsilon_{tk}}$ is the disturbance term. The parameters $\boldsymbol{\alpha}$ are the monthly scale factors for Alameda County in 1998, $\boldsymbol{\beta}$ are year-specific, multiplicative adjustments (relative to 1998), and $\boldsymbol{\delta}$ are county-specific, multiplicative adjustments (relative to Alameda County) to adjust the MEDS total caseloads to the true county totals. The parameters are estimated with non-linear least squares. These parameters are then used to scale cell specific estimates of the caseload (e.g., age*race*language).

Because we also have June extractions for many of the years, we modify equation (A2) slightly. First, we change the notation from using monthly dummies that correspond to a calendar month ($\boldsymbol{M}$) to monthly dummies that correspond to a data extraction month ($\boldsymbol{R}$),

$$y^*_{tk} = y_{tk}(\alpha' R)(1+\beta' Y)(1+\delta' C)+\varepsilon_{tk} \; . \qquad \text{(A3)}$$

Thus, the first dummy variable will equal one if it is the first month for a given extraction. Second, we now have two sets of data for certain months. For example, we have a June extraction (June 1998) and a December extraction (December 1997) for November 1997. However, the June 1998 extraction for November 1997 is substantially more complete than the December 1997 extraction because more of the new applications have been processed. We construct a data set that has the "best" report for a given month, where "best" is defined as the extraction that is furthest from the given month. We then calculate scale factors by

creating dummy variables consistent with (A3). However, we note that December always appears to be systematically less than the CA237, regardless of whether it is a June or December extraction. Presumably, the discrepancy is due to a processing lag for the holiday season.

SUPPLEMENTARY DATA SOURCES

We use a few supplementary data sources in this report. We rely on population information from the Intercensal Estimates of population, provided by the U.S. Bureau of the Census.

We use information on the population per square mile from the 1994 County and City Data Book to classify counties as urban, mixed, or rural, as shown in Table A.1. Because of it size, we classify Los Angeles County as its own category.

Table A.1

Classification of Counties into Urban, Mixed, and Rural

Designation	Counties
Urban	San Francisco, Orange, Alameda, San Mateo, Santa Clara, Contra Costa, Sacramento, San Diego
Mixed	Santa Cruz, Marin, Solano, Ventura, San Joaquin, Stanislaus, Sonoma, Riverside, Napa, Yolo, Santa Barbara, Placer, Fresno, Butte, Sutter, Monterey, Merced, Yuba
Rural	Nevada, El Dorado, Kings, San Bernardino, Kern, Tulare, San Luis Obispo, Amador, Madera, Lake, Shasta, Calaveras, Humboldt, Imperial, Del Norte, San Benito, Mendocino, Tuolumne, Glenn, Tehama, Colusa, Mariposa, Plumas, Siskiyou, Lassen, Trinity, Mono, Sierra, Alpine, Inyo, Modoc
Los Angeles	Los Angeles

SOURCE: Based on population per square mile in the 1994 County and City Data Book.

Finally, we provide analysis with respect to four geographic regions, as shown in Table A.2.

Table A.2

Classification of Counties into Regions

Region	Counties
North	Butte, Colusa, Del Norte, Glenn, Humboldt, Lake, Lassen, Mendocino, Modoc, Nevada, Plumas, Shasta, Sierra Siskiyou, Sutter, Tehama, Trinity, Yuba
Central	Alpine, Amador, Calaveras, El Dorado, Fresno, Inyo, Kern, Kings, Madera, Mariposa, Merced, Mono, Placer, Sacramento, San Joaquin, San Luis Obispo, Stanislaus, Tulare, Tuolumne, Yolo
Southern	Imperial, Los Angeles, Orange, Riverside, San Bernardino, San Diego, Santa Barbara, Ventura
Coastal	Alameda, Contra Costa, Monterey, Napa, San Benito, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma, Marin

APPENDIX B: SUPPLEMENTAL RESULTS

FORECASTING CHANGES IN CASELOAD

Accurately forecasting changes in the AFDC/TANF caseload is clearly important for planning purposes for the CDSS. In addition, accurate forecasts can aid in the evaluation process because they will allow us to assess the impact over a longer time period. In this section, we sketch the conceptual issues in forecasting the caseload.

One simple method for forecasting is to extend the time trend of total recipiency apparent in Figure 2.1. For example, we can forecast the caseload using a polynomial in time and allowing for seasonal variation.[23] We first present the actual change in caseload from fiscal year 1997-98 to fiscal year 1998-99 and the expected value of the change implied by the forecasting model in Table B.1, followed by one- and two-year forecasts.[24] First, note in Table B.1 that the actual decline from 1996-97 to 1997-98 is 13 percent. Because our forecasting methodology is based on the observed time trend, our forecasts for future declines continue: 19 percent between 1997-98 and 1998-99 and 25 percent between 1998-99 and 1999-00. Clearly, these forecasts follow from extending the time series, but it seems unreasonable to expect such large increases in caseload reduction rates.[25]

Another method of forecasting that will not force such large swings on the data is to explicitly model the caseload level as in equation (1). In these circumstances, reasonable forecasts of the underlying "forcing variables" must still be obtained, but many times these forecasts can be obtained from other sources.

---

[23] Specifically, we forecast the logarithm of the caseload using a fifth-order polynomial in time, monthly dummies to capture seasonal variation, and a second-order auto-regressive error structure. The reported forecasts are the expected value of the future caseload, and thus ignore the auto-correlated errors.

[24] The California fiscal year runs from July to June; thus, we forecast for this time period.

[25] For example, the CDSS forecasted declines of 11 percent and 8 percent respectively for these fiscal years. (See California Department of Social Services, 1998.)

Table B.1

Forecast of AFDC/TANF Caseload

Fiscal Year	Actual Level	Percent Change	Forecast Level	Percent Change	Difference Actual-Predicted
1996-1997	842,616		847,595		.59%
1997-1998	732,175	-13.11	734,628	-13.33	.33%
1998-1999			592,339	-19.37	
1999-2000			445,365	-24.81	

A final forecasting methodological point is that very short-term forecasts can be obtained fairly precisely by modeling flows rather than stocks. In particular, it is clear from Figure 2.5 that the approval rate of applications is quite stable over time. In addition, a processing lag exists for applications. Increases in the application rate in one period will cause the pool of applications to increase in the next period. Because the approval rate tends to be fairly stable, we can obtain fairly accurate forecasts of new welfare cases in the following periods. Thus, by focusing on changes in applications rates and the pool of applications to be processed, we can get fairly precise predictions of the caseload for the very near term.

TOTAL AFDC/TANF CASELOAD FOR CALIFORNIA BY MONTH

In Table B.2, we present the total caseload for California for January 1992 through March 1999. In addition, we present the Family Group (FG) and Unemployment Parent (U) caseloads separately.

CHANGES IN EXPENDITURE PER CASE

Figure 2.3 shows the change in AFDC/ TANF expenditure per case from July 1985 to March 1999, adjusted to 1998 dollars. Expenditures per case can change for three reasons. First, legislative action can cause the program generosity to change. Many of these actions can be observed as discrete movements in Figure 2.3. We note that cost-of-living adjustments (COLAs) are generally implemented infrequently (at most, once a year), thus they appear as discrete upward movements in the figure. Second, because cost-of-living-adjustments are not continually made, inflation will tend to erode the average payment per case, causing

a downward drift in the time series. Finally, changes in the distribution of case characteristics, such as family size or amount of earned income, will cause the expenditure per case to change; the direction of the change will depend on the particular distributional change.

Table B.2

AFDC/TANF Caseload for California by Month

Month	FG	U	Total	Month	FG	U	Total
Jan-92	701,164	124,402	825,566	Oct-95	771,914	170,675	942,589
Feb-92	700,406	127,108	827,514	Nov-95	764,719	170,049	934,768
Mar-92	705,779	131,005	836,784	Dec-95	762,305	171,224	933,529
Apr-92	707,496	132,517	840,013	Jan-96	763,397	173,192	936,589
May-92	707,672	132,645	840,317	Feb-96	761,139	174,951	936,090
Jun-92	709,863	132,559	842,422	Mar-96	762,262	176,873	939,135
Jul-92	712,951	132,943	845,894	Apr-96	760,988	177,240	938,228
Aug-92	716,450	133,412	849,862	May-96	758,184	176,368	934,552
Sep-92	719,686	134,640	854,326	Jun-96	753,876	174,020	927,896
Oct-92	723,952	135,459	859,411	Jul-96	751,314	171,772	923,086
Nov-92	721,117	136,350	857,467	Aug-96	745,386	168,055	913,441
Dec-92	726,363	140,244	866,607	Sep-96	739,998	165,371	905,369
Jan-93	725,937	143,429	869,366	Oct-96	735,829	163,676	899,505
Feb-93	728,598	147,094	875,692	Nov-96	727,321	161,172	888,493
Mar-93	737,733	151,845	889,578	Dec-96	721,380	160,943	882,323
Apr-93	741,110	154,099	895,209	Jan-97	715,820	160,821	876,641
May-93	742,539	154,971	897,510	Feb-97	709,166	159,805	868,971
Jun-93	745,835	155,616	901,451	Mar-97	705,173	159,453	864,626
Jul-93	747,196	155,728	902,924	Apr-97	697,195	156,669	853,864
Aug-93	751,813	157,005	908,818	May-97	688,742	153,351	842,093
Sep-93	754,398	157,625	912,023	Jun-97	680,041	149,892	829,933
Oct-93	756,891	158,333	915,224	Jul-97	671,931	145,706	817,637
Nov-93	757,490	159,404	916,894	Aug-97	665,516	141,812	807,328
Dec-93	760,645	162,363	923,008	Sep-97	658,532	138,609	797,141
Jan-94	763,262	165,047	928,309	Oct-97	655,256	136,358	791,614
Feb-94	765,278	167,388	932,666	Nov-97	644,704	133,201	777,905
Mar-94	772,888	171,182	944,070	Dec-97	641,509	133,001	774,510
Apr-94	774,162	172,449	946,611	Jan-98	631,542	132,409	763,951
May-94	775,873	172,989	948,862	Feb-98	621,891	131,844	753,735
Jun-94	777,508	172,658	950,166	Mar-98	618,453	132,636	751,089
Jul-94	773,916	171,189	945,105	Apr-98	610,129	131,000	741,129
Aug-94	778,640	171,865	950,505	May-98	602,310	129,174	731,484
Sep-94	778,284	170,650	948,934	Jun-98	596,460	127,179	723,639
Oct-94	778,339	170,533	948,872	Jul-98	589,028	124,473	713,501
Nov-94	775,681	170,073	945,754	Aug-98	581,518	121,582	703,100
Dec-94	778,914	171,645	950,559	Sep-98	575,361	119,216	694,577
Jan-95	780,670	173,849	954,519	Oct-98	570,797	117,663	688,460
Feb-95	780,041	175,113	955,154	Nov-98	561,779	115,678	677,457
Mar-95	784,500	177,560	962,060	Dec-98	559,150	115,961	675,111
Apr-95	780,206	176,985	957,191	Jan-99	555,375	115,982	671,357
May-95	780,273	176,898	957,171	Feb-99	546,817	115,211	662,028
Jun-95	778,582	175,347	953,929	Mar-99	544,719	115,585	660,304
Jul-95	773,268	173,205	946,473				

SOURCE: Authors' tabulations from the CA237.

Table B.3

Legislative Actions Affecting the AFDC/TANF Payment Levels

Date	Action
7/1/85	5.7% COLA
7/1/86	5.1% COLA
7/7/87	2.6% COLA
7/1/88	4.7% COLA
7/1/89	4.6% COLA
9/1/91	-4.4%
10/1/92	-4.5%
12/1/92	-1.3%
9/1/93	-2.7%
6/1/96	9.2%, -2.3%*
1/1/98	**
11/1/98	7.9%

Source: California Department of Social Services, Information Services Bureau (1997).

Notes: The action refers to the change on the Maximum Allowable Payment (MAP) for an Assistance Unit (AU) of three persons. COLA adjustments explicitly are tied to the rate of inflation.

* On 6/1/96, Exempt and Non-Exempt categories were created. The first number refers to the change for Exempt AUs and the second for Non-Exempt AUs.

** The change on 1/1/98 was the CalWORKs legislation. The major changes of this legislation are detailed in Zellman et al. (1999).

In Table B.3, we list the legislative actions that led to discrete generosity changes.

REFERENCES

Bane, M. J. and D. Ellwood (1994). *Welfare Realities: From Rhetoric to Reform*, Cambridge, MA: Harvard University Press.

California Department of Social Services, Estimates Branch and Financial Planning Branch (1998). "Caseload," November 1998 Subvention of Caseload Predictions.

California Department of Social Services, Information Services Bureau (1997). "A Comparison of California and National Public Assistance Statistics through December 1996," Sacramento, CA: California Department of Social Services.

Figlio, D. and J. Zilick (1999). "Welfare Reform, the Business Cycle and the Decline in AFDC Caseloads," March 1999 manuscript.

Hoynes, H. W. (1997). "Local Labor Markets and Welfare Spells: Do Demand Conditions Matter?" December 1997 manuscript.

Klerman, J. A., V. J. Hotz, G. Imbens, and P. Steinberg (1999). *Welfare Reform in California: Design of the Impact Analysis*, MR-1086-CDSS. Santa Monica, CA: RAND, 1999.

Wallace, G. and R. Blank (1999). "What Goes Up Must Come Down? Explaining Recent Changes in Public Assistance Caseloads," February 1999 manuscript.

Zellman, G., J. A. Klerman, E. Reardon, D. Farley, N. Humphrey, T. Chun, and P. Steinberg (1999). *Welfare Reform in Calfornia: State and County Implementation of CalWORKs in the First Year*, MR-1051-CDSS. Santa Monica, CA: RAND, 1999.