

Contract No.: 2002-23655  
MPR Reference No.: 8953-400

**MATHEMATICA**  
Policy Research, Inc.

**The Impact of the  
Children's Health  
Initiative (CHI) of  
Santa Clara County  
on Medi-Cal and  
Healthy Families  
Enrollment**

*Final Report*

*September 2004*

*Christopher A. Trenholm  
Sean Orzol*

*Programmer:  
Crandall Peeler*

Submitted to:

The David and Lucile Packard Foundation  
300 Second St., Suite 200  
Los Altos, CA 94022

Project Officer:  
Linda Baker

Submitted by:

Mathematica Policy Research, Inc.  
P.O. Box 2393  
Princeton, NJ 08543-2393  
Telephone: (609) 799-3535  
Facsimile: (609) 799-0005

Project Director:  
Christopher Trenholm



## ACKNOWLEDGMENTS

In preparing this report, we received support and assistance from many colleagues and evaluation stakeholders for which we are grateful: evaluation colleagues Embry Howell and Dana Hughes contributed their policy expertise and offered a number of thoughtful comments throughout the research; Michael Fitzwater and his colleagues at the Medical Care Statistics Section in the Department of Health and Human Services and Lesley Cummings and her staff at the Managed Risk Medical Insurance Board provided the enrollment data used in this report and expertly and patiently answered questions about these data as we conducted our analysis; members of CHI's Local Evaluation Group (LEG) provided ongoing guidance during the design and analysis phases of the study; in particular, LEG member Mary Cardenas (with the County of Santa Clara Social Services Agency) provided critical feedback on the selection and division of Medi-Cal eligibility codes for the analysis and LEG member Pat Posey (of the Santa Clara Family Health Plan) provided information on the counts of Healthy Kids members and other program features; Margo Rosenbach provided an excellent critical review; Cindy McClure adeptly produced the report; and Roy Grisham provided skilled editorial support. Any errors or omissions in this report are the sole responsibility of its authors.







## TABLES

<b>Table</b>	<b>Page</b>
1 CHILDREN’S INSURANCE OPTIONS IN SANTA CLARA COUNTY.....	2
2 SELECTED CHARACTERISTICS OF SANTA CLARA COUNTY AND THE MATCHED COMPARISON AREA.....	9
3 IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL IN SANTA CLARA COUNTY, BY ELIGIBILITY GROUP (2001-2002) .....	13
4 SENSITIVITY OF CHI IMPACTS TO QUALITY OF COMPARISON AREA MATCHES.....	18
5 SENSITIVITY OF CHI IMPACTS TO CHANGES IN ECONOMIC CONDITIONS .....	20



## FIGURES

<b>Figure</b>		<b>Page</b>
1	HEALTHY FAMILIES AND MEDI-CAL ENROLLMENT IN SANTA CLARA COUNTY .....	4
2	HEALTHY FAMILIES AND MEDI-CAL ENROLLMENT IN CALIFORNIA (STATEWIDE) .....	5
3	UNEMPLOYMENT RATE IN SANTA CLARA COUNTY AND STATEWIDE, 1999-2002 .....	6
4	IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002) .....	12
5	CUMULATIVE IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002) .....	15
6	DISTRIBUTION OF CHI IMPACTS ON NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002) .....	16



## A. INTRODUCTION

The Santa Clara County Children's Health Initiative (CHI), launched in January 2001, is an ambitious effort to extend health insurance coverage to all uninsured children in the county. The initiative combines the resources and expertise of several organizations, both public and private.<sup>1</sup> Through these organizations, CHI partner organizations direct staff and volunteers to reach families with uninsured children, evaluate their eligibility for public health insurance coverage, and assist them in enrolling in the appropriate program(s). Options for coverage include the two major statewide insurance programs, Medi-Cal and Healthy Families, as well a new county-based program, Healthy Kids, that CHI developed with local funding.<sup>2</sup>

As seen in Table 1, the new Healthy Kids program eliminates significant gaps in existing public coverage by extending eligibility to all uninsured children in the county with household incomes below 300 percent of FPL. This group includes children who are not eligible for existing programs because their family income is too high (between 250 and 300 percent of FPL), and children who are not eligible for existing programs because they do not meet legal residency requirements.<sup>3</sup>

---

<sup>1</sup>Major CHI partners include the Santa Clara Valley Health and Hospital System, the Santa Clara County Social Services Agency, the Santa Clara Family Health Plan, Working Partnerships USA (a local labor group), and People Acting in Community Together (a faith-based advocacy group). Among other CHI partners are many of the school districts in the county, along with a diverse group of public advocacy groups, community based organizations, and charitable foundations. For details on the formation of CHI and a description of the various activities that the initiative has undertaken, see Long (2001) and Howell and Hughes (2003). Also see CHI's website at [www.CHIkids.org](http://www.CHIkids.org).

<sup>2</sup>CHI activities are funded by a variety of city, county, and private funds. Among them are tobacco settlement funds allocated by Santa Clara County and the City of San Jose, Proposition 10 tobacco tax funds (controlled by the First 5 Commission of Santa Clara County), and several foundations (The David and Lucile Packard Foundation, The California Endowment, The Health Trust, and The California HealthCare Foundation).

<sup>3</sup>For a family of four, the FPL in 2004 is \$18,850 (Federal Register 2004). This corresponds to an income eligibility threshold for Healthy Kids of \$56,550 for a family of four.

TABLE 1  
CHILDREN’S INSURANCE OPTIONS IN SANTA CLARA COUNTY

Age	Income Limit as Percentage of Federal Poverty Level (FPL)	Program Eligibility
< 1 year	At or below 200 percent	Medi-Cal
	200-250 percent	Healthy Families
	0-300 percent	Healthy Kids <sup>a</sup>
1-5 years	At or below 133 percent	Medi-Cal
	133-250 percent	Healthy Families
	0-300 percent	Healthy Kids <sup>a</sup>
6-18 years	At or below 100 percent	Medi-Cal
	100-250 percent	Healthy Families
	0-300 percent	Healthy Kids <sup>a</sup>

Source: Children’s Health Initiative Website. <http://www.chikids.org/index.html>

<sup>a</sup>Children ineligible for Medi-Cal and Healthy Families are eligible for Healthy Kids up to 300 percent of FPL.

Through the creation of Healthy Kids, CHI has also instituted a fundamental change in the outreach message that families with uninsured children receive in the county. The message is now simple and direct—that your child(ren) will receive health insurance if you apply. This, in turn, has given Santa Clara County a potentially significant advantage in extending coverage to uninsured children by eliminating confusion over program eligibility, a factor long identified as a significant barrier to enrollment (Hughes et al. 2000 and 2001; and Perry et al. 1998). In addition, the creation of Healthy Kids has helped CHI to recruit and coordinate a large cadre of supporters, from paid outreach staff and volunteers at the local partnership organizations to key public officials and other local stakeholders.

In this paper, we evaluate the impact CHI has had on children’s enrollment in the two major state insurance programs, Medi-Cal and Healthy Families. Using enrollment data for these two programs over the period 1999-2002, we compare the pre- to post-change in enrollment in these programs in Santa Clara County to the change taking place in a carefully matched comparison

area comprised of zip codes elsewhere in the state. Findings indicate that CHI led to a significant gain (about 28 percent) in enrollment in both programs. These findings appear quite robust. They are evident across the county and throughout the two-year period (2001-2002) that we examined.

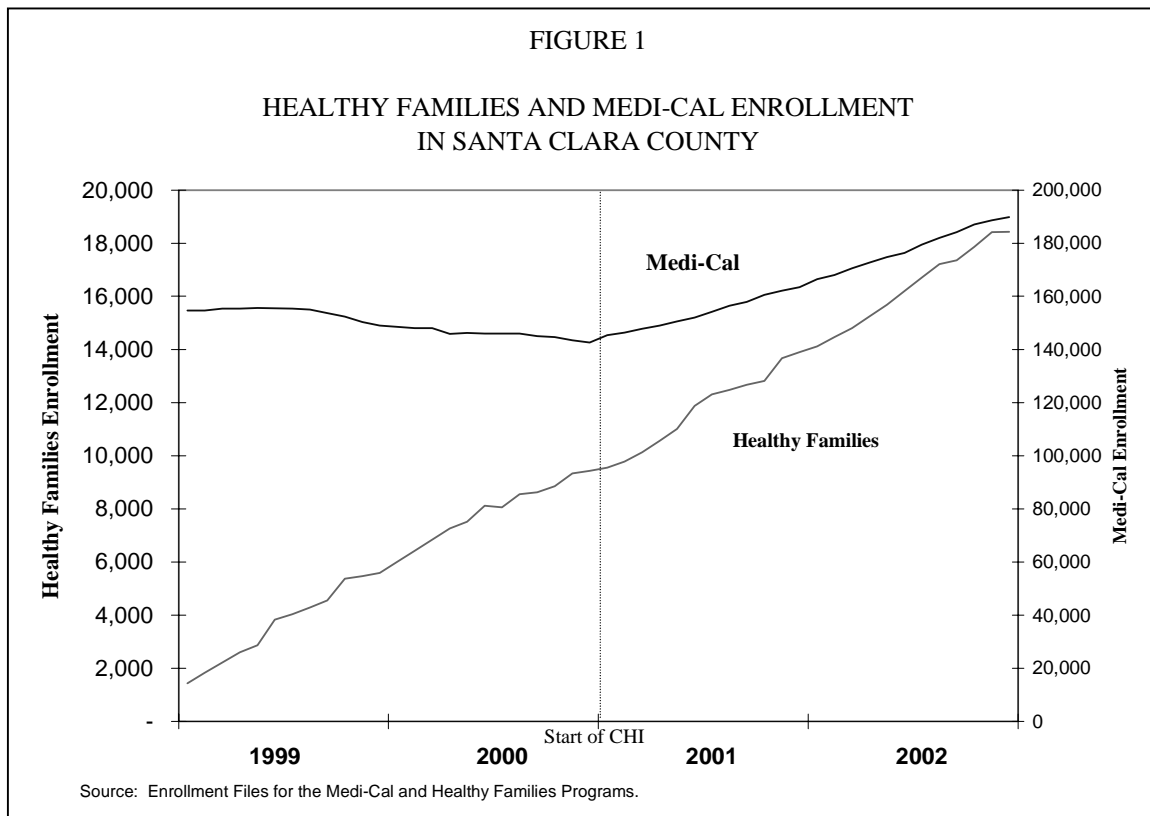
## **B. EVIDENCE TO DATE ON THE EFFECTS OF CHI**

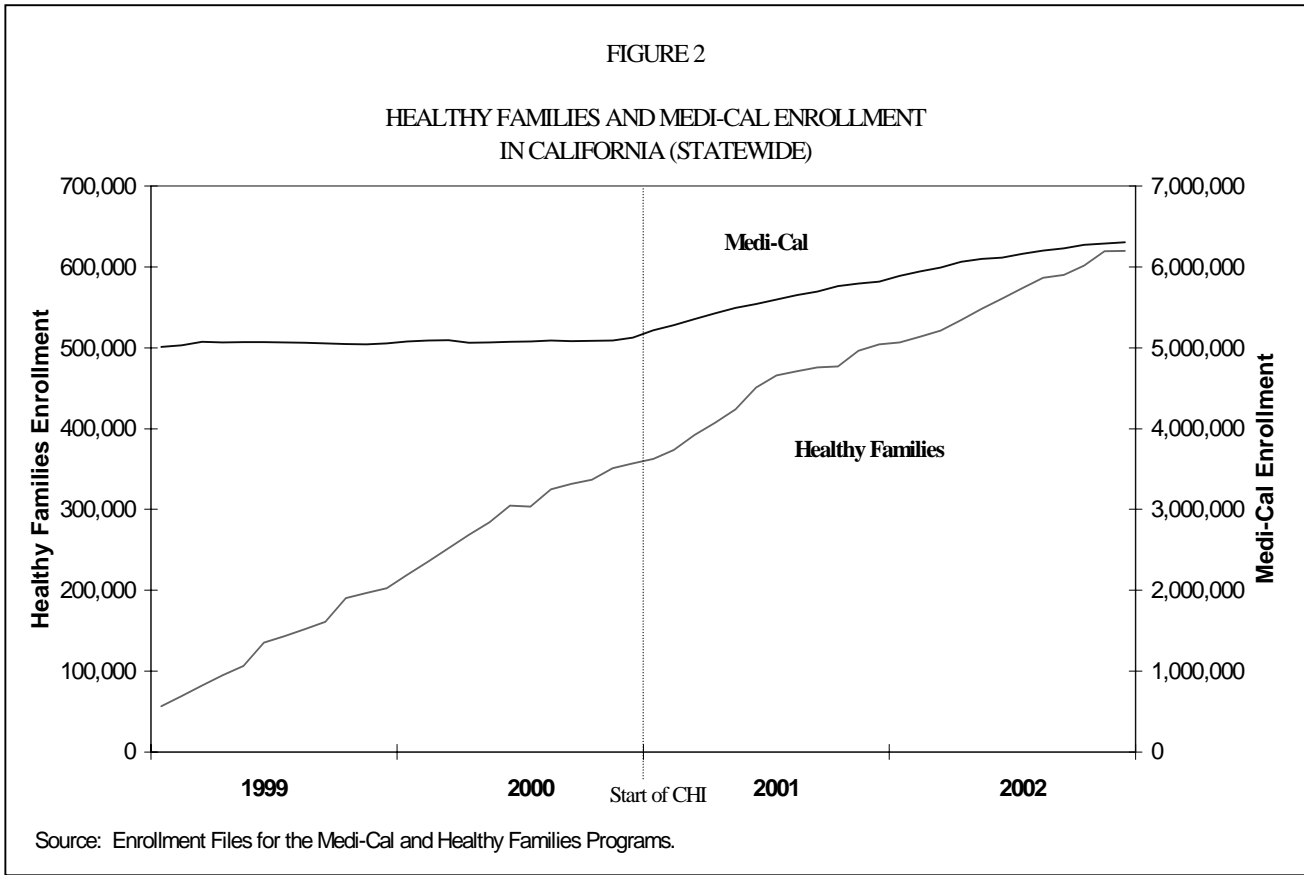
Qualitative evidence suggests that CHI began quickly and aggressively to meet its goal of extending health insurance coverage to all children in the county. Within the first six months of operation, CHI developed a simplified two-page application for Healthy Kids and worked to streamline and coordinate the application process for Healthy Families and Medi-Cal (Long 2001). It also began a significant “inreach” effort concentrated mainly on identifying uninsured children through the county’s health care system, other community clinics, and the social service system and enrolling them in health insurance. Families who had recently moved into the county were a focal target of these efforts. CHI also increased its community-based outreach (including local media advertising, informational campaigns, and community events) and eventually expanded into school-based outreach as well. These broader outreach strategies reportedly were uneven across the county, both in intensity and type, and were viewed by some as being less productive than the inreach efforts (Howell and Hughes 2003). Such efforts, however, were also likely to reach different types of families than those identified through inreach, most notably those inexperienced with the public health care system.

Enrollment in the county-based Healthy Kids program suggests that the initiative made significant progress toward its goal, at least for those who did not meet the eligibility requirements for Medi-Cal and Healthy Families. By the middle of its first year of operation, more than 4,500 children had enrolled in the Healthy Kids program, and by the end of the year, the program’s ranks had swelled to 7,800. Enrollment then grew steadily in 2002 and into 2003,

long after the point that CHI's inreach efforts had identified the most obvious candidates for coverage. Coverage leveled off in spring 2003 at roughly 13,000 children, but only because demand had begun to exceed funding, necessitating a cap in enrollment.

In contrast to the clear gains in coverage made through the county's Healthy Kids program, to date evidence that CHI has had an effect on enrollment in either Medi-Cal or Healthy Families is less compelling. As seen in Figure 1, enrollment in both programs did rise steadily from the start of CHI, presumably indicating that CHI has had some success. However, increasing enrollment on these programs is also seen statewide (Figure 2), making the size of CHI's contribution to the enrollment increase seen in Santa Clara difficult to assess. Indeed, at the same time that CHI began serving Santa Clara, employment in the county and throughout the state began a significant downturn, one that might easily account for much of the observed

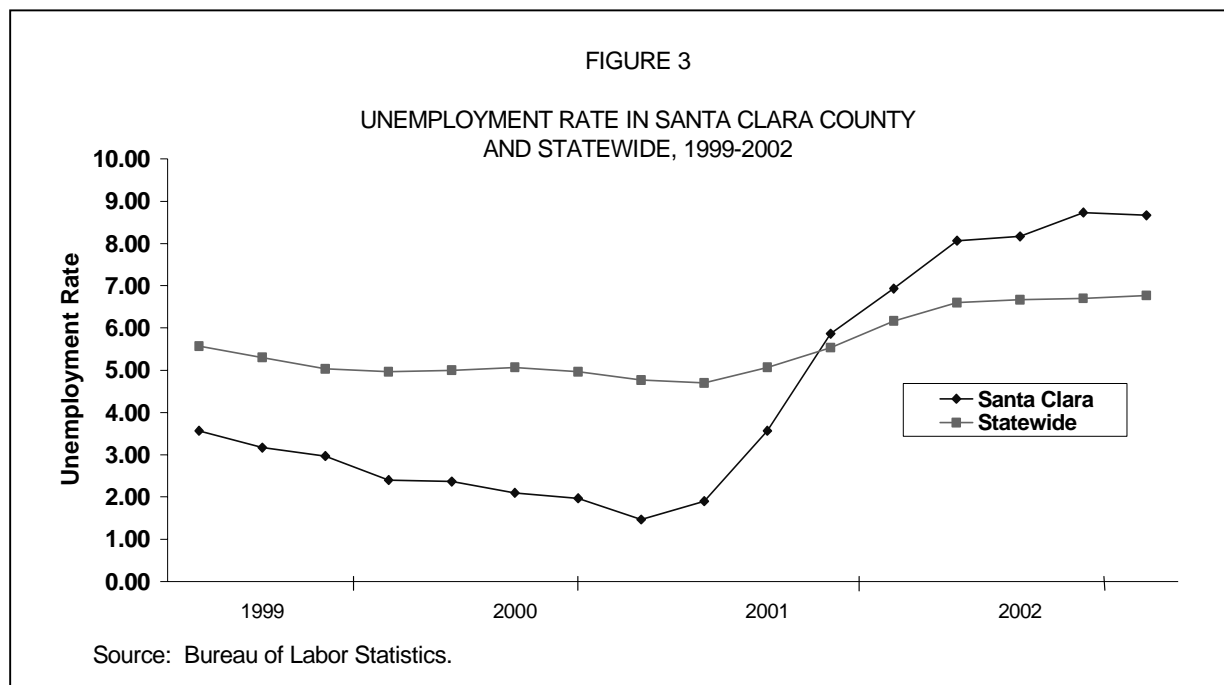




increases in enrollment (Figure 3). Teasing out these various factors to provide a credible estimate of CHI's effects on enrollment requires the type of rigorous evaluation that this study undertakes.

### C. STUDY DESIGN

In order to rigorously evaluate the impact of CHI, we conducted a quasi-experiment that compares the enrollment taking place in Medi-Cal and Healthy Families in the two years before CHI (1999-2000) and after the start of CHI (2001-2002) between Santa Clara County and a carefully matched comparison area. Below, we briefly summarize this design, focusing on the data sources and the general methods for developing the comparison area and conducting the estimation. For a more complete description on the study design, see the Methods Appendix (Appendix A).



## 1. Data Sources

To construct our focal measures of Medi-Cal and Healthy Families enrollment before and after CHI, we use data from monthly Medi-Cal enrollment extracts for the period January 1998 through April 2003. Each extract includes information on the current month's enrollees, including their eligibility and recipient status over the previous 12 calendar months and selected demographic information such as their zip code, birth date, and race/ethnicity. Each extract also includes a unique identifier that can be used to link enrollees across the extract files and identify the start and length of each spell of coverage.<sup>4</sup>

To measure the characteristics of the zip code regions used in the analysis, and to account for any important differences between them, we used the Census 2000 long form data for

---

<sup>4</sup>We used data from the Medi-Cal file to identify spells of coverage in both programs because they allowed us to distinguish children who newly enrolled in one of these programs from children who transferred between them. Healthy Families enrollment data over this same period provided the source to validate the counts of new entries that we obtained from the Medi-Cal records.

California. Among the characteristics identified from the Census and used in the analysis were the population counts and population distribution by age, race/ethnicity, income, and place of birth. In addition, to account for differences in economic conditions over our period of analysis, we used two sources: (1) monthly unemployment data available at the county level from the U.S. Department of Labor Bureau of Labor Statistics, and (2) counts of total employment by industry and occupation at the zip code level available from California's Employment Development Department of Labor Market Information. The former data were used in the main analysis, while the latter supported several sensitivity analyses.

## **2. Forming the Comparison Area**

To form our comparison area for the study, we carefully matched each of 42 zip codes in Santa Clara to a set of comparison zip codes elsewhere in the state.<sup>5</sup> This approach has several benefits over a more traditional design that would form the comparison area by selecting one or more counties that most closely resembled Santa Clara. First, by forming our comparison area at a more disaggregated level, we can better reflect the substantial variation that exists within Santa Clara in key demographic and other characteristics that may be associated with enrollment. Second, because there are far more zip codes than counties to choose from (2,827 zip codes compared with 58 counties), we can identify far more credible comparison zip codes to form the comparison area. Third, the approach yields more observations and, thus, greater statistical precision for estimating the effects of CHI on enrollment. Fourth, it allows for a far more rigorous assessment of the robustness of our results.

---

<sup>5</sup>Santa Clara County includes 58 zip codes, but some of these zip codes were combined or dropped in our analysis because they had few or no children enrolled in Medi-Cal and Healthy Families during 2001 and 2002. The resulting sample for the analysis included 42 zip codes (or groups of zip codes) in Santa Clara County, each of which we matched to a set of comparison zip codes elsewhere in the state.

To form the comparison area, we used an iterative process that, for each Santa Clara zip code, identified those zip codes elsewhere in the state that were most similar across nine key characteristics (see Table 2). At the first iteration, the “tolerance level” for identifying these comparison zip codes was set to zero, meaning that only zip codes that matched exactly across these nine characteristics would be selected as a comparison area member. At each successive iteration, the tolerance level was expanded slightly for each variable, creating the opportunity to identify additional comparison zip codes for each Santa Clara zip code. This process ultimately led to the selection of 512 comparison zip codes (at least 10 for each Santa Clara zip code).

As seen in Table 2, the resulting comparison area is similar to Santa Clara County across each of the nine characteristics. Perhaps most important, the rates of coverage prior to CHI (measured as a share of the children under 200 percent of FPL) are quite comparable for both Healthy Families and Medi-Cal. In fact, when we combine the rates for the two programs, we see that they are nearly identical between Santa Clara and the comparison area (both about 63 percent). The percentage of the population in poverty is also quite similar, although there is a larger difference in the share of children who are under 200 percent of the FPL (about 30 percent for Santa Clara and 35 percent for the comparison area). Santa Clara has a similar proportion of Hispanic children as the comparison area (36 percent versus 35 percent), but it has a somewhat higher share of Asian children and of foreign-born non-citizens.

### **3. Outcome Variables**

The central outcome for the analysis is the count of new entries to Medi-Cal or Healthy Families in each quarter from 1999 through 2002, a period that includes two years before and after the start of CHI. We define a new entry as any child who enrolled in Medi-Cal or Healthy Families during this period and did not have this coverage in the preceding 11 months. (With

TABLE 2  
 SELECTED CHARACTERISTICS OF SANTA CLARA COUNTY  
 AND THE MATCHED COMPARISON AREA

Characteristics	Santa Clara (Mean)	Comparison Area (Mean)
Percent of Eligible Children (Under 200% of FPL) Enrolled in Medi-Cal	54.7	52.5
Percent of Eligible Children (Under 200% of FPL) Enrolled in Healthy Families	8.2	10.8
Percent of Population with Income Below FPL	10.0	10.9
Percent of Children (0-17) in Household with Income Under 200% of FPL	29.7	34.2
Population Density (Pop./Square Miles)	6,964	6,114
Percent Hispanic	35.8	34.7
Percent White	33.8	37.2
Percent Asian	23.8	18.8
Percent Foreign Born – Not a U.S. Citizen	23.6	17.8
<b>N(Number of Zip Codes)</b>	<b>42</b>	<b>512</b>

Source: Census 2000 and enrollment data from Healthy Families and Medi-Cal.

Note: Statistics are computed as the mean value for the zip codes in Santa Clara County and the zip codes that comprise the matched comparison area used to estimate impacts. Statistics have been weighted, based on the number of children under 200 percent of poverty in each Santa Clara zip code. Some zip codes in Santa Clara County with very small numbers of children enrolled in public coverage have been combined. See the Methods Appendix for additional information.

FPL = Federal Poverty Level.

this definition, we eliminate from the study sample children who enrolled in Medi-Cal or Healthy Families as a result of a transfer and children who re-enrolled or “cycled” back onto one of these programs after only short time; neither group was a major focus of CHI during 2001-2002). This outcome is measured at the zip code level by adding up all the new entries occurring in each zip code in our sample for each quarter from 1999 through 2002.

The outcome is measured in three ways. First is the count of new entries in Medi-Cal and Healthy Families combined. Second is the count of new entries in these two programs separately—that is Medi-Cal alone and Healthy Families alone. Third is the count of new entries in each of four broad eligibility classifications within Medi-Cal: (1) poverty-related groups, (2) Medically Needy, 1931, and Transitional Medi-Cal groups; (3) TANF-related groups; and (4) SSI and other groups. (For more information on the definition of these groups, see Appendix A). We hypothesize that any impacts of CHI should be largest for the Healthy Families program and for the poverty-related Medi-Cal group because their families are the most likely to receive assistance with their application and/or to learn about their eligibility through outreach. In contrast, impacts should be least evident for the TANF-related groups and, to an even greater extent, the SSI/other groups, because these families typically obtain Medi-Cal coverage while applying for the other programs.

#### **4. Impact Estimation**

We estimate the impacts of CHI on enrollment through a “difference-of-difference” regression model.<sup>6</sup> With this model, the effect of CHI is measured as the difference in the *change* in quarterly enrollment (new entries) after the start of CHI between Santa Clara County and the comparison area. This model is particularly attractive when using an external

---

<sup>6</sup>For details on the exact specification of the model and the calculation of the CHI impact, see Appendix A, “Methods.”

comparison group design because it controls for all differences between Santa Clara and the comparison area that affect enrollment and are approximately fixed or stable over time.

The model is further strengthened through the addition of explanatory (or control) variables that might affect entry into public coverage. These variables are used to account for potential differences between the zip codes that may not have been addressed through our matching process or through our difference-of-difference estimator. They include: (1) the number of children below 200 percent of poverty; (2) the percentage of the population in poverty; (3) the zip code's racial and ethnic distribution; (4) the population density; and (5) the percentage of non-citizens. Each of these variables is created from the Census 2000 data, so that they measure only a single point in time. In addition, to account for time-varying economic factors, we included in each quarter a variable measuring the unemployment rate at the county level.

#### **D. FINDINGS**

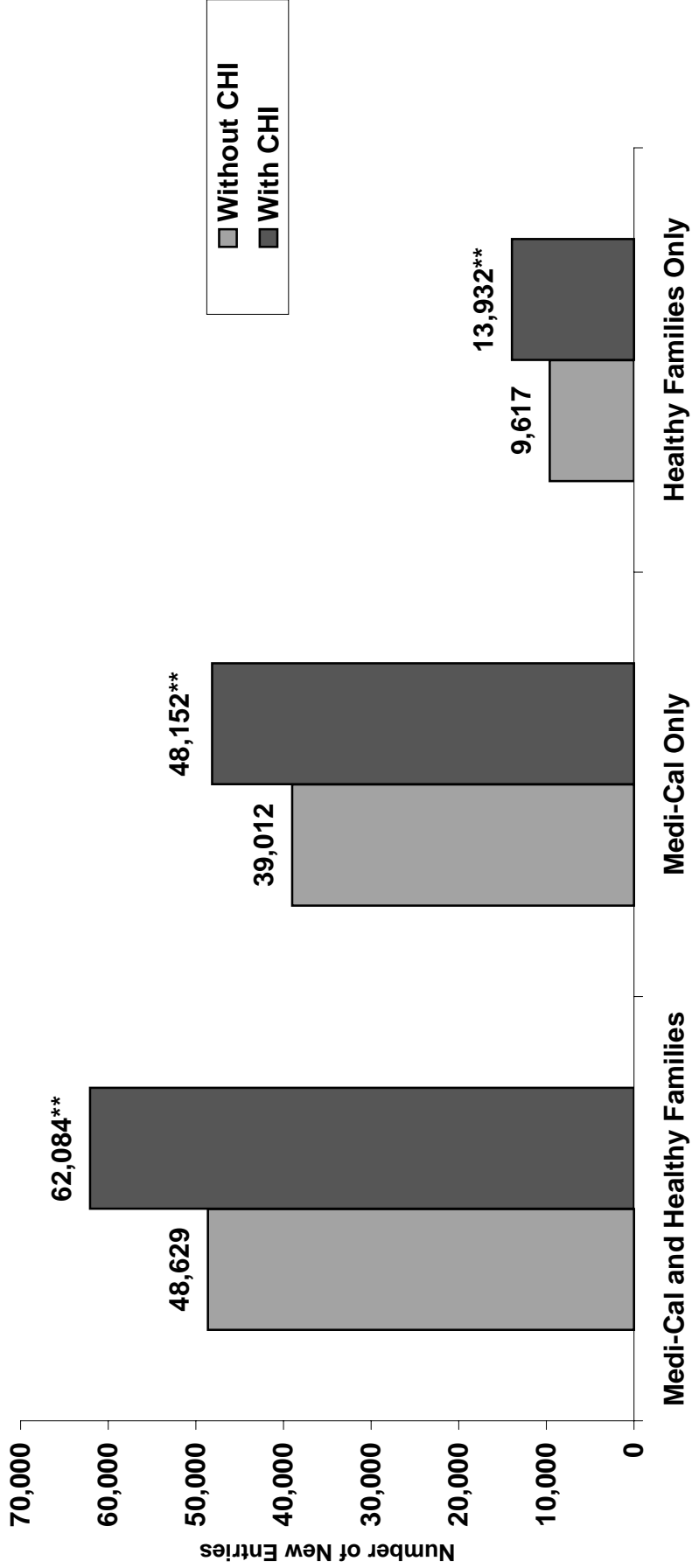
During its first two years, CHI increased the number of new entries in Medi-Cal and Healthy Families in Santa Clara County by a combined 28 percent, from 48,629 to 62,084 (Figure 4).<sup>7</sup> This increase resulted from significant gains in both programs. For Medi-Cal, the number of new entries rose from 39,012 to 48,152, or 23 percent. For Healthy Families, the percentage increase was even larger, from 9,617 to 13,932, or 45 percent.

Within the Medi-Cal program, CHI had by far its largest effect on children in the poverty-expansion group—72 percent (Table 3). This result is consistent with our expectation that children qualifying for Medi-Cal within this eligibility group will be reached most often through the outreach efforts of the initiative. In contrast, CHI had a relatively modest effect on children qualifying under the TANF-related groups (16 percent) and essentially no effect on children

---

<sup>7</sup>All impact estimates presented in this section have been regression-adjusted.

**FIGURE 4**  
**IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002)**



Source: Enrollment Files for the Medi-Cal and Healthy Families Programs.

Notes: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. The number of new entries without CHI is estimated from a matched comparison area. This estimate has been regression adjusted in order to account for observed differences between the comparison area and Santa Clara County. See Appendix A for additional details.

\*\*Difference is significantly different from zero at the .01 level.

TABLE 3

## IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL IN SANTA CLARA COUNTY, BY ELIGIBILITY GROUP (2001-2002)

	New Entries Per Quarter		
	Without CHI	With CHI	Impact (Percentage)
Overall Medi-Cal	39,012	48,153	23.4**
Medi-Cal Poverty-Related	6,153	10,570	71.8**
Medi-Cal, Section 1931/Medically Needy, Continuous Eligibility, TMC <sup>a</sup>	17,556	20,757	18.2**
Medi-Cal, TANF-Related	7,620	8,838	16.0**
Medi-Cal SSI/Other	7,683	7,988	4.0

Source: Enrollment Files for the Medi-Cal and Healthy Families programs.

Notes: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. The number of new entries without CHI is estimated from a matched comparison area. This estimate has been regression adjusted in order to account for observed differences between the comparison area and Santa Clara County. See Appendix A for additional details.

<sup>a</sup>Section 1931 requires Medi-Cal to be provided to low-income families who meet the requirements of the Aid to Families with Dependent Children (AFDC) State plan in effect July 16, 1996. TMC = Transitional Medi-Cal.

\*\* Impact significantly different from zero at the .01 level.

qualifying under the SSI/other categories (4 percent; not statistically significant). These findings are also quite consistent with our expectation, since many children qualifying for coverage within these groups will obtain their Medi-Cal coverage as part of their application process for the other programs.

## **1. Distribution of CHI Impacts**

Viewed over time, we see that the impacts of CHI appear almost immediately after the start of the initiative and continue, if not strengthen, over the two years that follow (Figure 5). During the first quarter following the start of CHI, 5,478 children entered Healthy Families or Medi-Cal in Santa Clara County, compared with an estimated 5,202 children in absence of CHI—a modest difference of about 276 children. However, this difference widened at a dramatic rate to more than 5,000 children by the end of 2001 and over 13,000 children by the end of 2002.

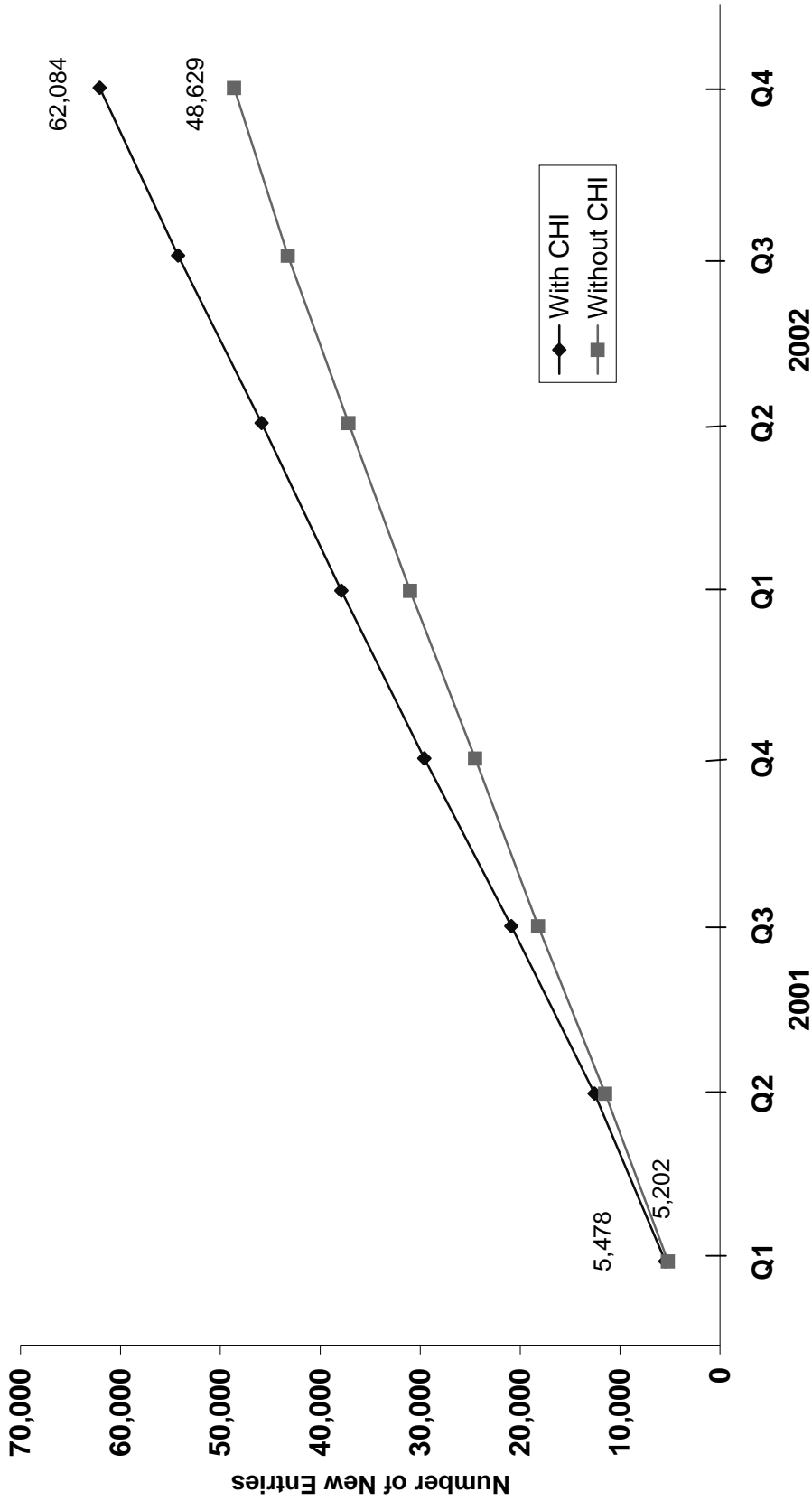
Looking across the county, we see the effects of CHI are widely dispersed and appear to be greatest in those sections of the county with a large number of low-income children (Figure 6).<sup>8</sup> Of the 42 zip codes or groups of zip codes that form our Santa Clara sample, 32 have impacts of 5 percent or more, and 17 have impacts in excess of 15 percent. Only four zip codes displayed negative effects. The largest impacts, those in excess of 15 percent, appear most often in the central parts of the county east of downtown San Jose. These areas include a large share of the low-income children in the county, making them natural areas for CHI to focus its attention. (See Appendix Figure A.1 for the distribution of low-income children in the county.) In contrast, the few negative impacts are located in the northwest sections of the county, where there are few

---

<sup>8</sup>The impact of CHI in each zip code shown in Figure 6 is estimated as the difference in the number of new entries between the individual zip code and its matched comparison area. Due to the small sample size, very few of these differences are statistically significant. Nevertheless, they provide a useful perspective on the distribution of CHI's impacts across the county.

FIGURE 5

CUMULATIVE IMPACT OF CHI ON THE NUMBER OF NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002)

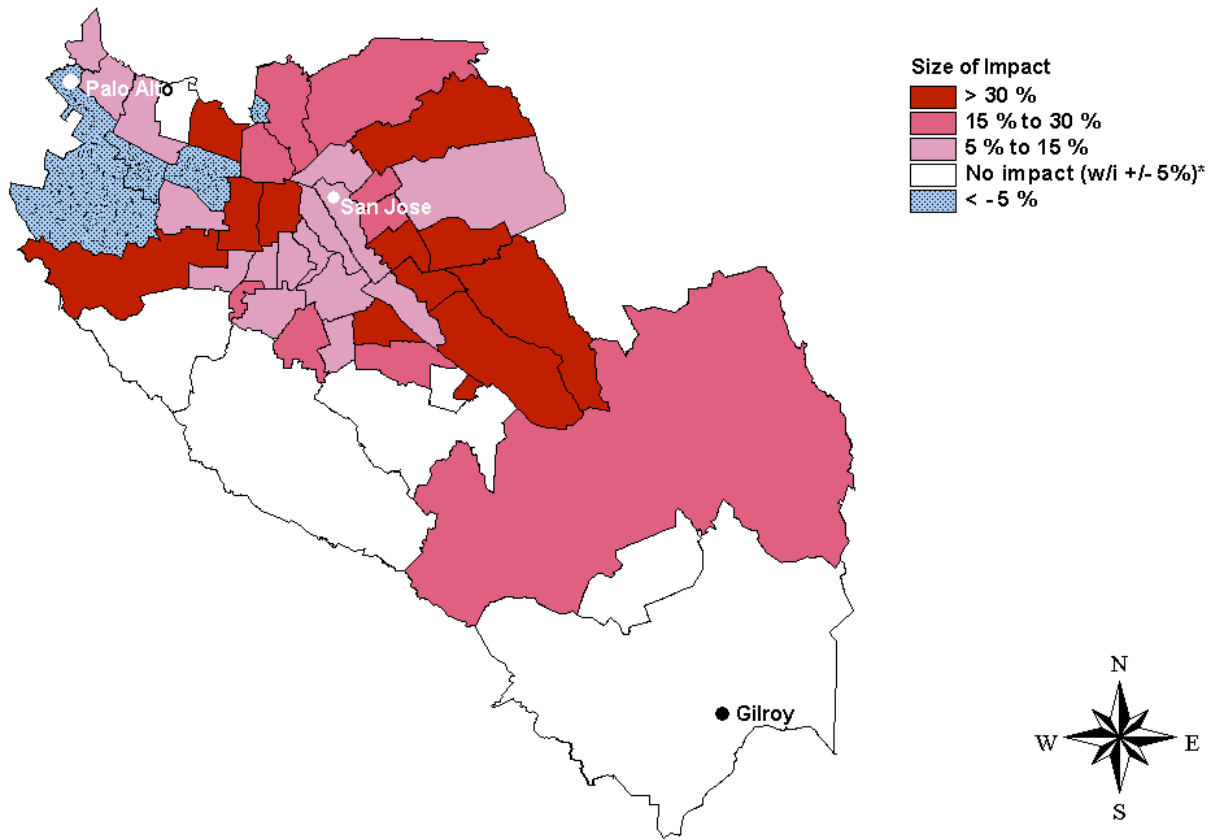


Source: Enrollment Files for the Medi-Cal and Healthy Families Programs.

Note: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. The number of new entries without CHI is estimated from a matched comparison area. This estimate has been regression adjusted in order to account for observed differences between the comparison area and Santa Clara County. See Appendix A for additional details.

FIGURE 6

DISTRIBUTION OF CHI IMPACTS ON NEW ENTRIES IN MEDI-CAL AND HEALTHY FAMILIES IN SANTA CLARA COUNTY (2001-2002)



Source: Enrollment Files for the Medi-Cal and Healthy Families programs.

Note: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. For each region (zip code) of the county shown in the figure, the impact of CHI is estimated as the difference in the number of new entries between the region and its matched comparison area. See Appendix A for additional details.

low-income children. Impacts in the southern and southwestern sections of the county are decidedly mixed, ranging from negligible to large. Much of this region is sparsely populated, which may present more of a challenge for CHI in trying to reach uninsured children and enroll them into public coverage.

## **2. Sensitivity Testing**

The consistent evidence of CHI impacts—both within and across the Healthy Families and Medi-Cal programs, and over the first two years of the initiative—strongly supports a finding that CHI generated large gains in coverage of uninsured children in the county. To further explore the robustness of this finding, we conducted two general sensitivity analyses. The first examines the extent to which our impacts might be driven by the selection of “low-quality” comparison zip codes that are dissimilar to those in Santa Clara County to which they are matched. The second analysis examines the extent to which our findings might be driven by changes in economic conditions that we have failed to account for.

To investigate the sensitivity of our impacts to the quality of the comparison area, we divided the comparison zip codes into four groups (from “A-level” through “D-level”) based on how closely they matched to their Santa Clara zip code and reestimated the CHI effects. As seen in Table 4, when we limit the comparison sample to the A through C matches (column 2) or to the A and B matches (column 3), the impact estimates display only a modest change. When we further limit the sample to A-level matches alone, the impacts display a more notable decline, from an original estimate of 28 percent to a revised estimate of 14 percent. However, the source of this decline appears not to be the match quality but rather the loss of several Santa Clara County zip codes with the largest impacts (because they lack an A-level match). While it is possible that the lack of these “highest-quality” matches might have contributed to the large impact estimates in these zip codes, we find no evidence that this is the case. In fact, impacts

TABLE 4

SENSITIVITY OF CHI IMPACTS TO QUALITY OF COMPARISON AREA MATCHES

	(1) Full Comparison Area	(2) Limited to A-C Level Matches	(3) Limited to A-B Level Matches	(4) Limited to A Level Matches
Overall	27.7**	25.3**	19.6**	13.5**
Medi-Cal	22.8**	23.4**	17.2**	13.4**
Healthy Families	44.9**	36.2**	32.8**	21.2**
N (Santa Clara)	42	42	39	26
N (Comparison)	512	455	369	154

Source: Enrollment Files for Medi-Cal and Healthy Families Programs.

Notes: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. The number of new entries without CHI is estimated from a matched comparison area. This estimate has been regression adjusted in order to account for observed differences between the comparison area and Santa Clara County. See Appendix A for additional details.

Level of match (A, B, C, or D) refers to the degree of consistency between the Santa Clara zip code and its matched-comparison zip codes across all of the demographic characteristics used to form the comparison area. A-level matches had the highest degree of consistency across all characteristics; for other matches, there was relatively less consistency across at least one characteristic. See Methods Appendix for additional information.

\*\* Impact significantly different from zero at the .01 level.

remain very large for each of these zip codes when we limit the sample to the best available B-level matches (not shown).

A second, more substantial sensitivity analysis examined the extent to which our impact estimates might be biased by economic conditions that are not fully accounted for in our model. This concern is heightened by the sharp and sustained increase in unemployment that took place in Santa Clara around the time that CHI was adopted (see Figure 3). In our main specification, we account for this change by including a control variable for the unemployment rate in each quarter. However, this rate is measured at the county level, thus providing a potentially weak estimate of the economic conditions for residents in a particular zip code. In addition, entry into public coverage may be affected by unemployment in a more complex way than the linear relationship assumed in our original specification.

As seen in Table 5, impact estimates are insensitive to a variety of alternative measures designed to control for changes in economic conditions over time (columns 1 to 3). The two alternatives shown in Table 5 include: the use of dummy indicators for unemployment, which allows for a non-linear relationship between enrollment and economic factors (column 2); and the use of additional controls for employment conditions, most notably an estimate of the rate of employment activity in the zip code based on employment counts over time (column 3). In both cases, the impacts of CHI display little change. Impacts also remained robust using several other alternative measures and specifications of economic conditions over time (not shown in Table 5).

Estimates do, however, decline notably when we limit the comparison sample to counties that experienced a rise in unemployment during 2001-2002 closest to the level experienced in

TABLE 5

## SENSITIVITY OF CHI IMPACTS TO CHANGES IN ECONOMIC CONDITIONS

	(1) Unemployment Rate (Original Estimate)	(2) Unemployment Rate (Dummy Indicator)	(3) Additional Employment Controls	(4) Similar Economic Downturns
Overall	27.7**	25.8**	26.5**	15.8**
Medi-Cal	22.8**	22.7**	24.9**	19.7**
Healthy Families	44.9**	40.9**	35.0**	5.9
N (Santa Clara)	42	42	42	37
N (Comparison)	512	512	512	148

Source: Enrollment Files for Medi-Cal and Healthy Families Programs. Unemployment data from U.S. Department of Labor Bureau of Labor, Statistics. Employment data from California's Employment Development Department of Labor Market Information.

Notes: New entries includes all children enrolling in either Medi-Cal or the Healthy Families program who have not been covered by one of these programs in the prior eleven months. The number of new entries without CHI is estimated from a matched comparison area. This estimate has been regression adjusted in order to account for observed differences between the comparison area and Santa Clara County. See Appendix A for additional details.

\* Impact significantly different from zero at the .05 level.

\*\* Impact significantly different from zero at the .01 level.

Santa Clara County (Column 4).<sup>9</sup> Overall, the impact estimates fall by close to half, from 28 percent to 16 percent, although they remain statistically significant ( $p\text{-value} < 0.01$ ). The estimated impact on Healthy Families experiences a particularly steep decline, from 45 percent to 6 percent (not significant;  $p\text{-value} = 0.36$ ). These estimates prove highly volatile, however, because they are based on only 148 of the original 512 observations in the comparison sample (and only 63 of the 282 unique zip codes in that sample). In fact, if we drop only 5 of the 63 zip codes from this sample, the impact estimates range from as low as 15 percent to as high as 20 percent and the effect for Healthy Families is as high as 12 percent ( $p\text{-value} = 0.08$ ). Moreover, estimates based on other restricted samples, including those limited to counties with unemployment rates closest to Santa Clara during 2001-2002, display impacts more consistent with the full comparison area (not shown). As a result, while these estimates provide some evidence that our impacts may be overstated, this evidence is limited.

## **E. DISCUSSION**

This paper finds strong evidence that CHI generated large increases in enrollment in the two major public insurance programs in the state, Healthy Families and Medi-Cal. Over our two-year observation period, from 2001 through 2002, we estimate that, in absence of CHI, the county would have added roughly 48,629 children to these two programs. With the adoption of CHI, it added about 62,084 children, an increase of about 13,500 children or 28 percent. Interestingly, the enrollment increase in these programs is very similar to the number of children enrolled in the Healthy Kids program by the end of this same period (12,155).

---

<sup>9</sup>This is one of several restricted samples that we examined to explore the possible effects of economic factors on our estimates. For this specification, we restricted the sample to 5 counties with the largest percentage increase in the unemployment rate during 2001-2002. The counties include San Mateo, Alameda, Marin, San Francisco, and Contra Costa.

Given that the Medi-Cal and Healthy Families programs are entirely state and federally funded, CHI's impacts on enrollment in these programs led to a substantial increase in funding for the county. Based on the per-member per-month funding for each child that CHI added to these programs, the initiative increased state and federal spending in the county by an estimated \$24.4 million during its first two years (2001-2002). Moreover, these spending gains continued to accrue beyond 2002, as more than half the new entries added to Medi-Cal and Healthy Families in these first two years remained enrolled into 2003.

By far the most notable threat to the credibility of these findings is the economic downturn experienced in Santa Clara County since the start of CHI. This downturn is not fully replicable in our comparison sample because Santa Clara County experienced a more severe downturn than any other county in the state. As a result, we cannot rule out that some of the estimated gains in enrollment attributed to CHI are the result of changes in economic conditions. At the same time, we find limited evidence that this is the case. The estimated effects of CHI were lower when we restricted the comparison sample to a select group of counties with downturns closest to Santa Clara. However, this sample includes only a small fraction of the original comparison zip codes, making it much more volatile and vulnerable to potential bias from a small number of poorly chosen matches.

While this paper provides evidence that CHI is experiencing success, it is important to caution that we can offer little insight about how close CHI has come to meeting its core goal of extending universal coverage for children in the county. Gaining such insight would require credible estimates of the number of uninsured in the county over time, which are currently

unavailable.<sup>10</sup> In addition, it would require a thorough assessment of other factors that affect uninsurance in the county, such as program retention and access to private coverage, both of which are beyond the scope of this analysis. To date, efforts to improve retention and address other factors related to uninsurance have been less a focus of CHI than efforts to reach and enroll eligible children. Nevertheless, they are likely to be key determinants of how close CHI can come to reaching its ultimate goal, making them important issues to examine in the future.

By all accounts, the creation of the Healthy Kids program has been central to the success of CHI. In their recent review of the initiative, Howell and Hughes (2003) note that Healthy Kids made it possible for CHI to articulate a “clear, achievable goal” of providing health insurance to children throughout the county. This, in turn, generated a catalyst for launching the initiative quickly and aggressively, as well as for generating the significant support that would be needed to expand its outreach and enrollment efforts. The implementation of Healthy Kids was itself facilitated by several key decisions, including the choice of a single health plan (the Santa Clara Family Health Plan) to administer the program and the decision to use an existing program, Healthy Families, as the basis for its design. Together, these decisions allowed the CHI partnership to simplify the process of establishing this new health insurance program and to move quickly toward its goal, once the program was adopted.

CHI also appears to have benefited significantly from the strong and sustained leadership of the lead partner organizations (Hughes and Howell 2003; and Long 2002). These organizations encompass not only the county’s entire public health and hospital system, but also its social

---

<sup>10</sup> Wong (2000) estimated the number of uninsured in Santa Clara prior to CHI based on data from the Current Population Survey (CPS). While it is difficult to assess the accuracy of these estimates, Czajka and Lewis (1999) have shown that local area estimates of the uninsured from the CPS and other national datasets are unreliable.

services agency, its major public health insurer, and several key outreach and advocacy organizations in the county. For communities with the interest and financial resources to replicate Santa Clara's CHI and Healthy Kids program, the ability to form and sustain this type of committed partnership may be a critical factor to their success.

## REFERENCES

- California Employment Development Department. "2003 County Snapshots: Santa Clara County, California." [<http://www.calmis.ca.gov/file/cosnaps/santcSnap.pdf>]
- Czajka, John, and Kimball Lewis. "Using National Survey Data to Analyze Children's Health Insurance Coverage: An Assessment of the Issues." Washington, DC: Mathematica Policy Research, May 1999.
- Duan, Naihua. "Smearing Estimate: A Nonparametric Retransformation Method." *Journal of the American Statistical Association*, Vol. 78, 1983, pp. 605-610.
- Federal Register, vol. 69, no. 30, February 23, 2004, pp. 7336-7338.
- Howell, Embry, and Dana Hughes. "CHI: Factors in Early Success and Challenges for the Future." Prepared for the David and Lucile Packard Foundation, November 2003.
- Hughes, Dana, et al. "Barriers to Enrollment in Healthy Families and Medi-Cal: Findings from Focus Groups with Chinese, Korean, and Vietnamese Parents." San Francisco: Center for Children's Access to Health Care Institute, Institute for Health Policy Studies, December 2000.
- Hughes, Dana, et al. "Parents' Perceptions and Knowledge of Healthy Families and Medi-Cal." San Francisco: Center for Children's Access to Health Care Institute, Institute for Health Policy Studies, February 2001.
- Kennedy, Peter. "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations." *American Economic Review*, vol. 71, no. 4, 1981, p. 801.
- Long, Peter. "A First Glance at the Children's Health Initiative of Santa Clara County, California." Menlo Park, CA: The Kaiser Commission on Medicaid and the Uninsured, August 2001.
- Perry, Michael, et al. "Barriers to Medi-Cal Enrollment and Ideas for Improving Enrollment: Findings from Eight Focus Groups in California with Parents of Potentially Eligible Children." Menlo Park, CA: The Kaiser Commission on Medicaid and the Uninsured, August 1998.
- Santa Clara Office of Education Center for Educational Planning. "Student Enrollment Data 1999 Through 2002." [<http://www.sccoe.k12.ca.us/programs/asc/docs/ASDData05-13-03.pdf>]
- Wong, Liane. "Background Data and Models for Expanding Health Insurance Coverage to Uninsured Children in Santa Clara County." Washington, DC: Institute for Health Policy Solutions, October 19, 2000.



**APPENDIX A**

**METHODOLOGY FOR ESTIMATING IMPACTS OF  
CHI ON MEDI-CAL AND HEALTHY FAMILIES  
ENROLLMENT**



This appendix details the methods used for estimating the impact of the Children’s Health Initiative (CHI) of Santa Clara County on the enrollment of children into the state Medi-Cal and Healthy Families programs. Section A discusses the data used for the analysis, including the specific data sources that were used, as well as the steps taken to construct analytic files from these data. Section B then details the methods used. This includes details on the sample design, including the formation of the treatment and comparison samples for the analysis, and a discussion of the data constructs and specifics on the estimation approach.

## A. DATA

### 1. Data Sources

The analysis draws on four data sources:

- **Medi-Cal enrollment files** for the period January 1998 through April 2003 provided information for estimating the count of new entries to the Medi-Cal and Healthy Families programs in each zip code over our four-year study period (January 1999-December 2002). Prior to August 2001, these data were available in six-month extractions, every January and July of the current year.<sup>1</sup> Starting in August 2001, these data were available on a monthly basis. Each extraction includes the current month’s enrollees as well as information on the eligibility and recipient status of these individuals over the previous twelve calendar months. They also contained key information on recipients including their zip code, birth date, and race/ethnicity. In addition, a unique identifier was provided for each recipient so that we could link files and identify the start and length of each enrollment spell over the study period.
- **Healthy Families enrollment files** through the period April 2003 provided a source for validating the enrollment information from the Medi-Cal files and for other supporting analyses.<sup>2</sup> These files contained zip code, program eligibility, and demographic information for each recipient similar to the Medi-cal data.

---

<sup>1</sup>Any enrollee whose enrollment spell occurs entirely between these six-month extracts cannot be identified, but their numbers are estimated to be small (less than 5 percent of the total) and to vary little across counties.

<sup>2</sup> In creating the enrollment spells used in the analysis, we used data on Healthy Families enrollees available from the Medi-Cal files because it could be linked with data on Medi-Cal enrollees. Among other benefits, this approach allowed us to identify the month in which a child enrolled in public coverage (rather than a particular program). Analysis of the Healthy Families

- **Census 2000 long form for California** provided detailed information on the population characteristics of each zip code tabulation area (ZCTA) in California. These ZCTAs closely approximate the U.S. Postal Service zip code service areas that form the unit of analysis for the study. Among the characteristics identified from Census and used in the analysis were the population counts and distribution by age, gender, race/ethnicity, income, primary language, and place of birth.
- **Labor statistics data** from U.S. Department of Labor Bureau of Labor Statistics and from California's Employment Development Department of Labor Market Information provided the sources to account for variation in economic conditions across zip codes and counties and over time. Specifically, the former data provided the source to estimate the county level unemployment rate for the period December 1998 through April 2003, which were used in the main analytic models. The latter data provided the source to measure employment counts, by industry and occupation, for each zip code, which were used in several supporting analyses examining the sensitivity of our results to possibly unmeasured changes in economic conditions.

## 2. Constructing the Main Data File

The main individual level data file contains a complete enrollment history for 2,436,332 children who received either Healthy Families or Medi-Cal coverage between January 1998 and April 2003. Each record in the file contains information on a unique recipient, including the recipient's type of coverage (if any) and eligibility type over this 64-month period, his or her location (both zip code and county), and birthday and other key demographic information.

To determine each child's eligibility type for a given month of coverage, we used data from the four eligibility codes. The codes we used are as follows:

1. **Primary eligibility/Aid** code is considered the most accurate of the eligibility codes available from the 12 months of retrospective data included with each Medi-Cal extract. Once a primary eligibility code is assigned by the state, it is retroactively applied back through the recipient's history. In this manner, missing codes are filled in and incorrect codes are overwritten. For this reason, we use the newest possible extract to look back at the primary eligibility for a given month. This also is the

---

*(continued)*

enrollment counts between the two files indicated strong similarity over time and across counties, establishing the credibility of this approach

reason that we limit the analysis to December 2002, allowing the extracts for 2003 to provide the updated primary eligibilities for these earlier months.

2. ***Special eligibility/Aid #1*** code is used in cases where the primary eligibility code remains missing. Special aid eligibility codes are often assigned to individuals who receive multiple forms of public insurance, such as health coverage and food stamps.
3. ***Special eligibility/Aid #2*** code is used if the primary eligibility is missing and special aid #1 eligibility is missing or illegitimate. As with the two codes above, it is revised and updated in each subsequent monthly extract.
4. ***Current eligibility code*** reflects an individual's eligibility from month to month and is not updated retroactively. Thus, the code may contain errors near the beginning of a public insurance spell and is only used in rare cases when the previous three codes are missing or illegitimate.

Once we assigned the code, we classified each enrollee month into one that is either valid or invalid for purposes of our analysis (see Table A.1). As detailed below, these classifications formed the basis for constructing the analysis file for the study. Those months deemed valid include seven distinct eligibility types within Medi-Cal and Healthy Families: (1) Medi-Cal Poverty Related, (2) Medi-Cal Section 1931, Medically needy, (3) Medi-Cal Continuous eligibility, TMC, (4) Medi-Cal TANF, (5) Medi-Cal Other, (6) Healthy Families, and (7) Temporary County Classification.<sup>3</sup> Those months deemed invalid include months with an eligibility code that was missing, unclassified, or reflected coverage that was only partial (such as Emergency Medi-Cal) or was for neither Medi-Cal nor Healthy Families.

---

<sup>3</sup> The 8E code is a temporary label that some counties assign to expedite enrollment when it comes through a single point of entry while processing an application. It serves as a temporary classification while the individual's eligibility is determined.

TABLE A.1  
MONTHLY INSURANCE CLASSIFICATION

Insurance Classification	Eligibility Codes	Aid Codes
<b>Valid</b>		
1. Medi-Cal, poverty-Related	From 0 to 499	72,7A,47,8R,8P
2. Medi-Cal, Non-TANF Section 1931/ Medically Needy	From 0 to 499	3N,34,82
3. Medi-Cal, Continuous Eligibility/TMC	From 0 to 499	7J,39,59,54
4. Medi-Cal, TANF-Related	From 0 to 499	30,35,33,3R,32,3A,3C,3P,3E,3M, 3U,3L,3W,3G,3H,39,38,1,2,8
5. Medi-Cal, SSI/Other	From 0 to 499	All Aid Codes
6. Healthy Families	Any	9H
7. Temporary County Classification	> 500	8E
<b>Invalid</b>		
8. Share-of-cost or non-health insurance programs	Missing or 500 to 998	All Codes
9. Ineligible for public assistance	999	Any or Missing Codes

Source: Med-Cal enrollment files.

### 3. Constructing the Analytic File

From the main file, we constructed an initial analytic file containing 1,241,505 records. Each record reflects a unique spell of coverage by a child in public health insurance coverage during the period December 1998 to March 2003. We define a unique spell of coverage as one of at least two months of continuous coverage with no coverage in the preceding 11 months. Through this latter restriction, we eliminate most cases of cycling within or transferring between coverage, allowing us to focus on coverage spells that have the greatest likelihood of being initiated by outreach activities. In addition, we avoid contaminating our measure of the effects of CHI with unrelated factors associated with program retention or transfer.

To decide whether a spell was to be included in our analysis sample, we reviewed the validity of the eligibility type in the first month (according to Table A.1 above) of the spell and

in each successive month thereafter.<sup>4</sup> The spells we included in the analysis sample began with a valid eligibility type and continue with that type, or another valid type, throughout the majority of the spell. For example, a spell that begins with enrollment under a Medi-Cal poverty eligibility code (for example, Aid Code of 34) and continues with that or any other valid eligibility type is labeled as an entry into the Medi-Cal poverty-related category.

Once we identified all the spells to be included in the analysis sample, we arrived at an interim analytic file that included 1,170,530 unique spells of coverage in either Medi-Cal or Healthy Families. As our next step, we classified these spells into five distinct eligibility types, based on their likelihood of being affected by the CHI intervention. They include:

1. Healthy Families
2. Medi-Cal, Poverty-related categories
3. Medi-Cal, Medically Needy/Section 1931/Continuous eligibility/TMC
4. Medi-Cal, TANF-Related
5. Medi-Cal Other/SSI (the least likely group to be impacted by CHI)

The final step in creating the analytic file was to aggregate the spell-level records to create a file containing the monthly counts of new entries, by type, for each zip code in the state. After dropping spells that began prior to 1999 (for which we lack a sufficient look-back) and spells after 2002 (for which we lack a sufficient time period to ensure that the codes have been updated), we arrive at a file containing 48 monthly observations (months) for each zip code. Each of these observations contains the sum of new Medi-Cal or Healthy Families entries, broken down by eligibility type, for that location and month. Thus, for example, for each zip

---

<sup>4</sup>Spells that begin with aid category 8E (see above) are an exception to this system. Because this initial value occurs in the month the spell began, it is still considered the starting point, but we look forward to the first non-8E code to classify each spell. For the vast majority of spells, the eligibility type remained constant in each month.

code in Santa Clara County, the file contains 48 observations, each containing variables measuring the count of children enrolling in each of the four Medi-Cal classifications (poverty-related, Medically Needy/Section 1931/Continuous eligibility, TANF, or SSI/other) and Healthy Families.

## **B. METHODS**

### **1. Sample Design**

To form treatment and comparison areas, we began with the 2000 the Census data file for California in which we had identified 1,664 zip codes in the state with valid Census data.<sup>5</sup> Of these 1,664 zip codes, 58 were located in Santa Clara County,<sup>6</sup> while the remaining 1,606 were located elsewhere in the state. For each of these zip codes, we then created a record containing key demographic data from the 2000 Census long form data set. Also added to each record was the count of children enrolled in the Medi-Cal and Healthy Families programs in each zip code as of July 2000. (This date was chosen to correspond roughly to the date of the Census.)<sup>7</sup> Records were then combined to form a master file for constructing the comparison area for the study.

**Defining the Treatment Group.** Once the data file had been constructed, we turned to defining the set of Santa Clara zip codes that would form the “treatment group” for our analysis.

---

<sup>5</sup>Census refers to these regions as five-digit ZCTAs. Each of these 1,664 ZCTAs was verified as a valid postal code in the state by matching it with data from the 2003 California pro-zip code file, obtained from GreatData.com.

<sup>6</sup> One Santa Clara County zip code (94086) split into two zip codes in July 2001 (94086 and 94085). We treated this as a single zip code for the analysis, combining both the enrollment data and the census information to form a single observation. See below for more information on how we addressed issues with individual zip codes.

<sup>7</sup>The count of Medi-Cal enrollees included only enrollees age 18 and under with full benefits coverage, which correspond to eligibility status codes between “001” and ”499.” See Researcher’s Guide to Medi-Cal (December 2002) for more information.

In total, there are 58 zip codes in the county with at least one enrollee in the Medi-Cal or Healthy Families program as of July 2000 (roughly the date to which our Census data correspond).<sup>8</sup> Of these 58 zip codes, we identified 38 that could be treated as individual observations for our treatment sample because they had at least 200 children enrolled in Medi-Cal and Healthy Families as of July 2000, making them suitable for retention in the analysis as individual observations.<sup>9</sup> For the remaining 20 “minor” zip codes, we conducted a case-by-case review to decide how they should be handled for the analysis. This process led to the following decisions:

- One minor zip code (95020) was added as an individual observation despite having enrollment of only 137 children in July 2000. This zip code had relatively unique demographic characteristics, making it a poor candidate to merge with other zip codes. However, it also had a non-trivial number of enrollees so that we did not want to remove it from the analysis.
- Nine minor zip codes were combined to form three additional observations—Los Gatos (95030, 95032, 95033); Saratoga/San Jose (95120 and 95070); and Palo Alto/Los Altos (94022, 94024, 94301, and 94306). For each observation, the combined zip codes were either contiguous or near contiguous and had very similar demographic characteristics. Enrollment in the three observations was still relatively small, from 246 to 150 in July 2000.

---

<sup>8</sup>Three Santa Clara County zip codes (94550, 95023, and 94303) cross into neighboring counties. In each zip code, the state Medi-Cal and Healthy Families files attribute all the enrollment to the neighboring county; however, enrollment data from Santa Clara social services office shows a fraction of these enrollees do reside in Santa Clara County and so should be considered for the analysis. In each zip code, the fraction is small—about 6 percent in 94303 and less than 2 percent in 94550 and 95023 (as of March 2003). Applying this ratio to the overall enrollment count in these zip codes for July 2000, we estimate that roughly 162 children were enrolled in 94303 in Santa Clara and about 11 and 2 were enrolled in 94550 and 95023 in Santa Clara County.

<sup>9</sup>The cutoff of 200 enrollees is arbitrary but is chosen because it is believed sufficiently large that CHI might have policy-relevant effects. In addition, as discussed below, each Santa Clara County zip codes is weighted in the analysis based on its share of children eligible for public coverage in the county. Zip codes with relatively few enrollees are therefore likely to take on relatively less importance in estimating the impacts of CHI, both in terms of the magnitude of any impacts and the associated precision.

- Four minor zip codes (94041, 95134, 95135, and 95139) were merged with larger, neighboring zip codes.<sup>10</sup> In each case, the minor zip code had demographic characteristics quite similar to those of the larger zip code.
- Six minor zip codes (94550, 94304, 94305, 95023, 95113, 95140) had enrollment of 2 to 25 children in July 2000, making them of negligible value for measuring impacts of CHI. In addition, their demographic characteristics matched poorly with other zip codes in the county, making them weak candidates for combining in some way. These zip codes were therefore dropped from the analysis.

The resulting treatment group includes 42 “members,” each comprising one or more zip codes in the county. For the main unit of analysis, quarterly enrollment, we have 16 quarters of observed data for each treatment group member; half of which are pre-CHI (from first quarter 1999 through the fourth quarter of 2000) and half of which are post-CHI (from the first quarter of 2001 through the fourth quarter of 2002). The total treatment group for the analysis therefore includes 672 observations (42 members × 16 quarters) and accounts for more than 99.8 percent of all the enrollment taking place in Santa Clara County over this period.

**Defining the Comparison Group.** To form our comparison group for the analysis, we matched each of the 42 treatment members (each comprising one or more Santa Clara zip codes) with at least 10 comparison zip codes elsewhere in the state. This matching was based on an iterative approach that, for each treatment group member, selected zip codes elsewhere in the state that were most similar in terms of their demographic characteristics and their pre-CHI enrollment levels. As described below, this iterative approach allowed us to quantify the relative quality of each comparison group member, creating opportunities to conduct various tests of robustness and to assess the quality of the overall comparison group.

---

<sup>10</sup>Minor zip code 95134 (San Jose) was merged with bordering zip code 95054; minor zip code 94041 (Mountain View) was merged with bordering zip code 94040; and minor zip codes 95135 and 95139 (San Jose) were merged with zip code 95138, which lies between the two.

As the first step in the matching process, we used Census data to construct seven continuous variables measuring key demographic characteristics of zip codes throughout the state. These variables included: (1) percent Hispanic ethnicity, (2) percent white, (3) percent Asian, (4) percent of children living in a household under 200 percent of FPL, (5) percent of the population under 100 percent of FPL, (6) percent non-citizen, and (7) the zip code’s population density expressed as its percent rank in the State of California. In addition, we developed two measures of the share of children who are eligible for public coverage but not enrolled: (1) the ratio of Medi-Cal enrollees in July 2000 to the number of children (aged 17 or younger) living in a household under 200 percent of FPL, and (2) the ratio of Healthy Families enrollees in July 2000 to the number of children living in a household under 200 percent of FPL.<sup>11</sup>

As the second step, we compared these nine variables over all of the zip codes outside Santa Clara County, searching for those with characteristics most similar to a given treatment group member.<sup>12</sup> At the first iteration, the “tolerance level” for matching to a treatment member was set to zero, meaning that only zip codes that matched exactly to the sample member could be selected. At each successive iteration, the tolerance level was then expanded slightly for each variable (see Table A.2), creating an opportunity to match additional zip codes to a given

---

<sup>11</sup>These ratios are based on a combination of census data and enrollment data as of July 2000. While they are not believed accurate as absolute measures of the share eligible for public coverage, they should provide a good proxy for comparing relative shares of eligibles across different zip codes.

<sup>12</sup>For treatment observations with large enrollments—1,000 children or more—we limited the pool of potential comparison zip codes to those with 500 children or more. For observations with medium enrollments—between 500 and 1,000 children—we limited the pool of potential comparison zip codes to those containing 200 to 3,000 children. For observations with small enrollments – fewer than 500 children – we limited the pool of potential zip codes to between 50 to 1,500 children. This step was not critical for the analysis, since all impacts were measured as percentage changes; however, it did improve the face validity of the matches identified and ensured that treatment observations with large enrollments were not matched with comparison observations that might be highly volatile because they had relatively few enrollees.

TABLE A.2

ACCEPTABLE MAXIMUM DIFFERENCES

Matching Characteristics	START: Run 1	Run 10	Run 15	Run 20	Rate of Increase per Iteration (+/-)
Medi-Cal Enrollment by Number Eligible	0.00	0.09	0.14	0.19	0.01
Healthy Families Enrollment by Number Eligible	0.00	0.09	0.14	0.19	0.01
Percent of Population Below 100 Percent of FPL	0.00	0.18	0.28	0.38	0.02
Percent of Kids Below 200 Percent of FPL	0.00	0.09	0.14	0.19	0.01
Percent Rank of Population Density	0.00	0.36	0.56	0.76	0.04
Percent Hispanic Ethnicity	0.00	0.09	0.14	0.19	0.01
Percent White	0.00	0.09	0.14	0.19	0.01
Percent Asian	0.00	0.09	0.14	0.19	0.01
Percent Non-Citizen	0.00	0.09	0.14	0.19	0.01

treatment member. If, at the end of each iteration, we had matched fewer than 10 zip codes for a given sample member, we expanded the tolerance level and iterated an additional time until we reached at least 10 matches. This matching was done with replacement--zip codes from our potential comparison pool could be selected as comparison areas for more than one treatment member.

In total, the matching process led to the selection of 512 comparison group members (that is, zip codes external to Santa Clara County), at least 10 for each member of the treatment group.<sup>13</sup> This group consists of 282 unique zip codes, meaning that each comparison zip code matched an average of 1.83 treatment group members. As seen in Table A.3, these zip codes are drawn from a total of 28 counties in the state. Counties with the largest shares of comparison zip codes include Los Angeles (38%), Alameda (13 percent and Orange (12 percent).

A comparison of the matching variables and selected other variables between the Santa Clara sample and the comparison sample suggests good similarity (Table A.4). For the full sample (column 1), the rates of coverage prior to CHI (measured as a share of the children under 200 percent of FPL) are comparable for both Healthy Families and Medi-Cal. Moreover, when combining the rates for the two programs, we see that they are nearly identical (both about 63 percent). The percentage of the population in poverty is also similar, although there is a larger difference in the share of children who are under 200 percent of FPL (about 30 percent for Santa Clara County and 35 percent for the comparison area). Santa Clara's share of Hispanic children is also similar to the comparison area (36 versus 35 percent), although it has a higher share of

---

<sup>13</sup> The matching process initially led to 517 comparison group members. Five of these areas (four unique zip codes) underwent changes in their structure (for example, the zip code split into multiple zip codes or they did not exist over the entire life of our analysis period) and were therefore dropped from the sample.

TABLE A.3  
MATCHED ZIP CODES, BY COUNTY

County	Unique Comparison Zip Codess	Total Comparison Observations
Los Angeles	103	192
Alameda	22	64
Orange	34	61
San Mateo	15	34
Contra Costa	17	30
San Diego	11	19
San Francisco	7	18
Ventura	12	16
San Bernardino	13	16
Sacramento	6	14
Monterey	6	8
Riverside	6	7
San Joaquin	3	5
Solano	5	5
Imperial	3	3
Stanislaus	2	3
Fresno	2	2
Marin	2	2
Merced	2	2
Sonoma	2	2
Yolo	2	2
Colusa	1	1
El Dorado	1	1
Placer	1	1
Madera	1	1
Napa	1	1
Santa Barbara	1	1
San Benito	1	1
<b>Total</b>	<b>282</b>	<b>512</b>

Asian children (24 percent versus 19 percent) and a higher share of foreign-born non-citizens (24 percent versus 18 percent). The latter statistics are most likely related.

These comparisons improve as the sample is refined to include only the better matches among the comparison sample (columns 2 through 4 of Table A.4).<sup>14</sup> For example, column 4 compares those zip codes in the Santa Clara sample that had at least one A-level match to the 154 A-level comparison zip codes. Examining those results, we see the mean percent of children (0-17) in households with income under 200 percent of FPL in the comparison sample is now within 2 percentage points of the Santa Clara mean (from a 5 percent difference in the full sample). In addition, we see improvement in the difference in means for each of the race/ethnicity characteristics, since each is within 1 percentage point when the samples are limited to our best matches. We also find that the difference between the two groups in the mean population density is reduced to less than 150 persons per square land mile (from over 800) in addition to the means of both samples declining to approximately 4,700 people per square land mile (out of approximately 6,500). Population density declines because the sample loses some of the largest urban zip code areas that were hardest to match. Findings presented in the report proved quite robust to comparison groups that were restricted to these higher-level comparison members only (see the main text for details).

---

<sup>14</sup>In order to examine the sensitivity of our results to the quality of the comparison group, we classified each member into four groups, from A through D, depending on the number of iterations required before it matched to a given treatment member. More than half (26) of the 42 treatment members had at least one A-level comparison group member, which included comparison members that matched within nine iterations. All but three (39) had at least one A- or B-level comparison group members, which included comparison members that matched within 14 iterations.

TABLE A.4

SELECTED CHARACTERISTICS OF SANTA CLARA COUNTY  
AND THE MATCHED COMPARISON AREA

Characteristics	(1) Full Comparison Area		(2) Limited to A-C Level Matches <sup>a</sup>		(3) Limited to A-B Level Matches <sup>a</sup>		(4) Limited to A-Level Matches <sup>a</sup>	
	Santa Clara Area (Mean)	Comparison Area (Mean)	Santa Clara Area (Mean)	Comparison Area (Mean)	Santa Clara Area (Mean)	Comparison Area (Mean)	Santa Clara Area (Mean)	Comparison Area (Mean)
	<b>Variables Used in Matching</b>							
Percent of Eligible Children (Under 200% of FPL) Enrolled in MC	54.7	52.5	54.7	53.0	51.3	51.3	49.0	49.7
Percent of Eligible Children (Under 200% of FPL) Enrolled in HF	8.2	10.8	8.2	10.9	7.8	9.9	7.4	9.5
Percent of Population with Income Below FPL	10.0	10.9	10.0	10.6	9.7	10.4	9.0	9.2
<i>Percent of Children in Household with Income Under 200% of FPL</i>	29.7	34.2	29.7	33.2	29.2	32.4	26.6	29.0
<i>Population Density<sup>b</sup></i>	6,964	6,114	6,964	6,176	6,521	5,610	4,807	4,677
<i>Percent Hispanic</i>	35.8	34.7	35.8	33.6	34.0	33.7	29.5	28.5
Percent White	33.8	37.2	33.8	36.9	36.6	39.0	45.0	45.9
Percent Asian	23.8	18.8	23.8	20.6	22.7	18.0	18.5	17.3
<i>Percent Foreign-Born, to a U.S. Citizen</i>	23.6	17.8	23.6	17.9	22.7	16.2	18.2	16.6
<b>Additional Variables</b>								
<i>Unemployment Rate</i>	3.2	4.9	3.2	4.8	3.2	4.8	3.2	4.7
<b>N(Number of Zip Codes)</b>	<b>42</b>	<b>512</b>	<b>42</b>	<b>455</b>	<b>39</b>	<b>369</b>	<b>26</b>	<b>154</b>

Source: Census 2000, enrollment data from Healthy Families and Medi-Cal, and unemployment data from U.S. Department of Labor Bureau of Labor Statistics.

Notes: Variables in italics were used in regression. Statistics are computed as the mean value for the zip codes in Santa Clara county and the zip codes that comprise the matched comparison area used to estimate impacts. Some zip codes in Santa Clara county with very small numbers of children enrolled in public coverage have been combined. See Methods Appendix for additional information.

<sup>a</sup>Level of match refers to the degree of consistency between the Santa Clara zip code and its matched-comparison zip codes across all of the demographic characteristics used to form the comparison area. For A-level matches, there was the highest degree of consistency across all characteristics; for other matches, there was relatively less consistency across at least one characteristic.

<sup>b</sup>total population / number of square land miles

## 2. Variable Construction

**Dependent Variables.** The dependent variables listed in Table A.5 were constructed from the analytic file described above. Our main dependent variable measures the number of new entries to public insurance (Medi-Cal or Healthy Families) in each zip code in our sample in each quarter from 1999 through 2002. We define a new entry as a child who enters one of these two programs, remains enrolled for two or more months, and did not have coverage under one of these two programs in the preceding 11 months. This definition excludes entry to these programs by children who transfer between the two programs or who cycle on and off over a short time period—both events that are likely unrelated to the CHI intervention. Including these children in our estimates is potentially misleading, for a number of reasons. For example, counties experiencing poor program retention may generate a larger pool of these children, leading to higher levels of cycling.

In addition to our main dependent variable, which measures entry into either Medi-Cal or Healthy Families, we created and analyzed variables that measured entry into these two programs individually. Within Medi-Cal, we also constructed variables measuring entry under four distinct eligibility categories: (1) poverty-related; (2) Medically Needy, Section 1931, Continuous eligibility; (3) TANF-related; and (4) SSI and other eligibility types. In general, the first group (the poverty-related group) is expected to be the most likely to be affected by CHI, while the last group is expected to be the least likely. Analyzing the individual eligibility types thus provides a more targeted measure of the impacts of the program, as well as a useful test of the credibility of the results.

**Explanatory (Control) Variables.** To account for potential differences between Santa Clara County and the comparison area that may not have been controlled for in our selection of the comparison areas, we construct several measures of demographic characteristics: (1) percent

TABLE A.5  
DEPENDENT VARIABLES

Spells of Public Insurance Classifications	Definition
<b>Overall Impacts</b>	
1. Healthy Families + Medi-Cal	Any new entry into either Medi-Cal or Healthy Families
2. Medi-Cal	Any new entry into a valid Medi-Cal insurance category: Poverty-related, Medically Needy, TANF, or other/SSI/disabled
3. Healthy Families	New entry into Healthy Families program
<b>Separate Impacts</b>	
4. Medi-Cal, Poverty-Related Categories	New entry in a valid Medi-Cal Poverty-related category
5. Medi-Cal, Section 1931/Medically Needy, Continuous Eligibility, TMC <sup>a</sup>	New entry in a valid Medi-Cal Section 1931-Medically Needy/Continuous eligibility/TMC category
6. Medi-Cal, TANF-Related	New entry in a valid Medi-Cal TANF-related category
7. Medi-Cal, SSI/Other	New entry in a valid “other” Medi-Cal Category, which includes SSI or disabled children

<sup>a</sup>Section 1931 requires Medi-Cal to be provided to low-income families who meet the requirements of the Aid to Families with Dependent Children (AFDC) State plan in effect July 16, 1996. TMC = Transitional Medi-Cal.

Hispanic ethnicity, (2) number of children (under the age of 18) living in a household under 200 percent of FPL, (3) percent non-citizen, and (4) the zip code's population density expressed as its percent rank in the State of California.<sup>15</sup> In addition, we used county-level, monthly unemployment rates from the BLS Local Area Unemployment Statistics (LAUS) database and a series of indicator variables measuring the quarter of the period starting in January 1999 and ending in December 2002. These variables control for local area trends in economic factors and seasonality in Healthy Families and Medi-Cal enrollment, which potentially could differ across zip codes and counties. (For the distribution of these control variables, as well as other selected variables used in the matching process, see Table A.4.)

### **3. Estimation**

Our approach to estimating the effects of CHI on enrollment is to use a difference-of-difference model. With this model, the effect of CHI is measured as the difference in the *change* in enrollment after CHI between Santa Clara County and the comparison area. This model is particularly attractive when adopting an external comparison group design because it controls for *all* differences between Santa Clara County and the comparison area that are approximately fixed or stable over time.

---

<sup>15</sup> Other variables that were included in the matching process, such as the racial distribution of the zip code and the percentage of the population in poverty, were dropped from the regression due to concerns of multicollinearity and very limited gains in the explanatory power of the model. Impact estimates varied little with or without the use of the control variables.

**Base Specification.** Our base model is specified as follows:

$$LN(NEWENTRIES)_{zt} = \alpha + SCZIP_z \beta_1 + \sum_{t=2}^{16} QRTR_t \phi_t + POSTCHI_t \beta_2 + (SCZIP_z * POSTCHI_t) * \left( \sum_{t=2}^{16} QRTR_t \right) \beta_{3t} + ZIPCHARS_z \beta_4 + CTYCHARS_{zt} \beta_5 + \sum_{t=2}^{16} QRTR_t \phi_t + \varepsilon_i$$

where:

$\alpha$  = an intercept term.

$NEWENTRIES_{zt}$  is a count variable measuring the total number of entries into Medi-Cal and Healthy Families in zip code  $z$  in quarter  $t$  (over 16 quarters, from first quarter 1999 through fourth quarter 2002). Additional specifications look at the programs separately. The variable is specified in natural log form to account for skewness in the distribution of new entries across zip codes in the sample.

$SCZIP_z$  is an indicator variable for whether zip code  $z$  is a Santa Clara County zip code. It equals 1 for all Santa Clara zip codes, and 0 for all the other (comparison) zip codes.

$POSTCHI_t$  is a post-CHI implementation indicator, which will be equal to 1 if time period (quarter)  $t$  is after CHI was implemented (i.e., from first quarter 2001 through fourth quarter 2002), and equal to zero otherwise.

$SCZIP_z * POSTCHI_t$  is an interaction term between the indicators for SCZIP and CHI. The variable equals 1 if the zip code is in Santa Clara County and time period  $t$  is after the adoption of CHI; otherwise it is 0.

$ZIPCHARS_z$  is a series of variables measuring the demographic characteristics of zip code  $z$  based on data from Census 2000.

$CNTYCHARS_t$  is the county-level unemployment rate for quarter  $t$ . Although it is time-varying, the rate is constant for all zip codes located within the same county.

$QRTR_t$  = binary indicators of calendar quarter  $t$ . These variables will control for the time trend in enrollment over the study period. The omitted (reference) category is quarter one (Jan-Mar 1999).

$\varepsilon_i$  = a random error term.

The specification is estimated as a log-linear model using weighted least squares (see the information below on weighting). SUDAAN software was used to account for the effects of this weighting and for the non-independence of the sample due to repeated observations of the zip codes over time and to clustering of the zip codes within counties.

The impact estimate for each quarter after the start of CHI is given by the coefficients  $\beta_{3t}$  (for t from 9 to 16). Each coefficient is exponentiated, following the transformation recommended by Kennedy (1981), to yield a measure of the percentage change in new entries in each quarter after adoption of CHI.<sup>16</sup> In findings presented in the report, the count of new entries under the treatment (with CHI) and the counterfactual (in absence of CHI) are based, respectively, on the actual count of new entries during the quarter of interest and the count dividing through by the estimated percentage impact. So, for example, if the impact of CHI is 10 percent in a given quarter and the number of new entries observed in the quarter is 11,000, the base number (in absence of CHI) is assumed to be 11,000/1.10, or 10,000.<sup>17</sup> The total estimated impact over the full post-CHI period (2001-2002) is given by the sum of the impacts for each individual quarter.

**Weighting.** Because zip codes in Santa Clara County vary substantially in the number of children who are eligible for public insurance (Figure A.1), our impact estimate should not be based on a simple average of the effect of CHI across each of the zip codes in the county.

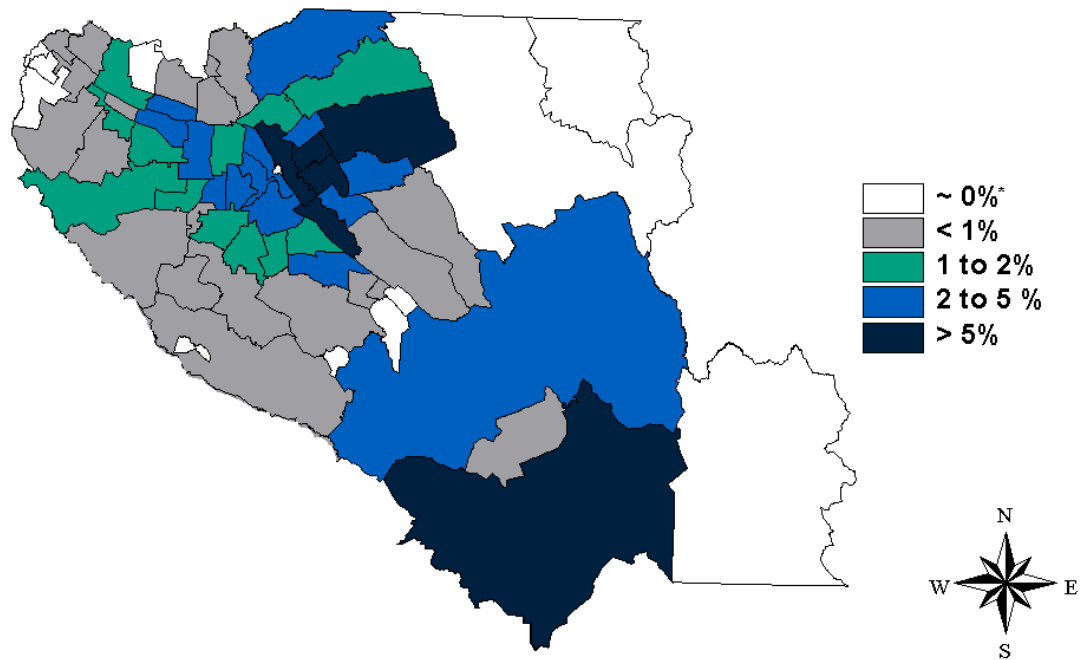
---

<sup>16</sup>The transformation by Kennedy (1981) is used to account for the potential bias that can result when simply exponentiating dummy variables given log-linear models. Results with this transformation were only slightly different from those derived by simply exponentiating the coefficient, suggesting that the extent of any bias is trivial.

<sup>17</sup> Alternative counts were derived using predicted values from the regression and applying the smearing estimator developed by Duan (1983). These counts were very similar. The approach above was chosen because it provides the actual numbers seen in the county.

FIGURE A.1

DISTRIBUTION OF CHILDREN BELOW 200 PERCENT OF THE FEDERAL POVERTY LINE IN SANTA CLARA COUNTY



Source: Census 2000 and enrollment data from Healthy Families and Medi-Cal.

\*Those zip codes in white include such small shares of low-income children that they were dropped from the analysis sample. The remaining zip codes account for more than 99.8 percent of the children under 200 percent of the Federal Poverty Line in the county.

Rather, it should be based on a weighted average, whereby the impact of CHI in each zip code in the county is weighted by the proportion of children who are eligible for coverage in each zip code. By doing so, we assign less importance to impacts in those zip codes with relatively few eligible children and greater importance to impacts in zip codes with relatively more eligible children.

Lacking a direct measure of the proportion of eligibles in each zip code in the county, we use instead the proportion of children living in a household under 200 percent of the FPL (from the 2000 Census). Thus, zip codes containing a large fraction of these children relative to the total for the county take on greater weight, while those containing a small fraction take on less. While this proxy measure will likely offer a poor estimate of the *count* of eligibles across the zip codes in the county, it can be expected to perform well as a measure of the *share* of eligibles, since there is little reason to believe that the rates of existing coverage among low-income children in the county (prior to CHI) vary dramatically across zip codes. Results from our regression models support this assumption. In every specification that we estimated, we consistently find that a 1 percent increase in this measure is associated with a 1 percent increase in enrollment, which is exactly what we would expect to find from a direct measure of the number of eligibles. This measure thus appears to proxy quite well for the actual distribution of eligible children across the county.

Weighting of the comparison sample follows directly from the weights assigned to Santa Clara zip codes. Specifically, for a given Santa Clara County zip code (“treatment group member”), each of its comparison zip codes is given a weight equal to the treatment group member’s weight, divided by the number of comparison zip codes for that member. For example, for a Santa Clara zip code with a weight of 0.15 and 10 matched comparison zip codes, each of the 10 comparison zip codes receives a weight of  $0.15/10$  or 0.015. The result is that

each set of matched comparison areas for a given Santa Clara zip code sums to exactly the same weight (that is, receives the same degree of importance) as the Santa Clara zip codes itself. And, in total, the weighted sample size of the treatment group sample and comparison group sample both sum to the same value.<sup>18</sup>

**Sensitivity Testing.** Several alternative specifications of the base model shown above were estimated to explore the distribution and robustness of the overall impacts. Among the main tests conducted and reported are:

1. To assess how impacts varied across the county, we estimated the model above for each of the individual Santa Clara zip codes and its corresponding matched comparison zip codes.
2. To assess the sensitivity of results to the comparison area definition, we reestimated the model above, constraining the sample to an increasingly better set of matches.
3. To assess the sensitivity of results to employment and economic conditions that might be unaccounted for, we conducted numerous respecifications. Among those presented in the report: we reestimated the model without the unemployment rate in the model; we explored alternative definitions of this variable, including one and two-quarter lags; and we replaced the unemployment variable with a measure of the count of employed people in each zip code over time. Because this count variable is based on the employment activity of businesses in the zip code (rather than the residents in the zip code), it is a less attractive measure for assessing the effects of CHI despite the fact that it is available at the zip code level. However, it is quite useful for sensitivity analysis.
4. We reestimated this model without control variables, yielding a measure based entirely on the actual counts of new entries in the sample (without regression adjustment).

As described in the main text, each of these models displayed results supportive of those found from the main specification. The results from these tests thus underscore the credibility of the findings.

---

<sup>18</sup> The sum of the weights for the treatment or comparison group members is 1.00, since this is simply the sum of the shares of eligibles across the county.