

CA-033-001 +

CA-033-002

GARYFILE Technical  
Documentation Series  
No. 1

USERS GUIDE FOR THE  
GARY INCOME MAINTENANCE EXPERIMENT  
PUBLIC USE FILE--GARYFILE

*Data & Program Library Service  
8313 Social Science Bldg.  
University of Wisconsin-Madison  
Madison, WI 53706*

July 6, 1979

MATHEMATICA POLICY RESEARCH, INC.  
2101 L St., N.W., Washington, D.C.

The research reported herein was performed pursuant to contract number HEW-100-78-0059 with the U.S. Department of Health, Education, and Welfare, Washington, D.C. The opinions expressed herein are those of the authors and should not be construed as representing the opinions or policies of the United States Government.

## CONTENTS

	Page
LIST OF TABLES . . . . .	ii
LIST OF FIGURES . . . . .	iii
I. INTRODUCTION . . . . .	1
The Income Maintenance Experiments . . . . .	1
II. GARY EXPERIMENT DESIGN AND ADMINISTRATION . . .	6
The Experimental Design . . . . .	6
Administration of the Experiment . . . . .	12
III. DATA OBTAINED FROM THE GARY INCOME MAINTENANCE EXPERIMENT . . . . .	25
Interview Data . . . . .	25
Income Report Form Data . . . . .	17
Administrative Records Data . . . . .	28
Creating a Public Use File . . . . .	29
IV. THE GARY INCOME MAINTENANCE EXPERIMENT PUBLIC USE FILE . . . . .	31
Format and Contents . . . . .	31
Documentation . . . . .	40
Uses of the Data . . . . .	46
V. OBTAINING THE GARYFILE . . . . .	54
APPENDIX A: MATHEMATICA POLICY RESEARCH PUBLICATIONS ON THE GARY INCOME MAINTENANCE EXPERIMENT . . . . .	A-1
APPENDIX B: SOCIOECONOMIC AND SOCIAL-PSYCHOLOGICAL MODULES . . . . .	A-4
APPENDIX C: VARIABLE LIST FOR GARYFILE . . . . .	A-10

TABLES

	Page
I.1 Characteristics of the New Jersey, Gary, Seattle, and Denver Income Maintenance Experiments and Available Public Use Data . . . . .	3
II.1 Gary Income Maintenance Experiment Payment Schedule, by Family Size . . . . .	8
II.2 Family Payment Schedule for NIT Plan with \$3,300 Guarantee and 40-Percent Tax Rate, by Earnings . . . .	10
II.3 Gary Experiment Sample Design: Families Enrolled by Negative Income Tax Treatments and Social Services/ Day Care Treatment . . . . .	14
II.4 Sample Attrition . . . . .	16
IV.1 Gary Income Variables by Subject and Placement on the File . . . . .	47

## SECTION I

### INTRODUCTION

A thousand black families in Gary, Indiana, were guaranteed a specified minimum income for three years as participants in the Gary Income Maintenance Experiment (GIME). Another 800 families served as controls, receiving only token payments, but providing the same kinds of detailed data as did participants in the financial experiment to permit evaluation of the experiment's effects on a wide range of economic and social behavior. This field test of a negative income tax (NIT) program was conducted between 1971 and 1974 by Indiana University under a contract with the U.S. Department of Health, Education, and Welfare, and the Indiana State Department of Public Health.

The Department of Health, Education, and Welfare (HEW) subsequently awarded a contract to Mathematica Policy Research, Inc. (MPR) to carry out a large number of research analyses of the Gary data. The studies focus on experimental effects with regard to employment, income, and family structure patterns of families who received payments compared to those who acted as controls, but cover diverse other topics as well. In addition, MPR's contract included a major data processing effort to build a public use data file that will enable other researchers to use the wealth of data collected in the experiment.

At the heart of the Gary data access system is a set of documentation that describes the full range of available data. This document, report number 1 of the set, represents a users' guide to the Gary experiment. Following this introduction, which briefly places the Gary experiment in the context of the series of income maintenance experiments sponsored by the Federal government, is a detailed description of the Gary Income Maintenance Experiment's design and administration, a description of the subsequent research analyses and files created to perform them, a description of the final public use file and the extensive documentation which accompanies it, and, finally, directions on obtaining a copy of the Gary Income Maintenance Experiment Public Use File (GARYFILE). This users' guide is intended to be a comprehensive, easy-to-use resource for persons approaching the data for the first time, and a convenient reference for persons already acquainted with the Gary experiment.

### THE INCOME MAINTENANCE EXPERIMENTS

The Federal government has sponsored a series of income maintenance experiments over the past decade, designed to provide policy makers with better information about the effects of alternative income support plans as possible replacements for the current welfare system. The initial test, the New Jersey Graduated Work Incentive Experiment, was carried out from 1968 to 1972 in four urban areas: Trenton,

Paterson-Passaic, and Jersey City, New Jersey, and Scranton, Pennsylvania. The sample included approximately 1,200 low-income white, black, and Spanish-surname families headed by working-age males.

To permit comparison of the New Jersey results with a rural population, the Rural Income Maintenance Experiment was conducted from 1969 to 1973 in selected rural counties of Iowa and North Carolina, representative of the Midwest and South farming regions. The North Carolina sample contained both black and white families, while the Iowa sample was all white. Three-quarters of the total number of 800 families enrolled were headed by prime-age males, as in New Jersey, but some female-headed and aged families were also enrolled.

Planning began in 1969 for the Seattle and Denver Income Maintenance Experiments (SIME/DIME) to study the effects of longer-range guarantees, and in 1970 for the Gary Income Maintenance Experiment to study the different effects of various NIT plans on single-headed versus two-parent families. Each of these locations was carefully selected to add dimensions to the findings of the original urban and rural experiments.

Seattle was chosen as an experimental site primarily because of the unique structure of its labor market. The dominance of one particular industry--aerospace--offered an opportunity to study adjustments made to the business cycle, particularly adjustments by lower-income families to periods of economic distress. In contrast, Denver was chosen as a site because its unemployment rate remained at or below the national average, it experienced a high rate of growth in total employment, and because of its geographic similarity to Seattle. In addition, Denver has a significant chicoano population. The Seattle sample included 2,000 low-income black and white families, of whom two-thirds were headed by prime-age males and the rest were female-headed; the Denver sample of 2,700 families was similar, with the addition of a sizeable number of chicanos. Three-quarters of the SIME/DIME families were enrolled for three years of payments, with most of the remainder enrolled for five years. A small group in Denver enrolled for twenty years, to determine the effects of a longer-lasting income support plan, but this effort was aborted and the families disenrolled on the schedule of five-year families.

Finally, Gary, Indiana, was chosen as an experimental site because its primarily black urban population provided a unique opportunity to study the effects of alternative NIT plans on a sample composed principally of low-income female-headed families. Sixty percent of the all-black Gary sample were single-parent families. As in the urban and rural experiments, all of the Gary experiment families were enrolled for three years. Table I-1 summarizes the basic characteristics of the samples in each site, including Gary where the income maintenance concept was tested.

TABLE I-1

CHARACTERISTICS OF THE NEW JERSEY, GARY, SEATTLE, AND DENVER  
INCOME MAINTENANCE EXPERIMENTS AND AVAILABLE  
PUBLIC USE DATA

	New Jersey	Gary	Seattle	Denver
Sample Size: Original Families	1,374	1,780	2,042	2,758
Percent of Total:				
By Family Type				
Husband-wife	100%	41%	61%	61%
Female head	0	59	39	39
By Race of Head				
White	32	0	44	35
Black	42	100	56	34
Chicano	26	0	0	31
By Income Level				
Below 150% of poverty	100	67	53	51
Above 150%	0	33	47	49
By Experimental Status				
Control	47	43	46	41
Financial	53	57	54	59
Financial Treatment Plans (guarantee as percent of poverty tax rate)	8 plans from 50-30 through 125-50	4 plans from 89-40 through 116-60	9 plans from 100- 50 through 150-80	9 plans from 100-50 through 150-80
Time Period of Payments <sup>a</sup>	3 years from enrollment	3 years from enrollment or thru August 1974	3 years (75% of sample) 5 years (25%)	3 years (75% of sample) 5 years (25%)
Other Treatments Offered	---	Social Services and Day Care	Job Counseling and Training	Job Counseling and Training
Public Use Files				
Current	—	GARY FILE: 48-month family-person file with additional pre- experiment and ad- ministrative <sup>c</sup> data	48-month family- person file <sup>b</sup> 48-month principal person file <sup>b</sup>	48-month family- person file <sup>b</sup> 48-month principal person file <sup>b</sup>
Forthcoming	Cross-site file	Cross-site file	72-month family- person file 72-month principal person file Cross-site	72-month family- person file 72-month principal person file Cross-site
Contact for further information:	Mathematica Policy Research, Inc. 2101 L St., NW Suite 416 Washington, DC 20037 Attn: Miriam Aiken (202/833-9510)			

SOURCES: Institute for Research on Poverty Data Center, User's Manual for the New Jersey Income Maintenance Experiment (Madison: University of Wisconsin, 1975; Kenneth C. Kehrer, et al., The Gary Income Maintenance Experiment: Design, Administration, and Data Files (Gary: Indiana University Northwest, August 1975); and Robert C. Spiegelman, "The Design of Social Experiments with Principal Reference to the Seattle/Denver Income Maintenance Experiments," paper presented at SIME/DIME Conference, Orcas Island, Washington, May 14-17, 1978.

<sup>a</sup>During SIME/DIME, a 20-year financial sample was added, taken from control terminating 3-year families, but aborted so the families disenrolled on the schedule of 5-year families.

<sup>b</sup>Family-person files are organized by family; principal-person files are organized around each person ever a family head.

<sup>c</sup>Administrative data were collected from external sources such as welfare files, IRS files and state employment.

## AVAILABLE DATA

Each of the income maintenance experiments was designed to generate a large volume of data from personal interviews and other sources to permit assessment of the effects of the NIT payments on the behavior of participating families. Generally, a couple of interviews were administered prior to enrollment to determine the family's initial or baseline status. These early data were used to select participating families and assign treatment levels. Following enrollment in the experiment, lengthy periodic interviews were administered roughly three times a year until the end of the income support program (three or five years). A post-experiment interview represented the last collection of data from participants. In addition to interviews, data were collected from other sources--monthly income report forms filed by the families for calculation of experimental payments were maintained separately, and copies of administrative records such as welfare office case files were obtained in some cases.

The income maintenance experiments are unique in following the participating families continuously over a period of several years and in gathering many key socioeconomic and demographic variables on a month-by-month basis (except for New Jersey, where key observations were quarterly). The interviews tended to follow a common pattern of obtaining employment and income information from each participant who was 16 or older, and obtaining nonwage income and family expenses data from the household head who responded for the family as a whole. The questions, in each of the experiments except New Jersey, asked for a month-by-month response or responses that were convertible to a monthly base. Each interview also collected basic demographic information for each family member, including children, such as birthdate, sex, relationship to the family head, and the timing of arrivals into and departures from the family. Finally, each interview typically included several socioeconomic modules asked on a one-time or occasional basis, covering such specific topics as marital history of family heads, health status, life satisfaction, financial aspirations, and so on.

Clearly, the data from the income maintenance experiments constitute a rich resource for conducting analysis of a wide range of behaviors. Of course, the data are expressly designed to permit assessment of the impact of the experimental payments on the behavior of financial families compared to controls with regard to labor supply and employment patterns, family composition changes, and so on. The samples of control families are themselves sufficiently large to allow analysis of income and family composition changes in response to a variety of "real world" conditions. For example, transfer programs such as AFDC and food stamps are important sources of support for low-income families. The experiments' data permit studying the dynamics of the decision to go on or off welfare in relation to other experiences such as layoff or divorce. The Gary data are unique in permitting analysis of income reporting patterns for a number of income types for which observations are available from several sources--interviews, income report forms, and administrative records.

It is true that the nature of the samples restricts many kinds of research. The samples are limited to low to lower-middle income families in poor central city (or rural) communities in only a few geographic locations, each with its own peculiar conditions. The Gary sample thus permits analysis only for black urban families which are primarily female-headed. Despite these constraints, the experiments' data can claim to be unique sources of longitudinal information on many aspects of behavior for a broad stratum of American society.

The U.S. Department of Health, Education, and Welfare recognizes the importance of the data generated by the income maintenance experiments for research in many areas going beyond the effects of an NIT per se. The Department has accordingly made provisions for public access to the data, particularly to the data from the Seattle, Denver, and Gary experiments, which have the largest samples and the most extensive information.

A public use file containing 48 months of information from the Gary experiment is now available to users. Monthly files containing four years of data from the Seattle and Denver experiments have been publicly available since December 1978. There are two files for each site, one of which is based on families, the other of which is based on family heads. These four-year files are currently being expanded to cover a full six years of data from Seattle and Denver, allowing analysis of three-year families which completed the financial experiment and five-year families continuing to participate. These expanded 72-month files are scheduled for completion by the fall of 1979, at which time they will be available to interested users.

MPR is also under contract to create a cross-site file with pertinent demographic and economic data from all of the income maintenance experiments formatted to be as comparable as possible. Table I-1 indicates the availability of data from each income maintenance experiment and whom to contact for further information. The remainder of this users' guide deals exclusively with the Gary data, from design of the experiment itself to details of the completed Gary public use file.

## SECTION II

### GARY EXPERIMENT DESIGN AND ADMINISTRATION<sup>1</sup>

#### THE EXPERIMENTAL DESIGN

The test of various negative income tax plans in Gary, Indiana, as in the other states, represented an experiment in which otherwise similar families were randomly assigned to either an experimental (payments-eligible) or control status and then asked to participate. Payments-eligible families agreed to abide by the rules of the experiment, which were designed to provide a model for the administrative regulations of a national program. In Gary, both experimental and control families reported income and demographic information on monthly income report forms (IRFs) and were interviewed in depth three times a year. Control subjects received only nominal payments for the information provided. By comparing the behavior of the experimental and control families, it is possible to determine statistically the effects of the experiment, since the only important difference between the two groups was the randomly-assigned treatment.

All of the income maintenance experiments were designed to examine the impact of alternative NIT plans on effort and family stability. In the Gary experiment, many other areas in which families could be affected by the NIT plans were investigated. These include:

- family expenditures
- human capital accumulation, including education and training, job search and migration
- school performance of children
- attitudinal change
- quality of life

In addition, the Gary design included a social service treatment in which selected families were eligible for day care subsidies and social service information and referral assistance.

#### The Payments Treatment

Negative income tax plans provide for partial supplementing of income below a given cutoff--the break-even level. Families with incomes below the break-even level receive negative income tax payments,

<sup>1</sup>This section represents an edited version of Kenneth C. Kehret, et. al., The Gary Income Maintenance Experiment: Design, Administration, and Data Files (Gary: Indiana University Northwest, August 1975), pp. 2-34.

just as families with incomes above a given level pay positive federal taxes. The NIT payments are determined by three parameters of the NIT plan: (1) the guarantee level--the payment made to a family if it has no other source of income; (2) the tax rate--the rate by which payment is reduced as a family earns income (or the share of the gap between family earnings and the break-even level which is filled by the NIT payments); and (3) the reporting and accounting system used for these calculations.

A central purpose of the income maintenance experiments was to determine the extent to which manipulating the incentive structure of the NIT (i.e., the guarantee and tax rate) would affect the amount individuals worked. In Gary, the experimental design included two guarantee levels and two tax rates.

The guarantee levels tested at Gary were \$4,300 and \$3,300 (for a family of four). Several factors were taken into account in choosing these two annual support levels. First, the guarantees had to be within a range that would be feasible for a national NIT program. Second, the less generous guarantee had to provide a clear alternative to Aid to Families with Dependent Children (AFDC). Finally, the difference between the two guarantees had to be large enough that participants might be expected to respond differently to the two plans. The \$3,300 guarantee was chosen because it provided \$1,500 a year more than the AFDC support level in Indiana for a family of four, yet it was still \$500 below the poverty line annual income for a family of that size. The \$4,300 guarantee was chosen to be about one-third greater than the \$3,300 guarantee for all family sizes. For a family of four, it was about \$600 above the poverty line.

The guarantee was adjusted by family size to reflect economies of scale (see table II-1). For example, a family of two received an annual guarantee that was two-thirds as much as a family of four received under the same support plan. Support levels were adjusted semi-annually to compensate for increases in the cost of living.

The tax rates on earned income used in the Gary experiment were 40 percent and 60 percent. For every \$100 earned, the payments of the experimental families were reduced by either \$50 or \$60. (Unearned income, such as bank interest or dividends, was also taxed at the treatment tax rate, except for AFDC and other forms of public assistance which were taxed at 100 percent.) The 60 percent tax rate on earned income was seemingly more generous than the AFDC rate of 67 percent; but since the first \$30 on earnings each month is disregarded in determining AFDC benefits, the two rates were probably quite similar.

The two experimental tax rates were designed to provide policy makers with information about the disincentive effects of taxing earnings at different rates. The higher 60 percent rate was considered more likely to discourage work than the 40 percent plan. However, it was anticipated that the lower tax rate would result in a more expensive program, because the break-even level would be higher (so that more families would be eligible for payments) and a family with earnings would receive greater payments than under the 60 percent tax

TABLE II-1  
 GARY INCOME MAINTENANCE EXPERIMENT  
 Payment Schedule, by Family Size

Family Size	\$4,300 Guarantee Level		\$3,300 Guarantee Level	
	Annual Guarantee	Monthly Guarantee	Annual Guarantee	Monthly Guarantee
1	\$ 1,433.00	\$ 119.41	\$ 1,100.00	\$ 91.66
2	2,866.00	238.83	2,200.00	183.33
3	3,582.50	298.54	2,750.00	229.16
4	4,300.00	358.33	3,300.00	275.00
5	4,872.20	406.01	3,740.00	311.66
6	5,302.10	441.04	4,070.00	339.15
7	5,558.70	465.72	4,290.00	357.50
8	5,732.00	477.63	4,400.00	366.66
9	5,817.97	484.83	4,466.00	372.16
10	5,903.95	492.00	4,532.00	377.66
11	5,989.93	499.16	4,598.00	383.16
12	6,032.93	502.74	4,631.00	385.91
13	6,075.91	506.33	4,664.00	388.66
14	6,118.91	509.91	4,697.00	391.41
15	6,161.89	513.49	4,730.00	394.15
16	6,204.89	517.07	4,763.00	396.91
17	6,247.87	520.66	4,796.00	399.66
18	6,290.86	524.24	4,825.00	402.41
19	6,333.85	527.82	4,862.00	405.16
20	6,376.85	531.40	4,895.00	407.91

NOTE: Dollar guarantees are as of January 1971 at the start of payments under the experiment.

rate. Nevertheless, the higher cost of the 40 percent plan could be partially offset if greater work effort and earnings were stimulated by that plan.

Some of the families with a \$3,300 guarantee were assigned a tax rate of 40 percent, and the remaining families with that guarantee were assigned a 60 percent rate. Similarly, some of the families with a \$4,300 guarantee had their earnings taxed at 40 percent, and the rest were taxed at the 60 percent rate. (See table II-3).

The break-even level of income, as may have become evident is determined by both the guarantee ( $G$ ) and the tax rate ( $t$ ). The simple NIT payment ( $P$ ) formula can be written as

$$P = G - tE,$$

where  $E$  is earnings.<sup>1</sup> Since payments become zero when  $G = tE$ , the earnings level where payments become zero is

$$E = \frac{G}{t}.$$

Table II-2 illustrates the basic payment calculation in the \$3,300 guarantee/40-percent tax plan. If a person in a zero-income family began earning \$120 a month, the monthly NIT payment dropped from \$275 to \$227. Notice, however, that total monthly income increased from \$257 to \$347. The NIT-payment schedule assured that, although payments declined with earnings, total income (the sum of earnings and payments) increased with earnings--it always paid to work.

The income reporting and accounting system in the Gary Experiment employed monthly retrospective reporting with a six-month carry-over. Payments in a given month were based on family income received two months previously. For example, early in June a family reported income received in May. This information was then used to calculate the family's July payments.

A record of income in excess of the break-even level received by the family in the current month was stored in the payments calculation as income carry-over. If in a subsequent month the family's income fell below the break-even level, the payment calculation added any previously unused income carryover to the income of the current month in calculating the current payment. If carry-over income was used in the calculation of a payment, the amount used up was subtracted from the family's carry-over file on a first-in/first-out basis. Otherwise, carry-over income was retained in the payment system for six months. For example, if any January carry-over was not used up by July, it was dropped from the system.

---

<sup>1</sup>This payment formula is a simplified version of the actual formula used in the Gary experiment. For example, the above formula is based on the assumption that a family had no tax withholdings, no income other than earnings or deductions.

TABLE II-2

FAMILY PAYMENT SCHEDULE FOR NIT PLAN WITH \$3,300  
GUARANTEE AND 40-PERCENT TAX RATE, BY EARNINGS

Monthly Earnings	Monthly NIT Payments	Total Monthly Income	Annual Earnings	Annual NIT Payments	Total Annual Income
\$ 0 . . .	\$ 275	\$ 275	\$ 0	\$ 3,300	\$ 3,300
120 . . .	227	347	1,440	2,724	4,164
240 . . .	179	419	2,880	2,148	5,028
360 . . .	131	491	4,320	1,572	5,892
480 . . .	83	563	5,760	996	6,756
600 . . .	35	635	7,200	420	7,620
687.50 <sup>a</sup> .	0	687.50	8,250	0	8,250
720 . . .	0	720	8,640	0	8,640
840 . . .	0	840	10,080	0	10,080

NOTE: This schedule is based on the simplified payment formula described in the text.

<sup>a</sup>The break-even level of income.

This system of income carry-over was designed to ensure that families receiving equivalent income over several months would receive roughly similar cumulative NIT payments, even if the income of some families fluctuated erratically while the income of others remained steady. By taking into account income over the past six months in determining current payments, it was possible to smooth out the income stream of the families with fluctuating income. Moreover, families that usually received high income but happened to receive low incomes in a particular month were not treated like families with consistently low incomes.

#### The Social Service Treatment

Participating families living inside the Model Cities neighborhood at enrollment were eligible for experimental day care subsidies and social services. Since approximately one-half of the enrolled families were headed by females, the Gary experiment provided the opportunity to determine (1) the extent of the demand for child care under varying subsidies, (2) whether the availability of subsidized child care arrangements would enable or encourage mothers to work, and (3) whether there would be a greater tendency for mothers to work when the NIT payments were combined with day care subsidies.

Families eligible for day care subsidies were randomly assigned subsidy rates of 100, 80, 60, and 35 percent. Assignment was disproportionate in order to counterbalance the expected effect of greater use of the program by families with higher subsidies. Within each subsidy rate, approximately 20 percent of the families was allowed to use the program for any purposes they chose. For the other 80 percent, the condition for receiving the subsidy was that the mother be employed, attending school, in a job training program, or actively seeking employment.<sup>1</sup>

Access workers were made available to participant families living in Model Cities neighborhoods at enrollment. These workers provided information about existing social services in Gary and made referrals to specific agencies when contacted by eligible families. The use of the access workers, however, was very low. After a careful review of ways in which the social service experiment treatment might be modified, the access-worker service was phased out in mid-experiment.

The social service research was designed to measure (1) the demand for social services under different NIT plans, (2) the extent to which inadequate information about social service agencies constituted a barrier to use, and (3) whether social services were more effective in assisting families when the families also obtained NIT payments.

---

<sup>1</sup>For discussion of the program, see Lois B. Shaw, "The Gary Subsidized Child Care Program," (Princeton, New Jersey: Mathematica Policy Research, Inc., October 1976).

## ADMINISTRATION OF THE EXPERIMENT

### Sample Selection, Enrollment, and Disenrollment

Initial selection of families. The first step in selecting participants was the administration of approximately 5,500 brief census interviews in the Model Cities project area and other low-income neighborhoods of Gary. Then, between July and November 1970, interviewers administered an economic baseline interview to nearly 2,400 of these families that met the original eligibility criteria. These criteria were that the household head(s) be black and there be at least one dependent under the age of 18 present. After they had completed the economic baseline, the families were administered the sociological baseline interview as a prerequisite to enrollment. Data from the baseline interviews were used to stratify households over three characteristics that were important to the experimental design.

The most important stratifying characteristic was "normal family income." A family's response to an NIT plan was expected to be related to its preexperimental income level. In addition, the cost to the experiment of making payments to families depended on the projected incomes (including NIT payments) of the participants over the course of the experiment. In order to select the sample, normal family income was measured by the ratio of the family's income as reported on the economic baseline to the poverty level income for a family of that size.<sup>1</sup> On the basis of this ratio, or "poverty score," families were divided into five poverty level groups (see table II-3). Families in the highest level were excluded from the initial enrollment.

The sample was also stratified over two other characteristics. Because of the research objectives of the experiment, a distinction was made between intact husband-wife families and those with female heads. Three female-headed families were to be enrolled for every two intact husband-wife families. A further distinction was made between families residing inside and outside Model Cities neighborhoods, because only the former were originally eligible for the experimental social service treatment. According to the experimental design, an equal number of families from inside and outside the Model Cities area were to be enrolled.

Original allocation to treatment cells. After being divided into 16 income and demographic groups, based on poverty level, family type, and residence, the sample families were randomly assigned to 5 payment treatments within each category. Thus the experimental design

<sup>1</sup> However, the economic baseline may have provided an imperfect measure of this variable since only five pay periods were covered by that questionnaire. For most families the period covered was ten weeks, but for others the period ranged from five days to five months.

<sup>2</sup> By the time of actual enrollment, much housing in the Model Cities neighborhoods had been demolished. Therefore, only about 35 percent of the originally enrolled sample resided in those areas.

included 80 cells. Families were listed in an enrollment queue and enrolled in the experiment until the desired cell sizes were reached. The number of families assigned to a particular treatment within each group was determined by a variant of the Watts-Conlisk Allocation Model, a program designed to maximize several desirable properties of the sample distribution, subject to a budget constraint. In order to minimize the cost of the experiment, a disproportionate number of families with the lowest levels of preexperimental income were allocated to the least generous treatment plans (i.e., the control treatment or the \$3,300 guarantee/60-percent-tax treatment).

Initial enrollment. Enrollment of the original sample took place between January and August 1971. Most families residing in Model Cities neighborhoods were enrolled between January and April; the remaining families were enrolled between April and August. Families with household heads over 58 years were excluded from the sample design, but some such families were enrolled by mistake at the beginning of the experiment. Halfway through initial enrollment, this restriction began to be strictly enforced.

The selection of additional families. Subsequent to the original sample selection, three special subsamples were incorporated into the experimental design--a high-income sample, a special control group, and a Hawthorne control group. (Table II-3 indicates the number of families actually enrolled in each cell, including the high income and special control families.)

Eligibility for participation had originally been limited to families in the poorest four poverty levels. However, this meant that families with incomes above the break-even level could be enrolled only in the most generous of the four plans to be tested. In order to study more families with incomes above the break-even level, 323 families were selected from among those which had been excluded from original enrollment solely because their incomes were too high. These poverty level 5 families were divided into groups of equal size by income ranges, randomly assigned among the five experimental and control treatments in an enrollment queue, and enrolled until a budget constraint was reached. As a result, 122 of these families were enrolled between December 1971 and February 1972.

In the original experimental design, all control and payments-eligible families that resided in Model Cities neighborhoods were eligible for the social service treatment. This meant that any effects of Model Cities residence would be indistinguishable from the effects of the social service treatment. In order to partially disentangle the two effects, another control sample was added to the design. This sample was composed of families which lived in the Model Cities neighborhood but which would be eligible for neither NIT payments nor experimental social services. These special control families were stratified only by sex of the head of household. Since preexperimental poverty level was not controlled during sample selection, a high concentration of families in the highest preexperimental poverty levels were enrolled as special controls. Enrollment of the 66 special con-

TABLE II-3

**GARY EXPERIMENT SAMPLE DESIGN: FAMILIES ENROLLED  
BY NEGATIVE INCOME TAX TREATMENTS AND  
SOCIAL SERVICES/DAY CARE TREATMENT**

Family Characteristics	Model Cities (Eligible for Social Services/Day Care)						Outside Model Cities (Ineligible for Social Services/Day Care)						Control Families	Special Controls	Totals			
	Experimental Families						Experimental Families											
	\$4,300		\$3,300				\$4,300		\$3,300									
	40%	60%	40%	60%			40%	60%	40%	60%								
<b>Husband-wife families</b>																		
Poverty level 1	5	2	2	4	8	0	5	4	2	1	8	1	41					
Poverty level 2	4	7	5	5	6	1	7	5	7	6	9	1	62					
Poverty level 3	7	6	8	19	60	0	9	9	9	23	53	23	203					
Poverty level 4	12	12	38	14	38	40	23	19	45	22	52	22	315					
Poverty level 5	10	5	10	6	24	0	4	9	5	8	27	8	108					
Total	38	32	63	48	136	41	48	46	68	60	149	14	729					
<b>Female-headed families</b>																		
Poverty level 1	19	18	33	8	66	0	38	33	50	26	121	12	412					
Poverty level 2	9	7	18	10	68	2	13	16	31	13	89	9	274					
Poverty level 3	8	8	8	53	14	4	7	12	10	65	17	12	202					
Poverty level 4	12	10	10	13	11	18	18	16	18	14	24	14	146					
Poverty level 5	0	1	0	1	6	0	1	1	2	1	4	1	17					
Total	48	44	69	85	165	24	77	78	111	119	255	25	1051					
Grand Total	86	76	132	133	301	65	125	124	179	179	404	40	1780					

SOURCE: John R. Maiolo, "G-X Design and Current Status of the Cells", Gary Income Maintenance Experiment, 1972, table 10.

NOTE: Poverty level categories were defined as follows:

Poverty level 1 . . . Families with pre-experimental incomes less than .5 times the official poverty level for their size and composition.

Poverty level 2 . . . Families with pre-experimental incomes  $\geq .5$  but  $< 1.00$  time the official poverty level.

Poverty level 3 . . . Families with pre-experimental incomes  $\geq 1.00$  but  $< 1.5$  times the official poverty level.

Poverty level 4 . . . Families with pre-experimental incomes  $\geq 1.5$  but  $\leq 2.4$  times the official poverty level.

Poverty level 5 . . . Families with pre-experimental incomes  $< 2.4$  times the official poverty level.

trol families began about halfway through the initial enrollment and extended through August 1971.

Approximately 250 families comprised the Hawthorne control group.<sup>1</sup> Some of these had been part of the preenrollment sample of 2,400 families; others were selected separately in late 1971. All members of the Hawthorne control group were given at least one baseline interview. They were not contacted again until the post-experiment interview.

The enrollment interview. All enrollments were conducted by trained personnel. Although a few families were enrolled in the office of the experiment, most were enrolled in their homes. The enrollers explained the design and rules of the experiment and answered questions about the program. The "filer," the family member designated to fill out the monthly Income Report Form (IRF), was assisted in filling out the family's first IRF. The enrollers gave each family a payment schedule, an abridged version of the rules of the experiment, and other descriptive material.

Multiple family units. Many households contained "multiple family units," that is, two or more family units sharing the same dwelling. The "secondary" family unit was often composed of a daughter who was an AFDC payee and her children. If discovered prior to or during enrollment, a secondary family unit was assigned a new identification number, administered a sociological baseline interview, enrolled as a separate family and treated as an independent household for the remainder of the experiment. In all 31 husband-wife and 145 female-headed secondary families were enrolled as separate units. After initial enrollment, any new secondary family unit discovered continued to be treated as part of the original household unit. Given that guarantee levels did not increase proportionally to family size, those families identified as secondary units were eligible for higher payments than secondary families in multiple household units who were not enrolled separately.

Spinoff family units. In March 1973 the rules of the experiment were changed to allow initially enrolled participants who left their households to "spin off" and form new and separate family units. In order to be considered a spinoff unit by the experiment, the individual(s) had to maintain a separate residence from that of the original participant family. The purpose of this rule change was to neutralize any family-stability bias built into the experiment and to simulate the mobility opportunities that might be available under a national NIT

---

<sup>1</sup>A Hawthorne control group is sometimes included in an experimental design because subjects are believed to respond, not only to experimental treatment, but also to the fact that they are under observation. To measure the latter response, some control subjects in a longitudinal panel study are interviewed only once or twice, thus minimizing their perception of being under observation. Because Hawthornes were only interviewed twice, their data are not included in the public use file.

program. However, only 14 spinoff family units were created before the experiment was completed.<sup>1</sup>

Disenrollment. Families were eligible to receive either NIT payments or IRF filing fees (\$10 paid monthly to each control family as an incentive to participate) for 36 months after their enrollment or through August 1974, whichever came first. Thus, the first group of families to be enrolled received their last checks in February 1974. Since high-income families had been enrolled from December 1971 through February 1972 and were disenrolled in August 1974, they did not participate in the experiment for the full 36 months. Spinoff family units were disenrolled on the same dates as their parent household units. Nearly every filer disenrolling from the experiment received a telephone call or field contact to ensure that he or she understood that the program was ending and to identify any special problems that this might have caused.

Attrition. Three hundred three families left the experiment prior to their scheduled disenrollment. Some of these withdrew voluntarily; others were dropped because they repeatedly violated rules of the experiment. As indicated in table II-4, sample attrition was inversely related to the generosity of the treatment.

TABLE II-4  
SAMPLE ATTRITION

Treatment Status	Number	Percent <sup>a</sup>
<b>Experiments</b>		
\$4,300 guarantee/40-percent tax	7	3.3%
\$4,300 guarantee/60-percent tax	21	10.5%
\$3,300 guarantee/40-percent tax	39	12.5%
\$3,300 guarantee/60-percent tax	<u>59</u>	18.9%
	<u>126</u>	
<b>Controls</b>		
Total	<u>177</u>	23.0%
	<u>303</u>	

<sup>a</sup>Percentage of originally-enrolled families in each category (see table II-3).

<sup>1</sup>Spinoff families have been deleted from the final Gary public use file.

Payments to Families

Northwest Council for Families. Northwest Council for Families (NCF) served as the agent for the experiment that received and processed the monthly IRFs, computed payments, and responded to participants' questions regarding payments or eligibility. This participant-contact agency was located near the Indiana University campus, a few miles from downtown Gary. Urban Developmental Services (UDS), the social service component of the experiment, was housed in the same building.

Designed to test the operation of a model NIT payment office, NCF experimented with alternative methods of administering payments. At the beginning of the experiment, staff members who had direct contact with participants were completely separated from the payment clerks who processed and edited IRFs. In mid-1973 some functions of the two groups were merged. Each payment clerk became responsible for processing the monthly IRFs for a fixed caseload of families. Moreover, the clerk was allowed to contact these families directly by mail or telephone when problems were encountered with their IRFs.

IRF processing. All families, both those receiving NIT payments and those serving as controls, were required to submit monthly IRFs, on which they were to report wage and nonwage income received by all family members over the age of 15, changes in household composition, and address changes. These data were used in calculating NIT payment checks.

On the first day of the month, each family was sent an IRF. The filer was supposed to report the family's income of the previous month and return the completed form within a week. When the completed form was received by the payment office, reported household composition and address changes were coded, keypunched, and added to the relevant data files. The income data were edited by payment clerks, who checked whether the reported amounts of wage and nonwage income were consistent with entries on the family's previous forms. Families were contacted in order to resolve any discrepancies. During the second half of the experiment, guidelines concerning the admissability of incorrect or inconsistent IRFs were tightened, thus limiting the discretion hitherto exercised by payment clerks in editing.

After being edited, the income data were submitted to the data processing department for keypunching, verification, and additional consistency edits, which became increasingly sophisticated over the course of the experiment. The data were then stored on computer-accessible files. Some of the information was extracted, stored on disc files, and used in calculating the NIT payments and filing fees. All control families and the experimental families whose calculated NIT payments would otherwise have been less than \$20 were sent \$10 filing fees on the first of the month. Monthly NIT payments were sent in two equal installments, on the first and fifteenth of each month.

Payments formula. The amount of a family's monthly NIT payment depended on the family's NIT treatment plan, income, size, AFDC status, and certain administrative factors. The specific formula used in calculating the payment changed several times during the first year but remained constant thereafter. The basic formula was as follows:

$$P = [G + T + A - t(W + N - D) - M - C] + R - L$$

where

P = payment

G = guarantee

T = federal, state, and FICA tax withholdings

A = year-end tax adjustment

t = treatment tax rate

W = gross wage earnings

N = nonwage income taxed at the treatment tax rate

D = allowable income deductions

M = nonwage income taxed at 100 percent (AFDC and other forms of categorical assistance)

C = carryover income

R = recovery or adjustment for previous errors in payments

L = penalty for late submission of IRF.

Earnings and nonwage income. At the outset of the experiment, all forms of income with the exception of AFDC were subject to the treatment tax rates.<sup>1</sup> Since the NIT payments were designed to replace rather than supplement means tested programs, in October 1971 the tax rate on AFDC, the other categorical forms of public assistance, and

<sup>1</sup> AFDC was completely untaxed at the beginning of the experiment because it was erroneously believed that no families covered by the NIT plans were receiving direct AFDC payments. The filers and filers' spouses in payments-eligible families had agreed to waive direct receipt of AFDC payments, but secondary family unit members had not been asked to do so. Consequently, some secondary members of such families continued receiving direct AFDC grants.

township trustee assistance (general assistance in Gary) was raised to 100 percent. The experimental tax rate on unemployment compensation was raised to 100 percent in December 1972 for similar reasons. Gross wage earnings and all other forms of nonwage income continued to be taxed at the treatment rate.

Federal, state, and FICA taxes. As part of the monthly payments process, the experiment reimbursed taxes withheld from the paychecks of wage earners when paystubs corroborating the withholdings were submitted with the IRFs. If such tax withholdings had not been added to the guarantee, the effective tax on wage earnings would have departed from the 40 and 60 percent treatment tax rates and would have depended on the tax rates of the state and federal systems. For the same reason, year-end tax adjustments were necessary when families received income tax refunds from the government or paid additional year-end taxes. Consequently copies of W-2, 1040, and Indiana income tax forms were collected from participating families each year. However, these adjustments were made for only the 1972 and 1973 tax years.

Deductions. Several income deductions could be applied against income that was taxed at the treatment rate if these deductions were suitably documented. While they varied somewhat over the period of the experiment, permissible deductions generally included alimony, child support, non-consumption court-ordered payments, unreimbursed medical expenses above an exclusion limit, and casualty losses.

Penalties. If a family submitted its IRF late, a \$5 penalty was deducted from its NIT payment. No NIT payments were made for IRFs submitted after the fifteenth day of the month. Penalties, including the withholding of the entire NIT payment, were also imposed when families refused to obey certain other rules of the experiment. For example, payments were sometimes withheld for failure to submit data required for year-end tax adjustments, failure to submit complete IRFs, or refusal to be interviewed.

Public assistance. All filers and their spouses were required to waive direct receipt of AFDC payments during their participation in the Gary experiment. However, they were promised a monthly NIT payment not lower than the AFDC grant they would otherwise have received. They were also allowed to remain on the AFDC rolls in order to retain their eligibility for Meidcaid and social services from the welfare department, listing the names and grant amounts of AFDC payees covered by experimental NIT plans. These data were keypunched and entered onto the disc file used to calculate the NIT payments.

Many secondary family unit members were unwilling to participate in the experiment if this meant that they would only be able to receive their AFDC grants as part of the family filers' NIT checks. Therefore, they were given the choice of either waiving direct receipt of AFDC or continuing to receive their AFDC grants directly. Such AFDC grants, as noted above, were taxed at a 100 percent rate by the experiment in computing the payments sent to the filers.

Since some families received their welfare grants through the experiment, the experiment needed continuing information about the eligibility benefit levels of the AFDC participants in its sample. Moreover, these participants' AFDC eligibility and benefit levels had to be determined according to special rules. (For example, NIT payments were not included as countable income in the calculation of AFDC grants, and AFDC families participating in the experiment were exempt from mandatory Work Incentive Program participation.) Several Lake County Department of Public Welfare (LCDPW) case workers were therefore assigned to work at the experiment, where they were responsible for determining the AFDC eligibility and grant amounts of individuals receiving NIT payments.

#### Day Care Services

Except for special controls, all families living inside the Model Cities neighborhood at enrollment were eligible for day care subsidies from July 1971 until the dates of their disenrollment. The subsidy levels ranged from 35 to 100 percent of the amount paid in child care fees. Some families were eligible to receive these subsidies if they were engaged in work-related activities (job search, employment, attendance in high school, college, or a certified training program), while others could use the program for any purpose. Each family could select formal day care arrangements, such as a day care center or licensed day care home, or informal care, defined as child care in the family's own home or that of a relative. Approximately 80 families participated in the day care program over the course of the experiment.

The experimental day care program was originally subcontracted to Community Coordinated Child Care (4 C's), an agency that coordinated day care services in the Gary area. In addition to administering the day care subsidy payments, 4 C's was responsible for developing a sufficient number of licensed facilities and furnishing information about providers of both licensed and unlicensed day care. In mid-1973, these responsibilities were assumed by Urban Developmental Services, the social service and day care agency of the experiment.

The day care program operated on a biweekly cycle. When a family submitted a form covering its children's day care attendance for the previous two weeks, the family was sent a subsidy check for that period. Two weeks later the family was required to submit a receipt from the child care vendor, documenting that the family had paid the fee, and an attendance form for the current period.

#### Family Interviews

During the course of the experiment, 13 interviews were administered to the participating families. These included the brief census interview; three baseline interviews (economic, sociological, and social service/day care) given to all eligible families prior to

enrollment; eight comprehensive periodic interviews (P1-P8) that were administered during the experiment; and a similar post-experiment interview (P9) administered after the phase-out. Each family was also visited at enrollment at which time one family member was given the responsibility of filling out the family's Income Report Form each month. This person, who had to be 16 or older and who was usually but not necessarily the family head, was designated the "filer." Finally, three special interviews were administered at the end of the experiment: one to families that terminated during the course of the experiment, another to families in the Hawthorne control group, and a brief telephone interview of families that moved out of town.

The normal interviewing cycle called for the administration of a new interview every four months. Some delays occurred early in the experiment because of the interview development problems, but the schedule remained regular throughout the remainder of the program.

Each interview cycle consisted of several stages. In the development stage, batteries of questions were conceptualized and refined. Then printed interviews were individually prepared for each respondent and administered. After administration, each instrument was hand checked, coded, keypunched, and subjected to computer checks for completeness and consistency.

Interview development. The responsibility for developing the interviews was shared by members of the senior research staff. Seeking to elicit the information of central interest to the project sponsors, the researchers worked in teams to develop questions that were then reviewed by the entire staff. The questions were reviewed and approved by the director of research, project officers of the U.S. Department of Health, Education, and Welfare, and staff of the U.S. Office of Management and Budget.

Each of the periodic interviews was designed to have a similar content. Common to the interviews was a labor supply core containing questions on employment status, hours, wages, and earnings that was to be asked of each family member 16 or over. The noncore section of each interview contained questions on nonwage income, expenditures, family life, and education for the family as a whole which were to be answered by the family member designated as the filer. In addition, each interview contained several different detailed socioeconomic batteries repeated once yearly. In some instances where a family response was adequate, these batteries were addressed only to the family filer; in other instances, for example, marital history or education and training, the batteries were addressed to specific individuals.

Interview preparation. On the basis of the information obtained in previous interviews, the printed instruments were individually prepared for each respondent. A label on the front cover of each filer's interview book identified all family members to be interviewed. Less detailed identification labels were included in the interview books of other family members. A flysheet inserted into each filer's interview book contained certain basic information on all members of the experimental family unit--name, relationship to filer, age, sex,

educational attainment, etc. During the interview these data were verified, and any changes were noted. This information then became part of the household composition file maintained for all families.

Several types of information were precoded into interview books in order to help interviewers omit questions not pertinent to particular individuals. A branching instruction might read: "Is the respondent male/female? If male, go to question 5; if female go to question 10." Another instruction might read: "If this box is checked, go to question 350." A member of the preparation staff would have checked the indicated box if, for example, the respondent was a member of the control group and not supposed to answer a set of questions designed for payments-eligible participants only.

Several batteries of questions included a precoded list of items to be presented to respondents. For example, the various expenditure batteries, which were repeated regularly, included a list of goods and services that may have been purchased by the respondent. If in the previous interview the respondent had reported the purchase of particular items, the interview preparation staff would mark those items for special probing. The interviewer could then ask specific questions: "you told us in the last interview that you had purchased an automobile. Do you still have it?" Thus, attention was focused on items of particular interest, and responses to previous questionnaires were verified.

Interviewing. Interviewing was conducted by Calumet Research Associates (CRA), an agency within the experiment which was separate both in name and location from the payment location. Interviewers were all local residents, and, with the exception of a few interviewers who conducted the census interview, all were black women.

The interviewer, who was responsible for initiating contact, would usually interview the entire family. In some especially large families, however, two interviewers worked together. The average family contained 2.75 persons age 16 or over to be interviewed. Because the interviewing instructions were somewhat complex, the instruments were not self-administered. Interviewers read the questions, and sometimes the response choices, to the respondents and recorded answers in the interview books. The average interview lasted approximately 75 minutes for the filer and 30-40 minutes for each additional family member. Originally, the filer was paid \$5 for the interview, and the entire family might receive up to \$10 for all the interviews its members completed. Late in the experiment the interview fee was increased to \$10 for the filer and \$2.50 for each other family member interviewed.

Occasionally, a family member failed to complete an interview. If that person was the filer, the filer's spouse, or the principal wage earner in the household, the family could be dropped from the experiment. If an interviewer could not find a respondent at home, subsequent attempts were made at different times of day. If these attempts proved futile, efforts were made to contact the individuals by mail. If such attempts were also fruitless, the participant's family could be dropped from the experiment.

A participant sometimes refused to be interviewed. In such a case, an interviewer supervisor telephoned and attempted to persuade the individual to consent to the interview. If this failed, the family was mailed two letters at two-week intervals requesting that an interview appointment be made. If no interview was arranged, the family was dropped.

Families that moved out of town during the course of the experiment were not interviewed thereafter until the end of the program. In the meantime, however, they continued to receive payments if they filed income report forms.

When completed, each interview was reviewed by a checking editor. Particular questions were checked to verify completeness, and the skip pattern was examined to determine accuracy. If any problems were revealed, the instrument was returned to the interviewer for correction. If no problems were found, it was sent to the coding department.

Coding and keypunching. The coding department coded most questions in preparation for keypunching. Those requiring the most effort were open-ended questions, which had to be fitted to standard scales. For example, the 1970 Census Index of Occupation Codes was used to code participants' occupations. Interviews were keypunched onto computer cards, which were then submitted to quality-control checks before being transferred to computer tape.

Quality control and data editing. Every interview that had been administered was computer checked for completeness and consistency. The information gathered was reviewed using three computer programs designed to check different aspects of the interviews. All three quality-control programs were run simultaneously for each case, and a listing of errors was printed. Quality control analysts then compared the output against the interview books to determine the nature of the errors. If a response noted as an error required verification with the respondent, the interview book was returned so that the individual could be recontacted. Errors in either coding or keypunching were returned to the appropriate department for correction. Unresolvable errors were notated in the data file.

The three quality control checks used by the experiment were the "limit check," "skip check," and "consistency check." The limit check program contained upper and lower numerical limits to the responses on all questions and generated error messages for any responses which fell outside those limits. Researchers were responsible for setting limits on the questions they designed.

The skip check was programmed to follow the branching instructions in interviews and generate messages when questions were skipped which should have been asked, or vice versa.

The consistency check program provided a review of the internal consistency of interviews by comparing the response to questions covering similar areas. For example, the respondent's reported wage rate was multiplied by reported hours worked and compared with the total wage figure to see if the two numbers fell within a certain range.

Interviewing was completed in June 1975, but quality control editing continued through September 1977.

The Gary experiment was designed to obtain data that would provide useful information for public policy formulation. The kinds of data collected and how these data were processed for research use are the subjects of the next section.

## SECTION III

### DATA OBTAINED FROM THE GARY INCOME MAINTENANCE EXPERIMENT

Data from the Gary experiment were obtained from three main sources: periodic interviews with participating families, Income Report Forms submitted by families for calculation of payments, and administrative records of tax and transfer programs provided by state and federal government agencies. It is convenient to describe the data from each of these sources and how they were processed in turn.

#### INTERVIEW DATA

Extensive information was collected from personal interviews of participating families. Most families were interviewed 13 times during the experiment, as described in the previous section; every person in the household 16 years or older was asked to complete some sections of most interviews. In addition, some school-age children were administered a special interview module as part of the eighth periodic interview.

Each interview consisted of modules of items on a particular topic. Family heads were asked to update information on the interview fliesheet regarding family composition, indicating whether any family members had departed since the last interview and whether any new family members had arrived and their sex, age, and relationship to the filer. A labor-supply core module was then administered to every family member 16 or older. This module was designed to ascertain the respondent's current labor force status, obtain detailed information on recent work or job search behavior, and a complete history of wages, hours, earnings, unemployment, and job search activity since the previous interview. The labor-supply core, somewhat revised over time was repeated at each interview.

The second major portion of each interview consisted of a set of questions about all sources of family income. The filer was asked to give a month-by-month accounting of income received by source, either for the family as a whole or for each recipient, for the time period covered by the interview. In addition, the filer was asked about family expenditures during the preceding month. Information on purchases of consumer durables was obtained from the filer once a year.

Finally, each interview concluded with a set of additional modules specific to that interview. Batteries of social-psychological items were usually repeated once a year on a rotating staggered basis. For example, items on educational aspirations for children were asked in the sociology baseline and the third and sixth periodic interviews. Items on social participation were asked in the sociology baseline, the first and fifth periodic interviews, and the post-experiment ninth

periodic interview. During most of the experiment, the social-psychological questions were posed only to the filer, but, in the later interviews, some batteries were also administered to the filer's spouse.

Other modules covered a wide range of possible income maintenance outcomes. Some of these, such as the housing expenditures battery, were asked yearly; others, such as verbal ability and marital history, only once. The scheduling of certain modules, such as income of teenagers and education training, reflected changing research priorities over the course of the experiment. In general, family information was obtained from the filer and information about children from the mother. Other modules, such as education and training, were administered to every person 16 or older or to some relevant subset of individuals or families (e.g., in the case of the welfare experience and migration modules).

As noted previously, the eighth periodic interview included a special student interview assessing attitudes toward school, home, and other aspects of the student's life. This module was administered to approximately 1,800 students, including all 16 to 18 year-olds participating in the experiment and a sample of students age 8 to 15. These data were supplemented by collection of school test and grade records for all children of participant families and collection of information from teachers in the Gary public school system via a self-administered questionnaire.

#### Source Data Files Created

All of the data from each periodic interview were keypunched and edited as described in the previous section and entered onto a set of computer files, one for each periodic. Collectively, these files are called the master file. From the master file, a variety of files were created for research purposes to permit analysis of the major experimental impacts.<sup>1</sup>

Several files were created on household composition, by month over the four years covered by periodic interviews, and other demographic characteristics of participants, including marital history and educational status. One file was created providing preexperiment information for every family as of January 1971; another file included so-called "current period" labor force variables for each of the nine periodics. Files were also created containing four years of monthly information from all of the periodics on family income, labor force status of persons, and wages and hours data for the longest job held by a person in each month. Some of these files were designed to include information of general-purpose interest, while others were created for specific research projects. Some were derived quite simply from the raw data, while others involved complex derivations

<sup>1</sup> See Glossary of Variable Descriptions for the Gary Income Maintenance Experiment Public Use File (Washington, D.C., Mathematica Policy Research, July 1979) for a listing and brief description of these files.

and required substantial editing and cross-checking of data. Most of the files described were referred to collectively as the intermediate files.

As is discussed further on in this section, data from the intermediate files were put together in a larger file, called MONSTA, that served as the basic source for the Gary public use file.

#### INCOME REPORT FORM DATA

All household units participating in the experiment, even those not eligible for NIT payments, were required to submit monthly Income Report Forms with data on wage and nonwage income received by family members 16 or older. The monthly IRF was a complicated instrument that was modified several times during the experiment. Wage and hours data were reported on the first four or five pages, one page for each week of the previous month. Although the format of these pages was changed from time to time, in general three kinds of data about wage earnings and employment were included for each person listed--gross wage receipts, payroll withholdings, and hours of work for each job. Families were asked to submit pay stubs to document these data. A high proportion of the wage earners in payments eligible families complied because income tax withholdings were reimbursed by the experiment when pay stubs were included with the monthly IRFs.

Entries on the last two pages of the IRF pertained to nonwage income, income deductions, and changes in family composition. The number of nonwage income items listed on the IRF was occasionally changed, but in general data covering the following components of income sources were sought: Aid to Families with Dependent Children (AFDC), Supplemental Security Income (SSI)--Disability assistance, Aid to the Blind, and Old Age Assistance,--Township Trustee assistance, social security, child support payments received, rent receipts, interest and dividends, veterans' benefits, food stamps, Unemployment Compensation (from both public and private sources), Workmen's Compensation, capital gains, alimony, insurance receipts, and income from self-employment. Uninsured medical expenses over an exclusion limit, capital losses, and child support or alimony payments were allowable income deductions when supporting documentation was provided.

Income Report Forms were required for every month of a family's participation in the experiment from enrollment through disenrollment--that is, in most cases, for three entire years. The high-income families enrolling late in the experiment had fewer than 36 months of data filed, and the 300 attrited families had, of course, fewer IRFs than most. The earliest IRFs were filed in January 1971 and contained data for December 1970; the latest IRFs were filed in July 1974 for June of that year.

The raw data from IRFs were put onto a file without further editing. Near the end of the payment phase of the experiment, an extensive IRF reedit was undertaken to create a complete and consistent

longitudinal Income Report Form stream of data for each participant family. In cleaning the data, efforts were made to correct mistakes made by filers and by payments clerks who initially processed the forms. The reedit process produced a clean file that had converted weekly person data to monthly figures and summed person wages for a family total for each month (nonwage amounts were already by month). Data from the reedited IRF file were included on MONSTA and subsequently on the Gary public use file.

In addition to the file of Income Report Form data obtained from the families, files were maintained of filing fees and negative income tax payments made to families by the experiment.

#### ADMINISTRATIVE RECORDS DATA

Near the end of the experiment, a number of external data sources were consulted. The information collected covered various types of income from transfer programs and tax payments. These administrative data provide verification of information collected in the experiment from periodic interviews and Income Report Forms.

At the beginning of the income validation activity, only a subset of the experiment population was chosen for this additional data collection effort. As it became obvious that the information obtained was potentially very useful, it was decided to perform the income validation steps on the entire experiment population.

Data were eventually obtained for most of the affected families from the following government agencies: the Indiana State Department of Public Welfare, the Lake County Department of Public Welfare, the Internal Revenue Service, the Indiana Department of Revenue, and the Indiana Department of Employment Security. Income sources covered include: Aid to Families with Dependent Children (AFDC), Supplemental Security Income (SSI), Food Stamps, Unemployment Insurance benefits, and wages, dividends, and other income as reported to state and federal tax agencies. Federal and state income tax payments were also obtained. In each case, the data provide observations prior to the start of the experiment and during some span of the four years of the experiment's operation. The length of time covered and the time unit of observation--year, quarter, month--vary by source.

The general approach followed for collecting administrative data was basically the same for each category. First, the agency was contacted for permission to collect data from its records. Some agencies (namely the IRS) required the signed consent of affected persons before their records could be addressed. Even where consent was not required, participants were informed of the impending data collection. Most families agreed; the three that objected were excluded from the effort.

Alphabetical lists of experiment participants were compared with similar lists from agencies. Addresses and household composition data were then compared to make sure the name-matches

were correct. Further checking with updated lists was done where possible. Discrepancies--such as participants reporting receipt of welfare income from an agency which has no record of such a person--were followed up more strenuously.

Available data were then taken from the administrative case files, coded onto specially-designed forms, and eventually--after several rounds of further checking and quality control steps--put onto a series of data tapes. Selected items from each file were put onto MONSTA which was the immediate source for administrative data on the Gary public use file.<sup>1</sup>

#### CREATING A PUBLIC USE FILE

As has been described, a variety of files were created of data obtained in the Gary Income Maintenance Experiment from interviews, Income Report Forms, and administrative records to permit analysis of the effects of the experiment on labor supply, marital stability, and other behaviors. Most of the so-called intermediate files created from interview data, the IRF Reedit file, and selected administrative data were merged onto a single file, dubbed MONSTA (for Monthly and Static variables). As the name implies, MONSTA is a large file that is, in fact, a patchwork of variables from many smaller files. The purpose of MONSTA was simply to gather the most-often used Gary data together in one place. Little effort was put into elimination of redundant information, and little was done to make the segments consistent. The handling of missing data, for example, varied greatly from subfile to subfile, according to the proclivities of the researchers responsible for a particular file's design; but MONSTA simply incorporated the data as they appeared on each subfile.

In creating a public use file of the most significant data from the Gary experiment for distribution to researchers, MPR re-formatted the MONSTA file to order the variables in logical sequence and apply a consistent treatment, to the extent possible, of coding schemes, values for missing data, and so on. Pertinent data were also added from several files containing preexperiment variables measured at the Econ Baseline and from a detailed file of household composition.

The Gary Public Use File--GARYFILE--is a longitudinal file for all families and persons participating in the experiments. It contains basic family composition information plus educational and marital status variables for adult family members. The bulk of the file is comprised of monthly observations of family and person income amounts by type and income taxes, plus items describing labor force status and wages and hours on the longest job held in each month by adult members 16 or older. The file has the unique feature that many variables are available from more than one source for many of the 48 months on the file--namely interviews, Income Report Forms, and administrative records.

Not all of the data collected in the experiment could be contained on a single public use file. Appendix A to this guide

<sup>1</sup> The kinds of information obtained from each source are described more fully in the Glossary of Variable Descriptions.

provides a complete list of research studies completed by MPR staff using data available on GARYFILE and also other data. The reports cited may inform the user about the totality of the Gary data base and indicate the extent to which certain kinds of data and certain subjects have already been researched. Appendix B provides a subject index to interview modules not included on GARYFILE. The contents of GARYFILE itself, the documentation developed by MPR for accessing the file, and the potential uses and limitations of the data are reviewed in the next section. Finally, section V describes how to obtain copies of GARYFILE and also how to obtain extracts of other variables in the experiment data base not contained on that tape.

## SECTION IV

### THE GARY INCOME MAINTENANCE EXPERIMENT PUBLIC USE FILE

In its final version, the Gary public use file--GARYFILE--contains information on 1,807 families and 5,325 persons 16 or older. The file is accompanied by extensive documentation which has been designed to give all necessary information to the user. This section of the users' guide describes the file's content and format and each piece of documentation, and provides an overview of uses to which the data can be put.

#### FORMAT AND CONTENTS

The Gary public use file is a hierarchical file with two record types. The file is arranged by family; each family record is followed by one person record for each family member 16 or older. Observations include key demographic, income, and employment/unemployment items from three sources--interviews, Income Report Forms, and administrative records--as described previously. By no means all of the variables found on the raw master files, the early research files, or even on MONSTA have been included; redundant and highly specialized variables were trimmed away, as were a few items known to contain incorrect information. Even after this pruning, the GARYFILE contains a wide range of data on records of over 10,000 characters (family) and over 7,700 characters (person) in length, respectively. Appendix C provides a full list of GARYFILE variables in the order they appear in the family and person records.

DPLS version  
Family and  
person recs  
in separate  
rectangular  
files.

#### The Family Record

The family record, identified by a unique family number, contains demographic and income variables. The demographic variables begin with experiment-specific items--the family's preexperimental income-poverty level ratio; the assigned treatment level; enrollment, payment, attrition, and disenrollment dates--and go on to describe headship status and family composition over time. Then come interview dates and completion codes; finally, information is provided for each child under 16 as to birthdate, relationship to the filer, and months present or not present in the family.

The next section of variables on the family record contains income data from interviews: preexperiment variables from the Econ Base (administered during the latter part of 1970) are followed by monthly

variables for nonwage, earned, and unemployment income from all nine periodics. These variables are all dollar amounts. The time period covered for each of the periodic variables is a 48-month span from January 1971 through December 1974, referred to as months 1 through 48 on GARYFILE. For each family, there is one or more months of pre-experimental observations (depending on when the family was enrolled), followed by up to 36 months of observations during the course of the income maintenance experiment itself, followed by four or more months of post-experiment observations.

Income variables from Income Report Forms follow, also providing nonwage, earnings, and unemployment income amounts. IRFs cover up to 36 months for each family, always falling in the range of month "0" (December 1970) to month 42 (June 1974). The end of the family record consists of administrative income data, including some nonwage income sources (Gary experiment payments, AFDC,<sup>1</sup> Supplemental Security Income, Food Stamps, and Unemployment Insurance benefits), state and federal tax data, and data on wages from the state employment security agency. Administrative variables include observations at the time of the Econ Base and yearly, quarterly, or monthly observations (depending on the source) over all or some portion of the time span of the 48 months covered by periodic interviews.

#### The Person Record

The person record, containing demographic, income, and employment variables, has a more complicated flow than the family record. It begins in the same way with demographic variables--monthly status codes and presence codes, monthly marital status and relationship codes, birthdate, sex, and education variables.

Then comes an extensive section of interview data on employment status and income. Variables from the Econ Base provide information on previous work history and on labor force status and current job at the time of that interview prior to the start of the experiment. These variables are available only on the person records for the male and female heads of each family. Next come monthly income variables for the 48-month time span from the periodic interviews. These variables are similar to those on the family record, though somewhat less complete for nonwage income sources. Next comes a series of "current period" variables. These are point-in-time employment variables--employment status, type of worker, hours worked, disability status--as of the period immediately preceding each of the nine periodic interviews. Finally, a series of 48-month variables is provided that detail the person's employment status and aspects of his/her main job in each month.

---

<sup>1</sup> The section of administrative data contains the only source of AFDC information on the entire file: both periodic-reported and IRF-reported AFDC income were unuseable as originally given.

The remainder of the person record is relatively straightforward: IRF data include wages, withholdings, and hours worked variables, while administrative variables are the same as those on the family record.

Illustration of File Format

Figures 1 and 2 show copies of actual dumps from the Gary public use file for a family and family head, with major segments of data demarcated by source and subject. Figure 3 shows a partial dump of a few lines from the same family and person records, highlighting demographic and income variables.

The assigned family number, 5502, appears at the beginning of both the family and person records. The treatment level code of 02 shows this to be a financial family, while the 48-month family type code of 02 shows that the family was female-headed throughout the experiment. Looking further into this family's composition, it appears from birthdates, relationship codes, and monthly family size, that there were five children under 16 present in the family for all months after enrollment.

Aside from the Gary experiment payment, this family seems to have had two major sources of income--Social Security and Veterans Administration payments, both of which are substantial.

There is only one person record for this family, that of the female head. (Had there been a male head at any time during the interview period, or had one of the children turned 16 before the last interview, that person would also have a person record.) Her unique person number 550201, immediately follows the family number. Her monthly presence indicator shows her to have been present for all months after enrollment, while her marital status and birthdate show she was widowed and in her early 30s during the experiment. There are no education data for her.

The Social Security and Veterans Administration amounts in the section of income variables, are the same as those on the family record for months in which the items were asked by person--other months have zeros or a 9997 missing data code. She never worked during the time covered by her interviews; the monthly percent not in the labor force was always 100.0 percent. This non-working status is confirmed by the appearance of the 9s...3 missing data code (indicating she had no data of any kind in this section) for all variables dealing with the principal job in each months.

Because she did not work, there are no positive IRF data on the person record, and because she did not receive income of the types detailed in the administrative records, there are no data for comparison here, either.

This small glimpse of one set of records may convey a sense of the available information and the way in which the data are ordered in the GARYFILE. To make the file as accessible and easy to use as possible, MPR has developed a comprehensive set of documentation, described in the following paragraphs.

**FIGURE 1**

## CARYFILE FAMILY RECORD DUMP SHOWING MAJOR DATA SEGMENTS

## **FAMILY**

## DEMOGRAPHIC VARIABLES

## I) INTERVIEW VARIABLES Econ Base, Income

### 48-month, Income

FIGURE 1 (continued)

**INCOME REPORT  
FORM VARIABLES**

Incomes

## **ADMINISTRATIVE VARIABLES**

Income

FIGURE 2

## GARYFILE PERSON RECORD DUMP SHOWING MAJOR DATA SEGMENTS

(continued)

**FIGURE 2 (continued)**

**INCOME REPORT FORM  
VARIABLES**

**Withholdings and Hours**

## **ADMINISTRATIVE VARIABLES**

## **Income**

**FIGURE 3**

**PARTIAL DUMPS OF GARYFILE FAMILY AND PERSON RECORDS  
SHOWING SELECTED DEMOGRAPHIC AND INCOME VARIABLES**

## **Character Positions**

Family Record

1 = 100  
101 = 100  
201 = 100

481	-	500
591	-	600
601	-	700
781	-	800

1000 1100

3901	-	4000
4001	-	4100
4101	-	4200
4201	-	4300
4301	-	4400

**CODES KEY**

**Treatment level code:**

**2 - Low guarantee, low tax rate**

**Family type code:**

2 - Female filer, no spouse

### **Relationship code:**

01 - File

03 - Son

04 - Daughter

**Monthly presence indicator:**

2 = Present in family

**Marital status:**

Q2 = Widowed

**Person Record**

I - 100  
101 - 200  
101 - 400

801	-	900
901	-	1000
1001	-	1100
1101	-	1200

•  
•  
3301 - 3400  
3401 - 3500

FIGURE 4

SPECIMEN RECORD FORMAT DESCRIPTION  
PAGE FOR GARYFILE

GARYFILE Family Record - 7

VARIABLE NUMBER	VARIABLE NAME	FIELD POSITION	LENGTH	VARIABLE DESCRIPTION
F1405	UMPAYECN	3174-3176	3	MONTHLY UNEMPLOYMENT PAY FROM EMPLOYER, DOLLARS, FROM ECON BASE
F1406	WCOMPECN	3177-3179	3	MONTHLY WORKMEN'S COMPENSATION, DOLLARS, FROM ECON BASE
F1407	INCOME69	3180-3184	5	REVISED FAMILY INCOME, 1969
		3185-3200	16	Blank (unassigned)

Periodic Interview Variables

The variables in this section are dollar amounts of income as reported by the family head in the periodic interviews. The maximum time period covered by these data is January 1971 (month 1) through December 1974 (month 48)--hence each variable is allowed space for 48 months of information for each variable. Nonwage income amounts are given first, followed by earnings and unemployment income information.

F1408- F1455	SSI1- SSI48	3201-3392	48 months x 4	SSI (BLIND, DISABLED, OLD AGE ASSISTANCE) INCOME, DOLLARS, MONTHS 1-48, FROM PERIODICS
F1456- F1503	TT1- TT48	3393-3584	48 months x 4	TOWNSHIP TRUSTEE INCOME, DOLLARS, MONTHS 1-48, FROM PERIODICS
F1504- F1551	FSVAL1- FSVAL48	3585-3776	48 months x 4	FOOD STAMPS, FACE VALUE, DOLLARS, MONTHS 1-48 FROM PERIODICS

## DOCUMENTATION

Aside from this users' guide, which is intended to introduce the Gary Income Maintenance Experiment and the resulting public data file, there are three major pieces of documentation accompanying the file: a Record Format Description, a glossary of variable descriptions, and a printout of the distribution of responses to each item produced with a computer program called STATS.

### Record Format Description

The record format description, or RFD, is the basic tool for accessing the data, whether using a standard package or a specially-written program. Technical characteristics are given at the beginning of the RFD, such as number of reels, number and sequence of logical records, and blocking factor. The rest of the RFD lays out the arrangement of variables on each record type. For each variable, the RFD provides a unique variable number and mnemonic name, the character position and length of the variable, and a variable title which briefly describes the item. Meanings of coded variables are included in the RFD: most are given just under the variable title, while extremely lengthy codes are found in RFD attachments.

Variable numbers are in ascending order from the beginning of each record type. These numbers can be used to reference the glossary for detailed information about particular variables. The variable name, or mnemonic, is never more than eight characters long to facilitate use with statistical packages such as SPSS. The field position indicates the location of the variable on the file and can be used to cross-reference the STATS printout. Variable length is the number of characters occupied by each field. The variable title provides a brief statement about each variable, including meanings of coded variables. This statement is meant to be descriptive rather than definitive, and the user should always reference the glossary description for a variable for a complete description.

Figure 4 shows a specimen RFD page from the Gary family record. Note that because many variables are simply repetitions of the same item, the convention has been to give one title to the entire set, such as "Township Trustee Income, Dollars, Months 1-48." Character positions are given to include all months, in this case 3393 - 3584, and the length is defined as "48 months x 4," indicating that there are 48 monthly fields for this variable (starting in month 1, January 1971), and that each monthly value of township trustee income is 4 characters long. The variable name, listed as TT1-48, also indicates the months pertinent to the particular variable, and can be used to define specified months, such as TT12 or TT36, which reference the dollar amounts received in the twelfth and thirty-sixth months, respectively.

As discussed previously, variables on GARYFILE cover different time spans and reference different units of time according to the source of the data. (For example, periodic interview data cover 48 months, IRF

FIGURE 5

SPECIMEN VARIABLE DESCRIPTION FROM GARY FILE GLOSSARY

a  
P1456-  
P1503  
  
b  
T1-  
T48

c  
VARIABLE TITLE  
TOWNSHIP TRUSTEE INCOME, DOLLARS, MONTHS 1-48  
FROM PERIODICS

d  
ALLOWABLE CODES  
0000 = No dollars  
0008-0696 = Dollar amount received  
9993 = Family has no periodic income data  
9994 = Missing data  
9999 = Family attrited, or month after  
P9 interview

e  
VARIABLE  
DESCRIPTION  
This variable shows, for each of the 48 months covered by periodic interviews, the dollar amount of assistance from the Township Trustee received in each month by the family and reported in periodic interviews. Township Trustee assistance is a local type of aid available on an emergency basis to persons or families ineligible for other assistance programs.

f  
TEXT DESCRIPTION  
Respondent. The respondent for interview questions on Township Trustee income was always the filer.

Universe. In P1, P2, and P6, the amount received from the Township Trustee was asked for the family as a whole; in P3-P5 and P7-P9, the question was asked of the filer for each recipient.

Wording. The Township Trustee as an income source was always addressed on a list of several possible income sources. Four categories of assistance from the Township Trustee were considered: (1) rent assistance; (2) food assistance; (3) assistance with utility bills; and (4) other assistance from the Township Trustee.

Variations in these categories were few: In P4, the category of rent assistance was omitted; in P5-P9, the category "food assistance" was changed to read "food stamp assistance."

In all periodics, the respondent was first asked whether he/she or anyone else in the family received income from this source in the time covered by the periodic. If the response was positive, he/she was

(continued)

FIGURE 5 (continued)

TEXT DESCRIPTION (cont'd)	asked for the dollar amount received in each month. As indicated above, the amount was sometimes asked for the family as a whole and sometimes asked for each recipient.
VARIABLE <sup>g</sup> DERIVATION	The amounts reported for each of the categories of assistance were aggregated by month.  For months in which the dollar amounts were collected by person, the amounts were then aggregated across all recipients in the family for each month.  The amounts thus arrived at for each month were assigned to the appropriate months.
COMMENTS <sup>h</sup>	The changes in the universe for this item mean that the data are not strictly comparable across periodics.  Fluctuations in the number of responses and amounts reported may also reflect the emergency nature of this type of assistance.  Other variables showing Township Trustee assistance are listed by subject in Appendix D.

<sup>a</sup>Variable Number. This is a unique number assigned to each variable on GARYFILE. Numbers on the family record are prefaced with the letter F, while those on the person record are prefaced with a P. The numbers shown on the glossary descriptions correspond to those on the RFD to facilitate cross-reference.

<sup>b</sup>Variable Name. This is a unique mnemonic of eight or fewer characters assigned to each variable. This mnemonic is also found in the RFD and the STATS printout.

<sup>c</sup>Variable Title. The variable title, also found on the RFD, is a brief one- or two-line description of the variable for easy reference.

<sup>d</sup>Allowable Codes. Listed for each variable are the response values that may be found in the records for that item, including the not applicable codes. In most cases the range given shows the actual range of responses; in some cases the range shows the possible minimum and maximum responses.

<sup>e</sup>Variable Description. This is a description which supplements the variable title.

<sup>f</sup>Text Description. This item gives either the exact text of a question as it appeared in the interview questionnaire, or gives a synopsis of the question and any other related questions. Where changes in wording occurred across periodics, the interviews involved are noted, along with a description of the wording change(s).

<sup>g</sup>Variable Derivation. This section describes in some detail how the GARYFILE variable was derived from the original raw data.

<sup>h</sup>Comments. Finally, in many cases, is a comments section which provides additional information which may help the user in properly interpreting and applying the responses, including warnings about sparse or anomalous responses and reference to related variables elsewhere on the file.

data cover a maximum of 36 months, some administrative data are annual, quarterly, or monthly.) To remind the user of the time reference of each type of variables, the RFD includes explanatory notes at the beginning of each major segment of data on GARYFILE.

### Glossary

The glossary which accompanies the Gary public use file provides detailed definitions of the variables on the file. The glossary begins with an introduction to the file and to using the glossary itself. Following this introduction is a section of overviews to the variable types--demographic variables, interview variables, Income Report Form variables, and administrative variables. The overviews describe generally how that class of data was collected and extracted for the public use file.

The bulk of the glossary consists of a specific variable description for each variable on the GARYFILE. These are organized by record type (family and person) and by the order in which the variables appear on the file. The listings repeat variable number, mnemonic name, and the variable title from the RFD, and indicate all allowable codes that can be expected for each variable. The listing then describes the manner in which the data were collected: in some cases the exact question wording is given, in others the context and questions are simply described. A derivation segment recapitulates the processes involved in creating the particular variable from the raw data. In some cases this was as simple as extracting an item; in others, the procedure was quite complex--involving, for example, cross-checks with other items on the file and returning to the actual instruments as filled out by interviewers to decipher marginal notes and calculations. Finally, a comments section describes peculiarities in the data, notes missing data codes, cautions about items of low response, and provides additional hints about derivation and use of a particular item. Figure 5 shows a specimen variable description.

The glossary is a large document designed to give an in-depth description of each variable on the Gary public use file, so that users can make valid inferences from the data and understand both their applicability and limitations. Of course, not all eventualities can have been predicted, and in the course of using the data, the user may encounter further anomalies.

### STATS Printout

The STATS Program, developed by staff of Social and Scientific Systems, Inc., readily provides, in a compact form, information about the distribution of responses to individual variables from a lengthy data set such as the Gary public use file. A STATS run has been made of both record types and included in the documentation.

For each variable, the program lists the mnemonic name and character position as in the RFD, counting as a separate variable each monthly observation for an item such as Township Trustee Income. For

each field, the program calculates and displays the minimum and maximum non-zero and non-missing data values (values greater than "0" and less than 9s...3), the mean and standard deviation of these values, and the number of records with positive responses in this range. Another column shows the number of records with a value of zero. Finally, there is a column for each possible missing data code, from 9s...3 through 9s...9.

For continuous variables, such as the monthly Township Trustee Income dollar fields, the STATS information gives quite a good picture of the distribution of responses. Figure 6 shows a page of the STATS printout for Township Trustee dollar income for months 1 through 20. As can be seen, the mean value of this variable fluctuated during this period, with a low of \$64 and a high of \$111. The standard deviation also fluctuated, reflecting the nature of this emergency type of aid. The 9s...3 column, with 68 families in each month indicates the number of families which simply had no data for the entire segment of periodic income data; 9s...4s indicate missing data of an unspecified type while the 9s...9s, beginning in month 14, show the number of families officially attrited. The "frequency" column gives the number of positive responses to the item--in this case a relatively small number of families.

For coded variables, STATS also show the minimum, maximum, mean standard deviation, and other information, but the measures are not as meaningful.

For example, the mean value of the relationship code in month 1 is 2.75, with a range of 1 to 26 for that month. What the user really wants in this case is the distribution of responses to the individual codes. Dates are another example of variables which have little meaning on the STATS--aside from seeing the frequency of responses and various missing data codes. For coded fields and dates, a series of tabulations has been run and appended to the STATS printout.

Working with the RFD, glossary, and STATS printout, the user should be able to make effective appropriate use of the data on the Gary public use files. MPR maintains backup records on procedures used in creating the file in order to assist with unexpected problems that may crop up when using the data.

SPECIMEN GARYFILE STATS PAGE

#### USES OF THE DATA

The demographic, income, and employment variables on the Gary public use file offer opportunities for research on a number of different topics, including replication studies of research already carried out (see appendix A for a list of completed reports by MPR staff from the Gary experiment) and studies that explore new ground. Effects of the negative income tax payments on labor supply behavior, patterns of receipt of different kinds of income over the course of the experiment and family composition can be studied comparing experimental and control families. Eligibility and participation in current welfare programs such as AFDC and Food Stamps can also be analyzed for the families in the control sample.

The GARYFILE is one of a small set of currently available data files that make possible analysis of behavior patterns over time. Monthly observations are provided for many variables over a four-year time span, permitting detailed analysis of intra-year and inter-year variations in income streams, job behavior, and arrivals and departures of family members.

A unique feature of the GARYFILE is that it includes observations on many different kinds of income from more than one source--periodic interviews, Income Report Forms, and administrative records. Income amounts can be compared across sources for consistency and differences analyzed according to such factors as frequency of reporting of Food Stamp data on Income Report Forms would correspond more closely to the amounts recorded in administrative agency records than the amounts reported in the periodic interviews. Families were interviewed at intervals of several months or longer, while IRFs were required to be submitted monthly for the calculation of experimental payments. On the other hand, one might posit that the tax rates imposed by the experiment on unearned income would be an incentive to misreport these sources on IRFs and to report more accurately on interviews. Having observations from multiple sources on GARYFILE permits these kinds of analysis. Table IV-1 indicates, for each

TABLE IV- 1  
GARY INCOME VARIABLES BY SUBJECT AND PLACEMENT ON THE FILE

Family Record				Person Record		
Subject	Interviews	IRFs <sup>1</sup>	Admin.	Interviews	IRFa <sup>1</sup>	Admin.
<b>Nonwage Income</b>						
AFDC	---	---	Preexp; months 1-18	---	---	Preexp; Months 1-48
Gary Payment	---	---	36 months	---	---	---
SSI	Months 1-48	36 months	Preexp; months 1-36	months 1-48	---	Preexp.; months 1-36
Township Trustee	Preexp; months 1-48	36 months	---	months 1-48	---	---
Food Stamps	Months 1-46	36 months	Preexp; months 1-36	---	---	Preexp; months 1-36
Social Security	Preexp;	36 months	---	months 1-48	---	---
VA	Preexp; months 1-48	36 months	---	months 1-48	---	---
Assets, etc.	Preexp; months 1-48	---	Annual,	--	---	---
Other/mis- cellaneous	Preexp; months 1-48	36 months	---	months 1-48 <sup>2</sup>	---	---
Unemployment	Preexp; months 1-48	36 months	Months 1-44	months 1-48 <sup>2</sup>	---	---
Strike pay/ Workmen's Comp	Preexp;	36 months	---	months 1-48 <sup>2</sup>	---	Months 1-48

Continued

TABLE IV-1 (continued)

Subject	Interviews	IRF	Admin.	Interviews	IRFs	Admin
<u>Wages</u>						
Regular job	Months 1-48	---	---	Preexp; months 1-48	---	---
Daywork	Months 1-48	---	---	Preexp; months 1-48	---	---
Odd jobs	Months 1-48	---	---	Months 1-48	---	---
Self-employment	---	36 months	---	---	---	---
Gross wages	---	36 months	Annual, 1969-73; quarterly, 1970-74	---	36 months	Annual, 1969-73; quarterly 1971-74
Taxes/with- holdings	---	---	Annual 1969-1973	---	36 months	Annual, 1969-73

<sup>1</sup> IRFs cover a maximum of 36 months for any one family; these 36 months always fall within the range of month 0 (December 1970) through month 43 (June 1974).

<sup>2</sup> Person records do not contain a full 48 months of data for these variables, as the questions were sometimes asked by family only.

income type, the availability of observations from interviews, IRFs, and administrative records, and the span of time covered in each case. Appendix C to this users' guide lists all of the variables on GARYFILE in the order in which they occur on the family and person records.

As further illustration of the potential of the GARYFILE data for detailed analysis of participating families over time, selected information on family composition, income, and job experience of two of the families found on the GARYFILE is presented in graphic form.

#### Wage Data for Two Gary Families

The two families shown in Figures 7 and 8 are a control family (family number 5580) and a financial family (family number 5599). Both are female-headed throughout, as was typical of the majority of Gary households enrolled in the experiments.

The control family consisted of five persons, a 41-year old female head with four children aged 8 through 12 (all ages are given as of the month of enrollment). The head worked during the entire 48 month period, while the family also steadily received AFDC and Social Security income. As shown in figure 7, combined Social Security and AFDC income totalled \$2,932, \$2,389, and \$3,358 in 1971, 1972, and 1973, respectively, and already totalled \$1,950 at the end of the first half of 1974.

The female head worked full-time for steadily increasing wages: though her 1973 wage income was slightly lower than that in 1972, a further look at the file shows that there was a period of a few months in 1973 when she was not working; when she returned to her job, her 1974 wage income rose. For years in which Income Report Form and administrative data (in this case from the Internal Revenue Service) are also available, it is interesting to note that there is no great difference between the annual amounts of wages received. Monthly differences are more marked, as seen in the graph of periodic-reported and IRF-reported gross wages where the latter shows more month-to-month fluctuation--probably due to lack of time lag in reporting.

The person's job stability is obvious: she reports the same industry code for the entire period covered. Her occupation changed twice during the period studied, apparently improving each time. Though she was out of the labor force in months 30 and 31, most probably due to illness she returned to the same job at similar wages for the remainder of the time.

The second family had a treatment level of a high guarantee and a high tax rate. Family composition remained stable throughout the period covered, consisting of a 40-year old head and two teenagers aged 14 and 15 at enrollment. Before the financial Gary experiment, the family depended upon AFDC for a steady income. During the experiment, the Gary payment replaced AFDC, to which the family returned, however, immediately after disenrollment from the experiment. In mid-1973, with the older child's entrance into the labor force, her wages became the other main source of family income. This family had,

FIGURE 7

INCOME VARIATIONS IN SELECTED GARY FAMILY  
CONTROL: FEMALE HEAD EMPLOYED

FAMILY NUMBER:  
5580  
TREATMENT LEVEL:  
Control  
HEADSHIP STATUS:  
Female headed

FAMILY COMPOSITION:  
AT ENROLLMENT:  
Female head, aged 41  
Four children, aged  
6 - 12  
Social Security,  
Job earnings, Social  
Security, Occasional Food Stamps  
"Other" income

GROSS EARNINGS:

Month	Income (\$)
6	5503
10	4204
15	3893
20	3660
25	4500
30	4500
35	4500
40	4500
45	4500

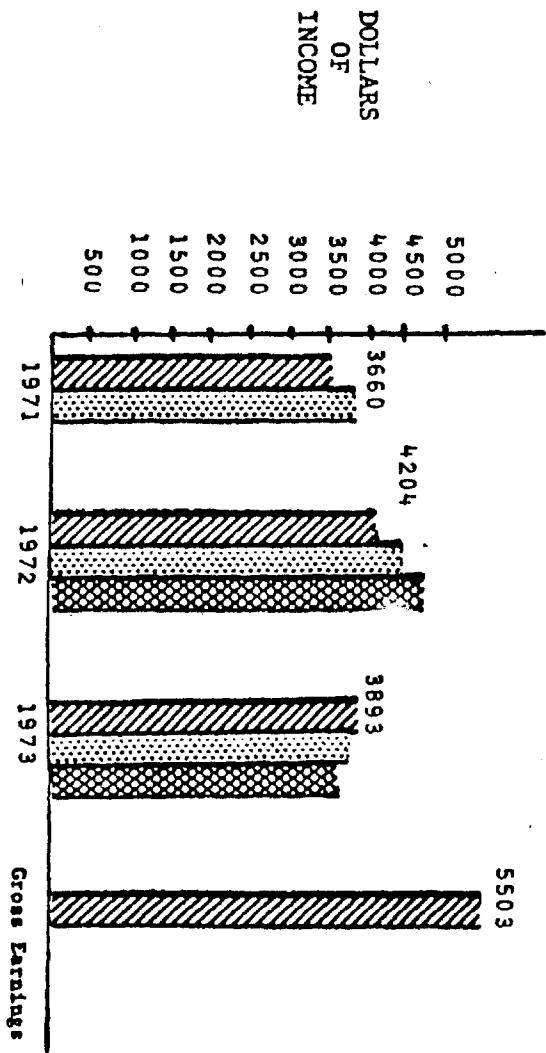
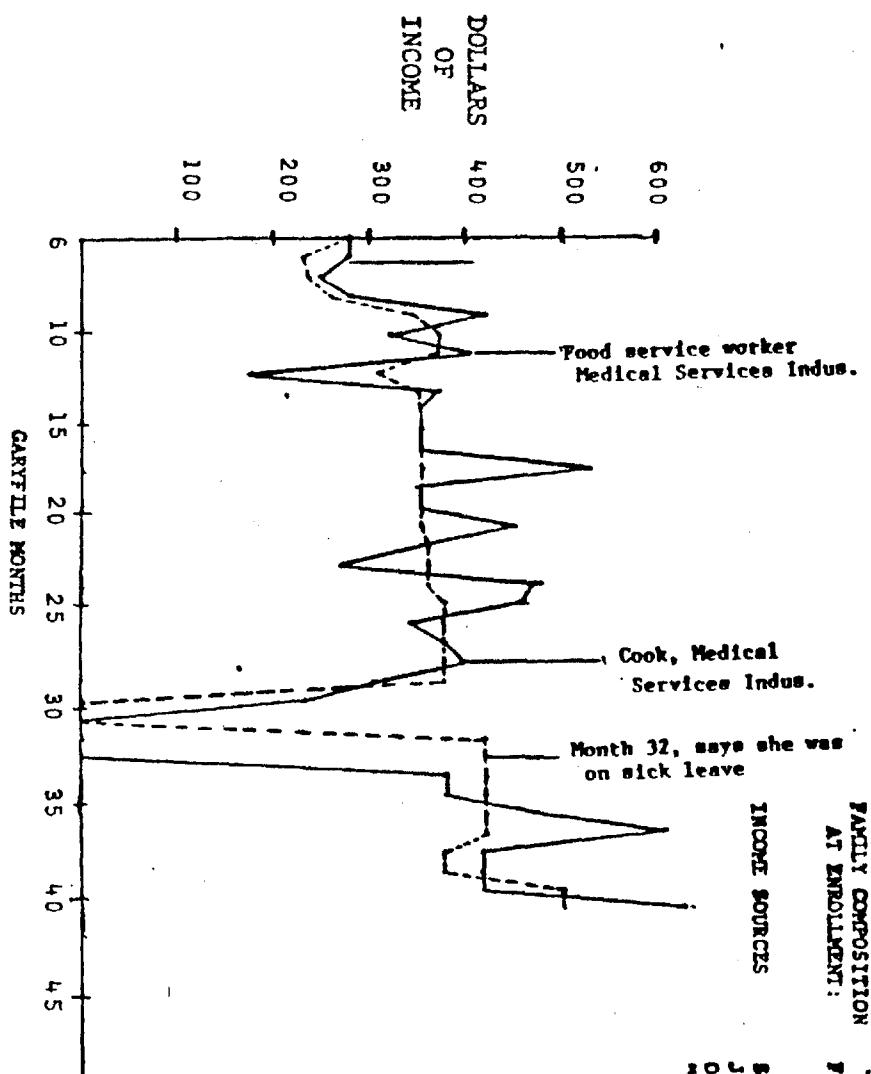
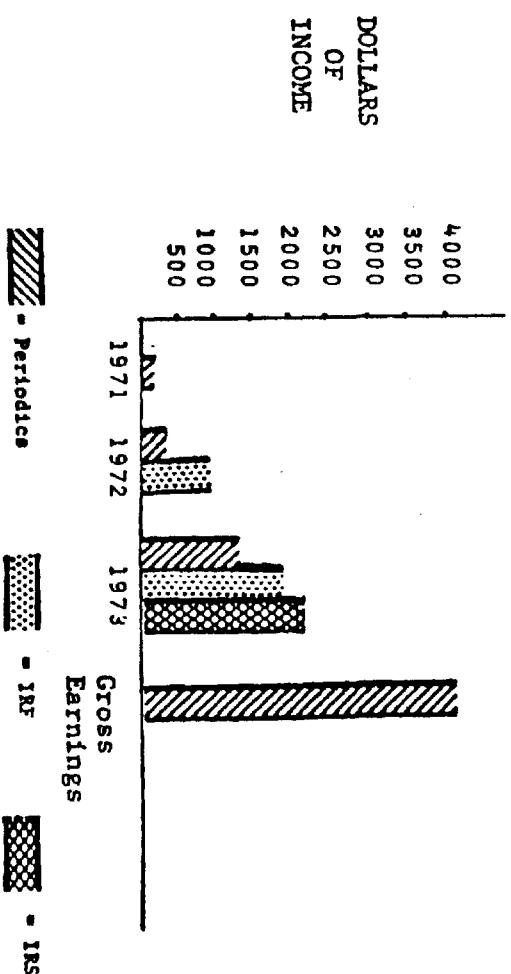
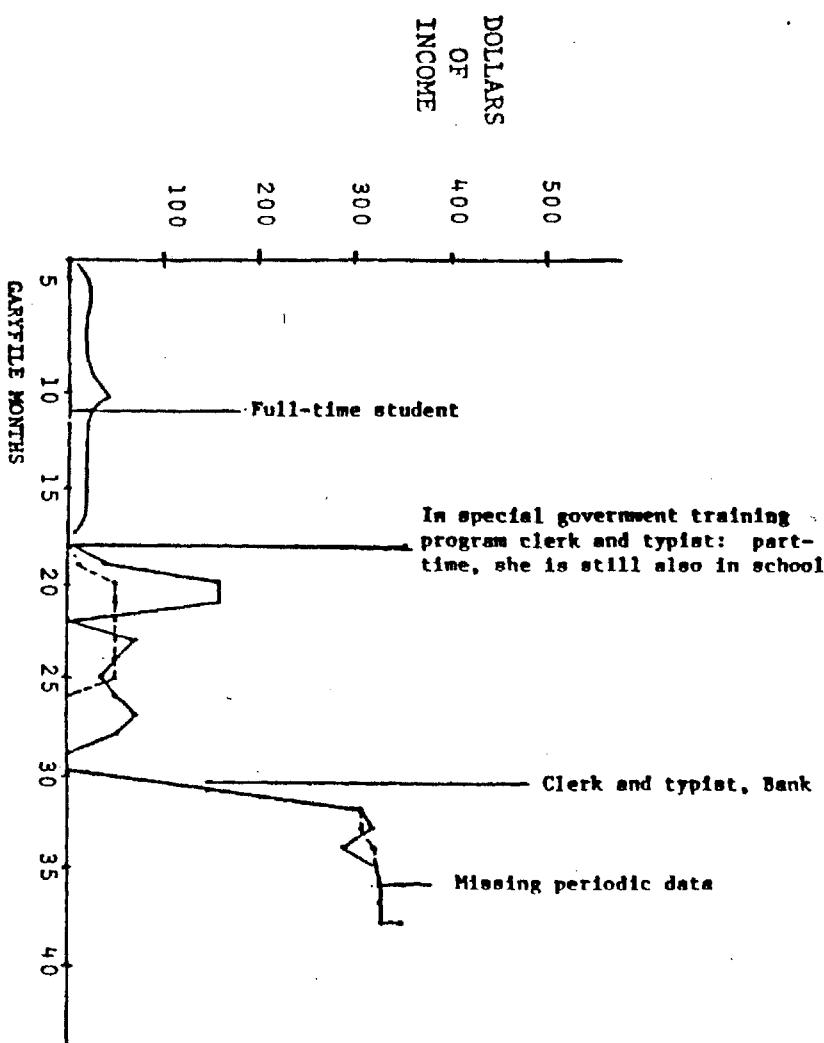


FIGURE 8

INCOME VARIATIONS IN SELECTED GARY FAMILY  
FINANCE: DAUGHTER EMPLOYED

FAMILY NUMBER: 5599  
TREATMENT LEVEL: High guarantee,  
high tax rate  
HEADSHIP STATUS: Female headed  
FAMILY COMPOSITION AT EN-  
ROLLMENT: Two daughters,  
ages 14 and 15  
INCOME SOURCES: Occasional Food  
Stamps/Township  
Trustee



in addition, a patchwork of occasional income from other sources--Food Stamps, Township Trustee Income (a local emergency aid), odd jobs, and several instances of "other" or unspecified income appear sporadically for the family.

Because Gary experiment payments were set to be equal to or higher than AFDC payments, it is possible, in this case, to compare what might have been with what was: that is, one can compare the family's "computed" AFDC income--what they would have received from the AFDC program in the absence of the income maintenance experiment, all other things being equal--with the amounts actually received. Summing the yearly amounts (thus including AFDC in early and late months), it can be seen that the family was far better off under the experiment than they would have been otherwise.

The graph of family wages in figure 8 shows a great change in wage income over the four years due, as discussed below, to the older daughter taking on a full-time job. Where comparison can be made between periodic-reported wages and wages from other sources, there are somewhat more marked discrepancies than with the control family observed.

The female head in this family did not have any job during the time covered by periodic interviews. Because she indicated disability in earlier months, it may be that this was what kept her from working.

Her two daughters, however, contributed to the household, especially the older one whose work pattern is shown in figure 8. Attending a vocational school through the 12th grade, she worked part-time during the summer of 1972 and during her final year of school. Her job was described as that of a clerk-typist in a special government training program. Upon finishing 12th grade in June of 1973, she began to work full-time in a bank. The nature of her work fluctuated between clerical and typing. Once started on this job, she worked steadily at it through the 48th month, and was the family's sole wage-earner.

Her younger sister also worked part-time in the summers of 1971 and 1972 in a special government training program. She was classified as a messenger. She was in the 12th grade in the school year of 1973-74, but apparently did not, like her sister, go immediately to a full-time job.

#### Concluding Note

The glimpse presented above of only two families out of the many on the file may give some idea of the potential the file has for further research. As noted before, GARYFILE does not contain all of the data collected in the Gary experiment and is not suitable for every kind of research analysis one might want to perform on the sample of participating families. Moreover, the sample itself represents a restrictive population--namely black, largely female-headed families in an industrial inner-city area. Nevertheless, the file affords a rich, longitudinal data base on the population included in the Gary experiment.

Section V provides information on how to obtain GARYFILE from Mathematica Policy Research. Users interested in data on subjects listed in Appendix B not included on GARYFILE are advised how to go about obtaining these data.

SECTION V  
OBTAINING GARY EXPERIMENT DATA

GARYFILE

File and Tape Format and Price

GARYFILE, the public use file of monthly demographic, income, and employment data from the Gary Income Maintenance Experiment, is available from Mathematica Policy Research. The file contains over 1,800 family records and over 5,000 person records for family members aged 16 or over. Each record is 10,400 characters long. The number of tape reels needed for the file depends on the density, as shown below along with the applicable prices for GARYFILE.

<u>Number of reels</u>	<u>Density</u>	<u>Price</u>
1	6250 BPI	\$ 300
2	1600 BPI	\$ 300
4	800 BPI	\$ 400

Included in the price are several items not obvious in the above pricing scheme. Tape reels themselves are furnished by MPR with standard labels unless otherwise requested. The comprehensive set of documentation includes a record format description, the glossary of variable descriptions, and a STATS run of GARYFILE variables--all as described earlier in this users' guide. Shipping will normally be by UPS; air freight entails extra cost. Finally, though it is hoped that the documentation will answer most questions, the price structure allows for several hours of consultation to handle questions or problems arising from use of the GARYFILE.

GARYFILE Extracts

It is possible that some users would prefer special extracts of GARYFILE variables--of a particular set of variables, or for a particular subgroup of participants, for instance. The variable list in appendix C should give some idea of whether an extract is preferred. Users considering this option may request a copy of the GARYFILE record format description to use in making detailed extract specifications. Cost and time estimates for extracts will depend upon the nature of the extract, but in general will run to at least twice the cost of a file copy.

How to order GARYFILE

An order form should accompany this user's guide. Further questions could be addressed to:

Mathematica Policy Research, Inc.  
2101 L Street, N.W., Suite 416  
Washington, D.C. 20037  
Attn: Miriam Aiken  
(202) 833-9510

SOCIOECONOMIC AND SOCIAL-PSYCHOLOGICAL MODULES

Appendix B lists the subjects of various socioeconomic and social-psychological modules available upon a special request basis. As with GARYFILE extracts, prices will be individually negotiated, and will run higher than GARYFILE costs to allow for the extra processing and documentation.

DATA FROM OTHER INCOME MAINTENANCE EXPERIMENTS

As shown in table I-1, public use files are currently available from the Seattle and Denver sites for four years, and are soon to be expanded to 72-month files. A cross-site file is also being developed for basic information from the New Jersey, Seattle, Denver, and Gary Income Maintenance Experiments. Information on these current and projected files may be obtained from MPR in Washington, D.C., at the above address.

**APPENDICES**

## APPENDIX A

### MATHEMATICA POLICY RESEARCH PUBLICATIONS ON THE GARY INCOME MAINTENANCE EXPERIMENT

Sponsored by the U.S. Department of Health, Education, and Welfare, Mathematica Policy Research, Inc., has completed a number of analyses of data from the Gary Income Maintenance Experiment. Publications resulting from these analyses are listed below, and may be obtained from:

Mathematica Policy Research, Inc.  
Publications Office  
P.O. Box 2393  
Princeton, New Jersey 08540.

#### The Gary Income Maintenance Experiment: Initial Findings Report

Volume I: Initial Labor Supply Findings. Edited by Kenneth C. Kehrer and Robert A. Moffitt, July 1976. \$22.05

Chapters in Volume I are:

"The Negative Income Tax and Labor Supply: Methodological Issues," John F. McDonald, Robert A. Moffitt and Kenneth C. Kehrer.

"Income Maintenance and the Labor Supply of Male Family Heads," John F. McDonald.

"Income Maintenance and the Labor Supply of Wives," Richard L. Kaluzny.

"Income Maintenance and the Labor Supply of Female Heads," Lois B. Shaw.

"Income Maintenance and the Labor Supply of the Family," Stanley P. Stephenson.

Volume II: Other Initial Findings, Edited by Kenneth C. Kehrer. \$14.25  
Chapters in Volume II are:

"Income Maintenance and Family Consumption and Wealth: An Initial Report on the Gary Asset, Debt, and Expenditure Data," V. Joseph Hotz and Charles E. Metcalf, March 1977.

"Income Maintenance and Housing: Baseline Consumption Patterns and Initial Findings on Changes in Housing Tenure and Mobility," Richard L. Kaluzny, January 1976.

"Income Maintenance and the Utilization of Social Service Agencies," Billy J. Tidwell, October 1977.

"Income Maintenance and the School-Enrollment and Labor Supply Decisions of Teenagers," John F. McDonald and Stanley P. Stephenson, Jr., January 1977.

"The Gary Subsidized Child Care Program," Lois B. Shaw, October 1976.

"Participation in the Access-Worker Program of the Gary Income Maintenance Experiment," Billy J. Tidwell, August 1975.

"Participants' Knowledge and Understanding of the Gary Income Maintenance Experiment," Billy J. Tidwell, Richard L. Kalyzny, Elise Bruml, and Dan DuRoss, 1976.

"The Gary Income Maintenance Experiment: Summary of Initial Findings," Kenneth C. Kehrer, March 1977. Also available as a separate paper, 91 pages, \$4.65

Other Mathematica Policy Research, Inc. papers and publications using Gary Income Maintenance Experiment data are listed below under the appropriate category of paper:

#### Technical Analysis Series

Selection Bias in the Estimation of Experimental and Quasi-Experimental Effects, Robert A. Moffitt, May 1977. 20 pages, \$2.00

#### Staff Papers

Fertility of Unmarried Females in the Gary Income Maintenance Experiment, Charles M. Wolin, May 1978. Staff paper 78A-01. 49 pages, \$2.45

Estimating Labor Supply Disincentives of a Negative Income Tax: Some Results and Lessons from the Experiments, Robert A. Moffitt and Kenneth Kehrer, December 1978, Staff paper 78A-06. 23 pages, \$2.00.

The Labor Supply Response in the Gary Income Maintenance Experiment, Robert A. Moffitt, December 1978. Staff paper 79A-02. 34 pages, \$2.00.

Working Papers

- The Gary Income Maintenance Experiment: Design, Administration, and Data Files, Kenneth Kehrer, Elise Brumel, Gary Burtless, and David Richardson. August 1975. Working paper A-10. 69 pages, \$3.45
- The AFDC Dominance Problem in the Gary Income Maintenance Experiment, Robert A. Moffitt, August 1976. Working paper A-14. 25 pages, \$2.00
- A Life-Cycle, Human Capital Model of Labor Supply Response to Negative Income Taxation: Evidence from the Gary Experiment, David L. Horner, October 1977. Working paper A-24. 50 pages, \$2.50.
- Findings from the Gary Income Maintenance Experiment: Testimony before the House Committee on Ways and Means, Kenneth C. Kehrer, October 11, 1977. Working paper A-26. 14 pages, \$2.00.
- The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Experiment. Gary Burtless and Jerry Hausman, November 1977. Working paper A-27. 40 pages, \$2.00.
- The Labor Supply Response to Tax and Transfer Programs, Robert A. Moffitt December 27, 1977. Working paper A-29. 27 pages, \$2.00.
- The Short-Run Labor-Market Replacement Effect of a Negative Income Tax, Robert A. Moffitt, revised November 1977. Working paper C-22. 27 pages, \$2.00.
- Will Welfare Reform Improve Children's School Performance?, Rebecca A. Maynard and Richard J. Murnane, January 1978. Working paper C-24. 51 pages, \$2.55.

## APPENDIX B

### SOCIOECONOMIC AND SOCIAL-PSYCHOLOGICAL MODULES

The last portion of each interview--following the sections on family finances and each adult's work experience during the interview period--was comprised of one or more modules on a specific subject, generally of a socioeconomic, sociologic, or psychological nature. Data from these modules were not used in creation of the GARYFILE; however, the responses as found on the raw data tapes are maintained at MPR, and extracts may be obtained on an individual basis.

The index on the next few pages illustrates the subject areas and the periodics in which these areas were approached. Use of the question numbers gives the user the general length of each module, or the number of pertinent questions in the subject area. Aside from putting the responses onto the raw data types, no further cleaning has been done on these items. Potential users should realize that it is typical of these "noncore" data that the responses have not been thoroughly checked and cleaned beyond the initial coding stages, so the items, though covering a variety of interesting subjects, will be neither as straightforward nor as consistently reliable as the variables on the GARYFILE.

Extracts of the socioeconomic and social-psychological modules are available on an individual basis. The terms would depend upon the length and complexity of the items required. Queries should be addressed to:

Mathematica Policy Research, Inc.  
2101 L Street, N.W., Suite 416  
Washington, D.C. 20037  
Attn: Miriam Aiken  
(202/833-9510)

APPENDIX B (continued)

INDEX TO INTERVIEW MODULES  
BY SUBJECT, PERIODIC, AND QUESTION NUMBER

Subject Area	Base-Line*	P1	P2	P3	P4	P5	P6	P7	P8	P9
<b>Economic Modules</b>										
Time allocation	SSDC Q101-102	---	Q512- Q522	Q621-24 Q713-41	Q640- Q708	---	---	Q961- Q1024	---	---
Education and Training	Econ Q15-17 Q126-130	Q523- Q601	---	---	---	---	---	---	Q1062- Q1070	Q1032- Q1098
<b>Expenditures &amp; Assets</b>										
Expenditures Assets (other than for autos or housing)	SSDC Q94-100	Q451- Q452B	Q654- Q665	Q550- Q620; Q665-711; Q747-58; Q768-79	Q576- Q587	Q758- Q769	Q791- Q807; Q1060- Q1091	Q791- Q807	---	Q1301- Q1449
Housing Expenditures & Houses Owned	SSD Q27-62	---	Q666- Q688	---	Q746- Q768	Q850- Q903	---	---	Q952- Q1018	---
Auto Expenditure & Autos Owned	SSDC Q93	---	---	Q625- Q664	---	---	---	Q1096- Q1128	---	---
Perceived Rate of Time Discount	---	---	---	Q744- Q746	---	---	---	---	---	---
Income of Teenagers	---	---	---	---	---	---	Q1012- Q1022	Q851- Q865	---	---
Migration	Econ Q6-10	---	---	---	Q714-31; Q769-90	---	---	---	Q1019-26; Q1101-11	---

\*NOTE: Abbreviations under the "Baseline" heading are: SSDC stands for the Social Services/Day Care Baseline Interview. Econ stands for the Economic Baseline Interview; and Soc stands for the Sociological Baseline Interview. P1 stands for the first periodic interview, and so on through P9 which stands for the ninth periodic. The ninth periodic was always a post-experiment interview.

SOURCE: Kenneth C. Keherer et al, The Gary Income Maintenance Experiment: Design Administration, and Data Files, (Gary: Indiana University Northwest, August 1975), pp. 2-34.

## APPENDIX B (continued)

Subject Area	Base-Line	P1	P2	P3	P4	P5	P6	P7	P8	P9
<b>Family Resource Pooling</b>	---	---	---	---	---	---	---	Q875- Q888	---	---
<b>Work History</b>	Q131- Q145	---	---	---	---	---	---	---	Q1033- Q1070	Q1135- Q1138; Q1500 Q1535
<b>Perceived Implicit Tax Rate</b>	---	---	---	---	---	---	---	---	---	Q1201- Q1218
<b>Social-Psychological Modules</b>										
<b>Social Participation</b>	Soc Q62- Q76	Q500- Q512	---	---	---	Q910- Q922	---	---	---	Q1536-Q1538; <sup>c</sup> Q1100-Q1105
<b>Voluntary Association</b>	Soc Q77- Q95	Q513 Q520	Q873- Q880	---	---	---	Q1030- <sup>b</sup> Q1039 <sup>c</sup>	Q845- <sup>c</sup> Q849	---	---
<b>Political Participation</b>	SS Q15- Q25	---	---	---	Q732- Q737	---	---	---	---	---
<b>Attitudes</b>										
<b>Political</b>	Soc Q118 Q124	---	---	Part IV <sup>a</sup> Q11- Q13	---	---	Q1040 <sup>b</sup> Q1045	Q849B <sup>c</sup>	---	---
<b>Self</b>	Soc Q126 Q130	---	---	Part IV <sup>a</sup> Q22; Q20; Q52; Q47	---	---	Q828; Q1041 <sup>b</sup> ; Q1047 <sup>b</sup> Q1052	Q1047- <sup>d</sup> Q1048	---	---
<b>Children</b>	SSDC Q26	Q523		Part IV <sup>a</sup> Q35	---	---	Q1046	Q1018- <sup>c</sup> Q1081; Q1049	---	---
<b>Job</b>	Soc Q137; Q12-14	Q534	---	Part IV <sup>a</sup> Q23- Q27	Q608- Q639	Q978	Q827- Q843; <sup>b</sup> Q1042- Q1044	Q833- Q834	---	---

a Part IV was a supplemental instrument administered simultaneously with the third periodic interview.

b Questions asked of female filers and spouses of male filers.

c Questions asked of filers and spouses.

APPENDIX B (continued)

Subject Area	Base-Line	P1	P2	P3	P4	P5	P6	P7	P8	P9
Education	Q3-Q51	---	---	---	---	---	---	---	---	---
Children	---	---	---	Q800-Q847	---	---	Q925-Q938	---	---	---
Self and Spouse	Soc Q52-Q61	---	Q848 Q860	---	---	---	---	---	---	Q1077 <sup>c</sup> Q1079
Family Life	Soc Q131-Q137	Q522	Q885-Q886	Part IV <sup>a</sup> Q48-Q55	---	Q978	Q1048 <sup>b</sup> Q1054	Q901 <sup>b</sup> Q957	---	Q1003-Q1006
Social Mobility	---	---	---	---	---	Q979-Q989	---	---	---	Q1120 <sup>c</sup> Q1131
Life Satisfaction	SS Q7-9	Q568-Q570	Q891-Q892	---	---	---	---	Q1080-Q1082	---	---
Relative Deprivation	Soc Q109-Q114; Q134	---	---	Part IV <sup>a</sup> Q14-Q15	---	Q740	---	---	Q1087-Q1088	Q1106-Q1111
Future Orientation	Soc Q108-Q114; Q134	Q521	Q884	---	---	---	---	Q828 <sup>c</sup> Q838	---	---

a Part IV was a supplemental instrument administered simultaneously with the third periodic interview.

b Questions asked of female filers and spouses of male filers.

c Questions asked of filers and spouses.

APPENDIX B (continued)

Subject Area	Base-line	P1	P2	P3	P4	P5	P6	P7	P8	P9
<b>Miscellaneous</b>										
Parental Characteristics	---	---	---	---	---	---	---	---	Q821- Q863	---
Financial Aspiration	---	---	---	---	---	---	---	---	Q1082- Q1086	---
Marital Experience	---	---	---	---	---	---	Q1055- <sup>b</sup> Q1056	Q833- <sup>c</sup> Q834; Q1043- <sup>d</sup> Q1046	---	Q975- Q1002
Verbal Ability	---	---	---	---	---	---	---	---	---	Q1168- <sup>c</sup> Q1183
Health of Respondents	SS Q1-4	Q543	---	---	Q792	---	Q854- Q933	Q1136- Q1137	---	---
Health of Children	---	---	---	---	---	---	Q939- <sup>b</sup> Q965	---	---	---
Time and the Project	---	---	---	---	Q710- Q713	---	---	---	---	---
Participant Knowledge	---	---	---	---	---	Q770 Q799	---	---	---	---
Participant Perception	---	---	---	---	---	---	---	---	---	Q850- Q961

a Part IV was a supplemental instrument administered simultaneously with the third periodic interview.

b Questions asked of female filers and spouses of male filers.

c Questions asked of filers and spouses.

**APPENDIX B (continued)**

<b>Subject Area</b>	<b>Base-line</b>	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P6</b>	<b>P7</b>	<b>P8</b>	<b>P9</b>
<b>Social Services Utilization</b>	SS Q6; Q9- Q10	Q568- Q590	---	---	Q791- Q806	Q800- Q843	---	---	---	---
<b>Day Care</b>	Soc Q5	---	---	Part III <sup>a</sup> Q850- Q1027	---	---	---	---	Q1120- Q1137	---
<b>Welfare Experience</b>	---	---	---	---	---	---	---	---	Q870- Q950	---
<b>Fertility</b>	Soc Q141- Q153	---	---	Part IV <sup>a</sup> Q28-34	---	---	Q981- Q1008	Q1051- Q1052	---	Q1014- Q1026

**a** Parts III and IV were supplemental instruments administered with the third periodic instrument.

## APPENDIX C

## VARIABLE LIST FOR GARYFILE

All variables on both family and person records of the Gary Income Maintenance Public Use File are listed here in the order in which they appear on the file. Variable format is briefly indicated.

Variable	Format
FAMILY RECORD	
<u>Demographic Variables</u>	
Family number	Assigned number
Poverty level	Coded
Treatment level	Coded
Cell number	Coded
Model Cities status	Coded
Enrollment date	Julian date (YYDDD)
First transfer payment number	Coded
Last transfer payment number	Coded
Attrition date	Julian date (YYDDD)
Disenrollment date	Julian date (YYDDD)
Multiple family flag	Coded, months 1-48
Family type code	Coded, months 1-48
Continuity of household headship	Coded, months 1-48
Person number of principal male	Assigned number
Filer status of principal male	Coded
Person number of principal female	Assigned number
Filer status of principal female	Coded
Time on financial experiment	Number of months, months 1-48
Total family size, each month	Number of persons, months 1-48
Total number of original enrollees	Number of persons
Total number of subsequent enrollees	Number of persons

APPENDIX C (continued)

Variable	Format
Number of children under 6 in family, each month	Number of children, months 1-48
Number of children under 16 in family, each month	Number of children, months 1-48
Total number of core person records	Number of records
Total number of noncore persons	Number of persons
Date of Economic Baseline Interview	Julian date (YYDDD)
Date of periodic interviews	Julian date (YYDDD), periodics 1-9
Interview completion code	Coded
Person number, each noncore	Assigned number
Birthdate, each noncore	Year and month (YYMM)
Relationship to filer, each noncore	Coded
Monthly presence indicator, each noncore	Coded, months 1-48

Income Variables from the Econ Baseline

Township Trustee income, monthly amount	Dollars
Social Security income, monthly amount	Dollars
Veterans Administration income, monthly amount	Dollars
Assets, property, family business income, monthly amount	Dollars
Miscellaneous nonwage income, monthly amount (includes six categories)	Dollars
Other nonwage income, monthly amount	Dollars
Unemployment benefits, monthly amount	Dollars
Workmen's Compensation, monthly amount	Dollars
Revised family income for 1969	Dollars

APPENDIX C (continued)

Variable	Format
<b><u>Income Variables from Periodic Interviews</u></b>	
Supplemental Security Income per month	Dollars, months 1-48
Township Trustee income per month	Dollars, months 1-48
Food Stamps, face value per month	Dollars, months 1-48
Food Stamps, amount paid per month	Dollars, months 1-48
Social Security income per month	Dollars, months 1-48
Veterans Administration income per month	Dollars, months 1-48
Assets and property income per month	Dollars, months 1-48
Miscellaneous nonwage income per month (sum of several categories)	Dollars, months 1-48
Other specified income per month	Dollars, months 1-48
Job earnings per month, working for someone else, all adult family members	Dollars, months 1-48
Dayworker earnings per month, all adult family members	Dollars, months 1-48
Odd job income per month, all adult family members	Dollars, months 1-48
Unemployment compensation/supplemental unemployment benefits per month, all adult family members	Dollars, months 1-48
Strike benefits/Workmen's Compensation per month, all adult family members	Dollars, months 1-48

**Income Variables from Income Report Forms<sup>1</sup>**

First month of family's IRF filing period	Month number
Last month of family's IRF filing period	Month number
Supplemental Security Income per month	Dollars, maximum 36 months

<sup>1</sup> IRF variables cover a span of 43 months from month 0 (December 1970) through month 42 (June 1974); each family has a maximum of 36 months of data with that span.

APPENDIX C (continued)

Variable	Format
Township Trustee income per month	Dollars, maximum 36 months
Food Stamps, face value per month	Dollars, maximum 36 months
Food Stamps, amount paid per month	Dollars, maximum 36 months
Social Security income per month	Dollars, maximum 36 months
Veterans Administration income per month	Dollars, maximum 36 months
Other miscellaneous income per month	Dollars, maximum 36 months
Self-employed earnings per month	Dollars, maximum 36 months
Gross family wage earnings per month	Dollars, maximum 36 months
Number of wage earners in family each month	Number of persons
State unemployment compensation per month	Dollars, maximum 36 months
Unemployment compensation from employer per month	Dollars, maximum 36 months
Strike benefits per month	Dollars, maximum 36 months

Income Variables from Administrative Sources

Gary Experiment Payment Files

Gary check amount per month	Dollars, months 1-48
-----------------------------	----------------------

AFDC Records

Total family AFDC amount, 1970	Dollars
Number of months AFDC received, 1970	Number of months
Family AFDC amount of Econ Baseline	Dollars
Family AFDC received per month	Dollars, months 1-48
Family AFDC calculated eligibility per month	Dollars, months 1-48

APPENDIX C (continued)

Variable	Format
<b><u>Supplemental Security Income Records</u></b>	
Total family SSI amount, 1970	Dollars
Number of months SSI received, 1970	Number of months
Family SSI amount for month of Econ Baseline	Dollars
"Current" SSI award per month	Dollars, months 1-36
"Retroactive" SSI award per month	Dollars, months 1-36
<b><u>Food Stamp Records</u></b>	
Food Stamps, total face value, 1970	Dollars
Food Stamps, total amount paid, 1970	Dollars
Number of months received Food Stamps, 1970	Number of months
Food Stamps, face value for month of Econ Baseline	Dollars
Food Stamps, amount paid for month of Econ Baseline	Dollars
Food Stamps, face value per month	Dollars, months 1-36
Food Stamps, amount paid per month	Dollars, months 1-36
<b><u>Internal Revenue Service Records</u></b>	
Annual gross wages, family	Dollars, annual amount, 1969-1973
Annual dividends and interest, family	Dollars, annual amount, 1969-1973
Annual amount of other income/loss, family	Dollars, annual amount, 1969-1973
Total federal tax, family	Dollars, annual amount, 1969-1973
Total federal tax withholdings, family	Dollars, annual amount, 1969-1973
<b><u>Indiana Department of Revenue Records</u></b>	
Annual gross wages, family	Dollars, annual amount, 1969-1973
Annual dividends and interest, family	Dollars, annual amount, 1969-1973

APPENDIX C (continued)

Variable	Format
Annual amount of other income/loss, family	Dollars, annual amount, 1969-1973
Total state tax on income, family	Dollars, annual amount, 1969-1973
Total state tax withheld, family	Dollars, annual amount, 1969-1973

State Division of Employment Security Records

Total wages paid to adults in family, 1970	Dollars
Numbers of quarters wages received 1970	Number of quarters
Wages paid to adults in family in quarter including Econ Baseline	Dollars
Wages paid to adults per quarter	Dollars, 1971-1974
Unemployment benefits to adults per month	Dollars, months 1-48

PERSON RECORD

Demographic Variables

Family number	Assigned number
Person number	Assigned number
Person status code by month	Coded, months 1-48
Marital status by month	Coded, months 1-48
Relationship to filer by month	Coded, months 1-48
Birthdate	Year and month (YYMM)
Sex	Coded
Enrollment date	Julian date (YYDDD)
Attrition date	Julian date (YYDDD)
Disenrollment date	Julian date (YYDDD)
Highest grade completed at school year '71-'72	Coded

APPENDIX C (continued)

Variable	Format
Current school type, at school Year '71-'72	Coded
Highest grade completed at school year '72-'73	Coded
Current school type, at school year '72-'73	Coded
Highest grade completed at school year '73-'74	Coded
Current school type, at school year '73-'74	Coded
 <u>Econ Baseline Variables: Employment</u> <sup>1</sup>	
Ever been employed?	Yes/No
Ever worked in family business	Yes/No
Year began to work	Year (YY)
Year began to work full-time	Year (YY)
Number of jobs in working career	Number of jobs
Number of jobs in last five years	Number of jobs
Number of jobs in last 3 years	Number of jobs
Number of jobs in last year	Number of jobs
Total 1969 income	Dollars
Currently employed?	Yes/No
Current labor force status	Coded
Number of months of experience on current/most recent job	Number of months
Industry code	Coded
Occupation code	Coded
Do work hours vary with the season?	Yes/No
Season you are out of work	Coded

<sup>1</sup> Econ Base variables on the person record are given for male and female heads only.

APPENDIX C (continued)

Variable	Format
Hours per week, most recent pay period, working for someone else	Number of hours
Overtime hours per week, most recent pay period, working for someone else	Number of hours
Hourly wage rate, most recent pay period, working for someone else	Cents
Overtime wage rate, most recent pay period, working for someone else	Cents
Gross weekly pay, most recent pay period, working for someone else	Dollars
Normal weekly overtime hours, working for someone else	Number of hours
Normal wage rate, working for someone else	Cents
Normal overtime wage rate, working for someone else	Cents
Normal gross weekly pay, working for someone else	Dollars
Hours worked per week, most recent pay period, dayworker	Number of hours
Hourly wage rate, most recent pay period, dayworker	Cents
Gross weekly pay, most recent pay period, dayworker	Dollars
Normal weekly hours, dayworker	Number of hours
Normal hourly wage rate, dayworker	Cents
Normal gross weekly pay, dayworker	Dollars
Do you plan to return to work in next 30 days?	Yes/No

APPENDIX C (continued)

Variable	Format
<u>Income Variables from Periodic Interviews</u>	<sup>1</sup>
Supplemental Security Income per month	Dollars, months 1-48
Township Trustee income per month	Dollars, months 1-48
Social Security income per month	Dollars, months 1-48
Veterans Administration income per month	Dollars, months 1-48
Miscellaneous nonwage income per month (sum of several categories)	Dollars, months 1-48
Other unspecified income per month	Dollars, months 1-48
Job earnings per month, working for someone else	Dollars, months 1-48
Dayworker earnings per month	Dollars, months 1-48
Unemployment compensation/supplemental unemployment benefits per month	Dollars, months 1-48
Strike benefits/Workmen's Compensation per month	Dollars, months 1-48
<u>Current Period Employment Variables</u>	
Employment status at time of interview	Coded, periodics 1-9
Type of worker at time interview	Coded, periodics 1-9
Reason not working at time of interview	Coded, periodics 1-9
Total hours worked in week before interview	Number of hours, periodics 1-9
Overtime hours worked in week before interview	Number of hours, periodics 1-9
Normal regular hours worked per week at time of interview	Number of hours, periodics 1-9
Extent of disability at time of interview	Coded, periodics 1-9

<sup>1</sup> Nonwage income variables were not always asked by person: about third of the months will show missing data codes for these variables.

APPENDIX C (continued)

Variable	Format
Extent disability limits work at time of interview	Coded, periodics 1-9
<u>Employment Variables from Periodics</u>	
Type of worker each month	Coded, months 1-48
Percent of month employed each month	Percent, to one decimal, months 1-48
Percent of month worked each month	Percent, to one decimal, months 1-48
Reason employed but not working each month	Coded, months 1-48
Percent of month not in labor force each month	Percent, to one decimal, months 1-48
Percent of month unemployed each month	Percent, to one decimal, months 1-48
Monthly employment status	Coded, months 1-48
Disability status each month	Coded, months 1-48
Number of jobs held each month	Number of jobs, months 1-48
Number of months of experience on longest job in each month	Number of months, months 1-48
Industry code for longest job each month	Coded, months 1-48
Percent of month worked on longest job each month	Percent to one decimal, months 1-48
Regular weekly hours on longest job each month	Number of hours, months 1-48
Hours change flag, each month	Coded, months 1-48
Weekly overtime hours on longest job each month	Numbers of hours, months 1-84
Hourly wage rate on longest job each month	Cents, months 1-48

APPENDIX C (continued)

Variable	Format
Wage rate change flag, each month	Coded, months 1-48
Overtime wage ratio for longest job each month	Coded, months 1-48
<u>Income Report Form Variables<sup>1</sup></u>	
Gross wages earned per month	Dollars, maximum 36 months
Federal tax withheld per month	Dollars, maximum 36 months
State tax withheld per month	Dollars, maximum 36 months
FICA withheld per month	Dollars, maximum 36 months
Number of jobs held each month	Number of jobs, maximum 36 months
Total hours worked each month	Number of hours, maximum 36 months
Overtime hours worked each month	Number of hours, maximum 36 months
"Special" hours each month (hours paid for, not worked)	Number of hours, maximum 36 months
Special hours code each month	Coded, maximum 36 months
<u>Administrative Income Variables</u>	
<u>AFDC Records</u>	
Total AFDC income, 1970	Dollars
Number of months received AFDC, 1970	Number of months
AFDC received, months of Econ Baseline	Dollars
AFDC income, each month	Dollars
Calculated AFDC eligibility per month	Dollars, months 1-48

<sup>1</sup>IRF variables cover a span of 43 months from months 0 (December 1970) through month 42 (June 1974); each person has a maximum of 36 months of data within that span.

APPENDIX C (continued)

Variable	Format
<b><u>SSI Records</u></b>	
Total SSI income, 1970	Dollars
Number of months received SSI, 1970	Number of months
SSI income for month of Econ Baseline	Dollars
"Current" SSI award per month	Dollars, months 1-36
"Retroactive" SSI award per month	Dollars, months 1-36
<b><u>Food Stamp Records</u></b>	
Food Stamps, total face value, 1970	Dollars
Food Stamps, total amount paid, 1970	Dollars
Number of months Food Stamps received, 1970	Number of months
Food Stamps, face value per month	Dollars, months 1-36
Food Stamps, amount paid per month	Dollars, months 1-36
<b><u>Internal Revenue Service Records</u></b>	
Individual annual gross wages	Dollars, annual amount, 1969-1973
Joint annual gross wages	Dollars, annual amount, 1969-1973
<b><u>Indiana Department of Revenue Records</u></b>	
Individual annual gross wages	Dollars, annual amount, 1969-1973
Joint annual gross wages	Dollars, annual amount, 1969-1973
<b><u>Division of Employment Security Records</u></b>	
Total wages, 1970	Dollars
Number of quarters received wages, 1970	Number of quarters
Wages earned in quarter including Econ Baseline	Dollars
Wages per quarter	Dollars, 1971-1974
Unemployment benefits per month	Dollars, months 1-48