Statistical Policy Working Paper 1 - Report on Statistics for Allocation of Funds

---

Statistical Policy
Working Paper 1

Report on
Statistics for Allocation of Funds

Prepared by
Subcommittee on Statistics for Allocation of Funds
Federal Committee on Statistical Methodology

U.S. DEPARTMENT OF COMMERCE
Juanita M. Kreps, Secretary
Courtenay M. Slater, Chief Economist

Office of Federal Statistical Policy and Standards
Joseph W. Duncan, Director

Issued:March 1978

For sale by the Superintendent of Documents,

U.S. Government Printing Office

Washington, D.C. 20402

Office of Federal Statistical
Policy and Standards
Joseph W. Duncan, Director
George E. Hall, Deputy Director, Social Statistics
Gaylord E. Worden, Deputy Director, Economic Statistics
Maria E. Gonzalez, Chair, Federal Committee on Statistical Methodology

Preface

This working paper was prepared by the members of the Subcommittee on Statistics for Allocation of Funds, Federal Committee on Statistical Methodology. The Subcommittee was chaired by Wray Smith, Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education, and Welfare. The members of the Subcommittee are the authors of this report and their names are listed below. It is hoped that this report will aid administrators and drafters of future legislation in recognizing some characteristics of data and formulas used in distributing Federal funds to State and local governments. The Subcommittee plans to discuss these results with many interested parties to further disseminate the findings of this report.

Members of the Subcommittee on
Statistics for Allocation of Funds
(January 1976-June 1977)

* Wray Smith (Chair)
Office of Technical Support and Statistics, ASPE

(HEW)

Richard B. Clemmer
Division of Economic Policy, OPDR (HUD)

* Edwin J. Coleman
Bureau of Economic Analysis (Commerce)

Martin M. Frankel
National Center for Education Statistics (HEW)

* Fred J. Frishman
Mathematical Statistics Branch, IRS (Treasury)

Forrest W. Harrison
National Center for Education Statistics (HEW)

* Thomas B. Jabine
Office of Research and Statistics, SSA (HEW)

* Charles D. Jones
Bureau of the Census (Commerce)

Eli S. Marks
Bureau of the Census (Commerce)

Rajendra Singh
U.S. Postal Service

Charles D. Troob
Compensatory Education Division, National Institute of Education (HEW)

Martin Ziegler
Bureau of Labor Statistics (Labor)

Maria E. Gonzalez (ex officio, Chair, FCSM)
Office of Federal Statistical Policy and Standards
(Commerce)
Tore E. Dalenius (Consultant)

Brown University and University of Stockholm

---

*Member, Federal Committee on Statistical Methodology

Acknowledgments

Chapter 1, 11, 111, and IV were written by Eli Marks, Charles Troob, and Wray Smith on the basis of Subcommittee outlines and discussions, as well as on the five case studies of formula programs. Chapter V (Subcommittee recommendations) is a joint product of the Subcommittee. The appendixes were prepared by individual members of the Subcommittee and their names are given both on their papers and in the table of contents. Research assistance was provided to the Subcommittee in 1977 by Henrietta Hyatt.

Members of the Federal Committee on Statistical Methodology

Maria Elena Gonzalez (Chair)
Office of Federal Statistical Policy and Standards (Commerce)

Barbara A. Bailar
Bureau of the Census (Commerce)

Norman D. Beller
Statistical Reporting Service (Agriculture)

Barbara A. Boyes

Bureau of Labor Statistics (Labor)

Edwin J. Coleman
Bureau of Economic Analysis (Commerce)

John E. Cremeans
Bureau of Economic Analysis (Commerce)

Marie D. Eldridge
National Center for Education Statistics (HEW)

Fred J. Frishman
Internal Revenue Service (Treasury)

Thomas B. Jabine
Social Security Administration (HEW)

Charles D. Jones
Bureau of the Census (Commerce)

Alfred D. McKeon
Bureau of Labor Statistics (Labor)

Harold Nisselson
Bureau of the Census (Commerce)

Monroe G. Sirken
National Center for Health Statistics (HEW)

Wray Smith
Office of the Assistant Secretary for Planning and
Evaluation (HEW)

Executive Summary

The Subcommittee on Statistics for Allocation of Funds prepared five case studies selected from the ten largest grant-in-aid programs that use data on population and per capita income. These five programs were then analyzed in terms of the variables "Need", "Capability", and "'Effort". These factors were selected by the Subcommittee as the key elements to be considered in analyzing both the formulas and data employed by grant-in-aid programs for allocation of funds. The report discusses the types of formulas used for allocation purposes, the required statistical data, and the impact of errors in the data on the actual allocation of funds. Based on the review of the case studies, the recommendations are as follows:

(1) That program goals be specified as clearly and completely as possible in the statement of purpose of each grant-in-aid act and that program drafters guard against over-specification of the statistical data and procedures to be used.

(2) That provisions be made for an active, continuous interface between legislative program drafters and the statistical community.

(3) That statistical and program agencies provide to program drafters an analysis of the sensitivity over time of formulas and of the statistics they incorporate so that possible effects on allocations can be anticipated. Also, that provisions be made for testing, monitoring, and assessing by program agencies of the performance of each specific formula or allocation rule prior to enactment.

(4) That legislative drafters and program designers be advised of data problems and the existence of statistical practices, as exemplified in the five case studies, which may lead to formulas with consequences that are generally recognized as undesirable.

(5) That a limited program of applied research and development be initiated to attack some critical problems and fill certain identifiable gaps in the present state-of-the-art of formula design.

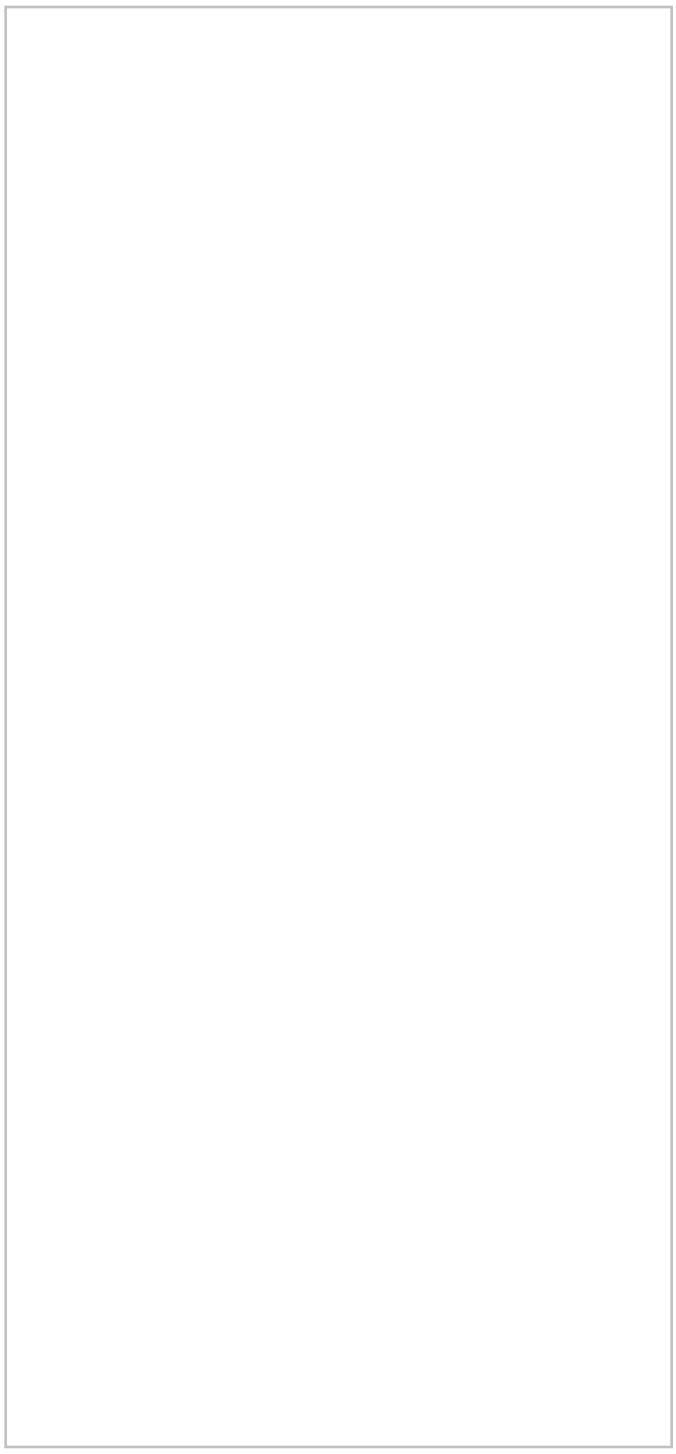
(6) That the Office of Federal Statistical Policy and Standards, with the assistance of the statistical agencies, designate a limited number of additional official statistical series for use in fund allocation. These would be kept as current and as accurate as possible for States and for local areas.

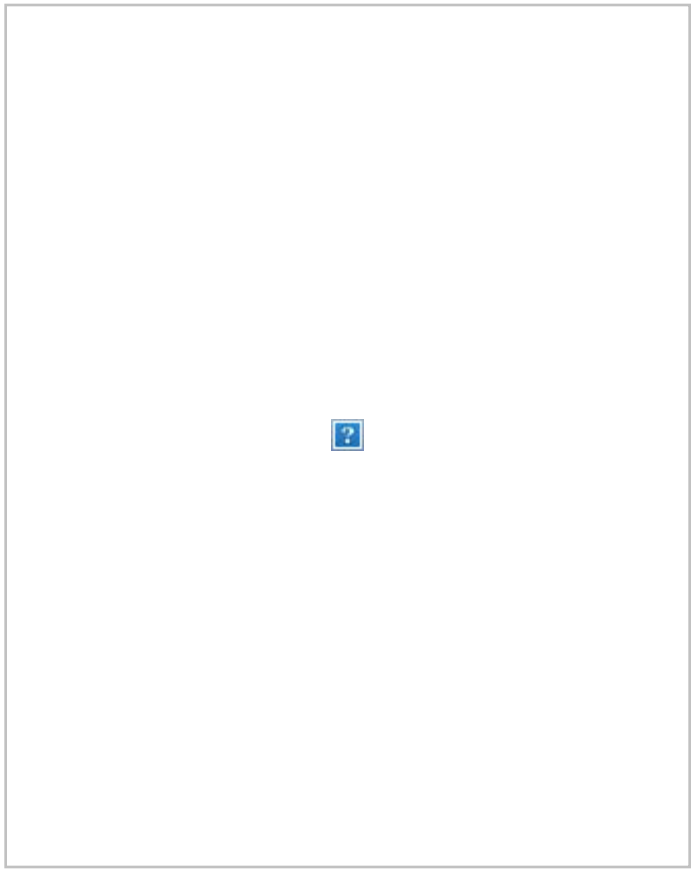
(7) That in tiered allocation programs comparable data be used for allocation to States, but policy flexibility be allowed for sub-State allocations. When

the Federal Government allows this flexibility it should be subject to the formulation of specific Federal statistical and administrative guidelines, concerning the designation of the responsible governmental unit for choosing among statistical series, for declaring the specific types of statistical series from which such a choice is permitted to be made, and for establishing administrative mechanisms for consideration of appeals from area governments.

(8) That since data errors are inevitable and since statistical resources are necessarily limited, priority be given to minimizing the very large errors which may occur in data used for the allocation of funds.

(9) That, to minimize the effects of data errors, eligibility cutoffs be such that there is a gradual transition from receiving no allocation to receiving the full formula amount.





CHAPTER I

Overview and Description of Allocation Techniques

Introduction

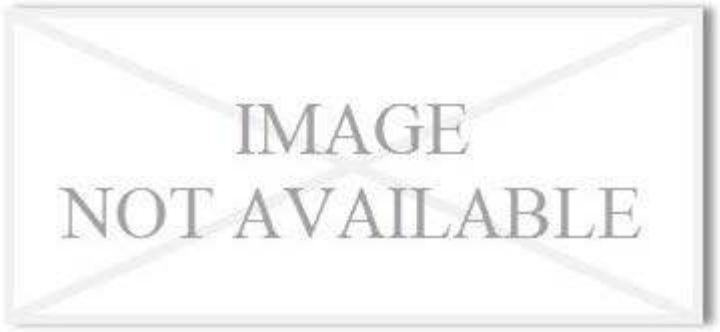
This report examines the formulas used in allocating Federal funds to States and local areas. To understand the behavior of these formulas, one must understand the various aspects of the data, such as definitions, methods of collection, and methods of analysis. The objective of the Subcommittee was to study from the statistical standpoint, possible principles or guidelines which could be used to insure that the intent of Congress is fulfilled in the allocation of Federal funds. For the purposes of this study it is assumed that whatever Congress specifies in the authorizing legislation for a grant-in-aid program on the manner of allocation of Federal funds is in prin-

ciple an equitable distribution, although anomalous and unanticipated results may emerge in some instances. In connection with the guidelines, the Subcommittee was also to identify possible improvements in statistical data and allocation processes that might be made either by better selection of the data, changes in data collection or tabulation methods, or statistical adjustments to compensate for known errors. The report is organized as follows: Chapter I gives an overview and description of allocation techniques; Chapter II examines the consequences of using existing data in allocation formula techniques; Chapter III presents the findings; Chapter IV discusses ways to reduce allocation errors; and Chapter V presents the recommendations of the Subcommittee based on its study of the deficiencies of existing data and allocation formulas and of possible alternatives.

We will now elaborate on some specific topics in these chapters. Chapter I and Chapter II are based on the five case studies presented as Appendixes A-1 to A-5 of this report. These five cases were selected from the ten largest grant-in-aid programs that use data on population and per capita income. In FY 1975, total formula grants for, all programs amounted to nearly 36 billion dollars. Fiscal year 1975 grants for the five case study programs range from 1.6 to 6.2 billion dollars and account for 47 percent of the total of formula grants.

Some of the findings of Chapter III are tentative and many of the recommendations of Chapter V are long-term goals which may never be achieved in the exact form presented. However, as an interim measure, some standard practices and guidelines are needed to aid policymakers and statisticians involved in constructing or revising allocation schemes for grant-in-aid programs. At the very least, such guidelines should warn practitioners away from some of the more dangerous, practices with disagreeable consequences that may be found in some existing formula

programs. For example, under some circumstances such guidelines might advocate the use of a particular population or economic statistic that was neither the most recent nor the most adequate from the standpoint of geographic detail but which had other statistical properties, such as stability from one time period to the next or uniform quality across



geographic areas. The Subcommittee believes that the development of some state-of-the-art guidelines will lead to a general simplification and increase in the transparency of allocation schemes to be adopted in the future.

The case studies show that many of the allocation formulas also contain constraints and special rules. For example, for administrative reasons it is necessary to impose some type of limitation on how often the allocations can be recomputed. Also, since the States and local areas must be able to prepare their own budgets and decide upon tax levies, capital investments, hiring, etc., some constraints may be imposed to prevent extreme year-to-year fluctuations in the allocation to individual jurisdictions. Sometimes, the restraints may prevent even moderate fluctuations in individual allocations.

Many of the formulas contain implicitly or explicitly a restriction designed to insure that every State or local area gets some amount. Sometimes this is coupled with a restriction on the maximum amount

to be allocated to any area. The limitation is usually not distinguishable from the limitations designed to damp or prevent fluctuations in individual allocations over time. To some extent, the restrictions may represent a well-justified distrust of the behavior of the allocation formula and of the appropriateness of the statistics used in it.

The Nature of Fed eral Grant-in-Aid Formulas

All of the allocation formulas studied deal with activities which are recognized as functions of State or local government but over time a feeling has developed that Federal assistance is appropriate to insure more equitable handling of the problem among local jurisdictions. That is, while there is recognition that the given function must be carried out locally and adjusted to the realities of local conditions, it is also recognized that financial resources available for handling the problem vary considerably among State and local governments so that it is appropriate for Federal funds to be used to supplement local funds.

Informally it is possible to adopt a helpful statistical paradigm for allocation formula research, in which the allocation is taken to be a function of "Need, Capability, and Effort", each of which is assumed to be at least approximately observable at the State or local level. There are, however, serious definitional and interaction problems imbedded in this model--the fact that a Need may appropriately have different components in two geographic areas, that taxable real estate and personal income may not give an adequate basis for Capability, that local tax revenue Effort may need to be analyzed in terms of the purposes to which the revenues are applied, and so forth. Frequently, one or even all of the factors in the model are defined neither in the statute nor in the legislative history or, if all the

factors are defined, the measures of Need, Capability, and Effort are inconsistent with the definitions or with each other. Thus, the terms are used to refer to statistical abstractions which apply only approximately (if at all) to the actual elements that make up a given allocation formula.

There are other elements of allocation formula problems, for example, the sensitivity of a particular formula to small, perhaps irrelevant, changes in the specified data over time. Some programs may require almost immediate reaction to the changes while, for other programs, insensitivity to short-term changes may be imperative. One wants the formula to respond fast enough to changing conditions but not too fast. Local government must be given some reasonable assurance of the general level of Federal funding they are to receive in future fiscal periods in order to keep local planning from becoming chaotic.

Another important question is the transparency of an allocation formula--can it be understood? Can citizens understand it? Politicians? Statisticians? Some formulas we have examined in existing Federal programs deserve to be called opaque--their behavior over time cannot be simply explained and may even exhibit some surprising and unanticipated results.

The general statistical approach used in this report conceives Federal grant-in-aid formulas as starting with some activity which the Congress perceives as properly a function of State or local government. In our statistical model we use the term Need to designate the activity required. For the purpose of the present report, Need is always to be understood in terms of the services (or goods) to be supplied--e.g., for food, shelter, etc., for AFDC (Aid to Families with Dependent Children); or police and fire protection, street and highway maintenance, etc., for General Revenue Sharing (GRS). While Need can be defined in money terms, this

definition involves the total amount required, whether or not that amount is available at the State or local level (or even whether it is available at any level). Thus, the Need in Title I, ESEA (Elementary and Secondary Education Act) might be defined as the total amount required to attain a given educational level in a local area, regardless of whether the funds are available at the local, State or national level, or perhaps, not at all.

Capability is used for an area's prospective ability to meet a stated Need--i.e., the possibility of meeting the Need from local or State (or private) funds. For example, Capability might involve the amount that could be raised by some (standard) taxing program whether or not actual tax revenues reach this level. Finally, Effort is used for the actual amounts available for the Need from local revenues. Frequently, Effort is measured relative to Capability.

Measures of State and local (relative) Need, and Capability of meeting the Need, are components of almost all allocation formulas. The measure of Need is often stated (at least approximately) in terms of the population to be served. Many allocation formulas also recognize that there may be considerable variation in the proportion of the available local resources actually devoted to meeting the Need and include some measure of Effort.

An important (but usually implicit) aspect of all allocation formulas is the time reference. Some programs are dealing with immediate objectives--to provide adequate food and shelter here and now. Others are dealing with a more distant time reference--to equip all of the Nation's children with the education and skills necessary to their functioning effectively in the Nation's economy as it is in 1977 (let alone as it will be in 1990). There is also a time reference or ability to meet a given Need. The

United States can, fortunately, meet our requirements for food and some sort of shelter immediately; but building sanitary, safe, and comfortable housing on the massive scale required in many communities takes at least 3 years and building even a partial rapid transit system for a major metropolitan center takes at least 6 years (from the time the system is designed and approved in principle).

Structurally, the formulas vary considerably. General Revenue Sharing (see Appendix A-1 of this report, "The General Revenue Sharing Program") uses the ratio of a measure of Effort (taxes as a proportion of aggregate personal income) to a measure of Capability (per capita money income). and multiplies this by (total) population. Essentially, this says that the share of a State or local area increases proportionally with the increasing population, increasing Effort, and decreasing Capability. In the General Revenue Sharing formula, per capita income serves as an indicator of Capability and total population as a measure of Need. This is equivalent to assuming that all jurisdictions have an equal Need per capita for the services covered by General Revenue Sharing. The General Revenue Sharing formula is complicated for sub-State distributions by lower and upper limitations on the per capita share of any locality (not less than 20 percent and not more than 145 percent of the average per capita share for the State). GRS allocations are also complicated at all levels by options relating to the specific measures to be used, but these do not affect the basic formula structure.

Like the General Revenue Sharing formula, the formulas of the other programs also involve measures of the basic factors (Need, Effort, Capability--with Capability entering inversely). However, the other formulas usually show some measure of Need explicitly. Often total Need in (dollars required) is used, so that population does not appear explicitly in the formula. Also, the formula may ignore Cap-

ability or use a single measure which reflects both Effort and Capability (with results which the Congress has found at times quite frustrating).

Thus, in the ESEA formula (see Appendix A-3 of this report, "The Authorization and Allocation of Funds Under Title I, ESEA") there is a measure of Need (the number of economically disadvantaged children multiplied by (a percentage of) the State average expenditure per pupil. Unlike General Revenue Sharing, this measure of Effort (per capita expenditure for the specified Need) does not relate it to Capability. However, an adjustment for low Capability is provided by substituting 80 percent of the national expenditure per pupil for the State expenditure per pupil, whenever the State expenditure per pupil is less than 80 percent of the national expenditure per pupil (presumably on the basis that low State expenditure per pupil characterizes the poorer areas, and was, therefore, a reflection of low Capability rather than of lower Effort or of lower unit costs for education of a given quality level). Corresponding to this floor on the allowance for per

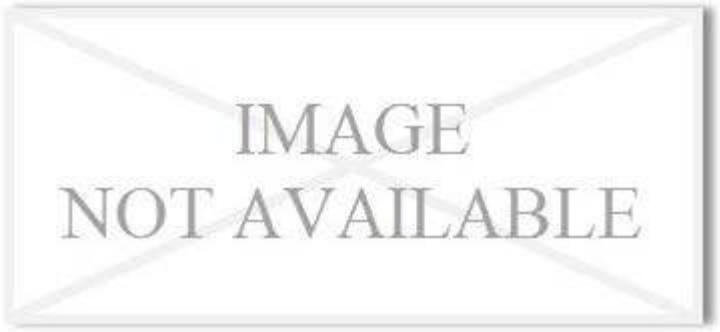
pupil expenditure, the (present version of) ESEA also provides for a ceiling of 120 percent of the national expenditure per pupil. There is, however, no allowance below the State level for variation among school districts in either Effort or Capability.

The AFDC formulas (regular and alternate) (see Appendix A-4 of this report, "Aid to Families with Dependent Children (AFDC) as a Formula Grant-In-Aid Program") resemble the ESEA formulas in starting with a measure of total Need times Effort--i.e., the total of money payments to families with dependent children plus payment for foster care. These payments are multiplied by the complement of a measure of Capability. However, in the regular AFDC formula there are provisions for using a fixed multi-

plier for part of the Federal payment (5/6 of the first $18 per recipient) and a maximum ($32 per recipient) above which no Federal reimbursement is made. As with the floor on per pupil expenditures in the ESEA formula, the use of a fixed multiplier for the first $18 has the effect of increasing the payments to States with very low Capabilities (measured by the State per capita income). Payments to States with very high Capabilities tend also to be decreased by the maximum of $32 per recipient in the regular formula. Since States have a choice between the two formulas, all but a few States with very low payments per eligible child elect to use the alternate formula based on actual payments and the computed percentage (Federal Medical Assistance Percentage, FMAP).

The formulas for CETA (see Appendix A-2 of this report "The Comprehensive Employment and Training Act") and CDBG (see Appendix A-5, "The Community Development Block Grant Program") are complicated by (a) a provision for a substantial proportion of the funds to be allocated on a discretionary basis and (b) so called hold harmless provisions for preventing sudden and drastic changes in an area's allocation. For CDBG, the hold harmless clauses provide for a gradual changeover from the previous (average annual) allocation level to the Basic Grant amount determined by the new CDBG formula. Communities whose new allocations would exceed their prior level would receive the full new allocation in the third year and the higher of the previous allocation and one-third or two-thirds of the new grant level during the first two years. Communities whose new allocation is less than their previous allocation would receive the previous allocation for the first three years of the program and would be cut back to the higher of the new level and two-thirds or one-third of the prior level during the fourth and fifth years, getting only the new allocation for the sixth year.

For CETA, the hold harmless provision involves use of a moving average of the current formula results and the previous period's allocation. This is similar to the exponential smoothing techniques used in economic predictions (for market planning, production inventory control, etc.) to obtain results which will reflect the real changes in basic economic conditions but will be insensitive to temporary fluctuations and disturbances. These averages are of the form:



The value of b is frequently determined (as is apparently the case for CETA) on the basis of expert opinion. There are methods for using past experience to determine an improved value of b (where a criterion for improved performance can be established).

CETA is further complicated by the existence of three titles with three different allocation formulas. All three formulas use unemployment level but one also considers adults in low-income families and has the hold harmless provision for smoothing short-term fluctuations; one has a lower limit for eligibility (6.5 percent unemployment rate for three consecutive months); and the third formula is a mixture of the other two.

While the primary effect of the smoothing (hold harmless) provisions of CETA is to reduce the effects of short-term fluctuations in unemployment, the smoothing provisions of the CDBG program are really phase-out/phase-in provisions designed to make a gradual transition to CDBG from the various

housing and community development programs it replaces. After the fifth year of the CDBG program,

the previous allocations are no longer considered in the formula.

The CDBG formula is different from the others considered in that it is additive. The basic allocation involves the weighted sum of three measures of Need, one of which is the total population of the area. The population measure receives a weight of 1/4 in the formula and the two other measures of Need (poverty count and number of overcrowded dwelling units) receive weights of 1/2 and 1/4.

The Advisory Commission on Intergovernmental Relations reviewed the allocation provisions of all federal formula-based categorical grants to State and local governments existing in 1975. Formula-based programs then numbered 146 out of 442 categorical programs. A review of the 146 programs shows that about 130 include some measure of Need; 41 programs include a measure of Effort; and 24 programs include a measure of Capability. These data show that there are few formula-based allocation programs that include all three measures: Need, Effort, and Capability. More than half of the programs include, only a measure of Need. However, there are many programs which combine two kinds of measures.

Definition and Measurement of Need in Grant-in-Aid Programs

As mentioned above, the term Need is used here to refer to the services which a given program is designed to provide. The measure of Need would usually be something proportional to the total cost of providing the services in a given jurisdiction.

The specificity of the Need to be met varies considerably between Federal grant-in-aid programs. In the examples of the Appendixes, AFDC probably has the most specific Need, that of providing adequate food, shelter, medical care, etc. for (non-institutionalized) children whose families are financially unable to provide for these needs adequately. At the other end of the spectrum is General Revenue Sharing where the Federal funds are to provide fiscal assistance for the general functions of local government.

The other programs fall in between AFDC and General Revenue Sharing with respect to the specificity of the Need to be met, but are, in general, nearer to AFDC than to General Revenue Sharing. For example, the ESEA is directed at establishing special education programs to help educationally deprived children. Most of the assistance is concentrated on improving basic skills such as reading, writing, and arithmetic but ESEA also includes funding for a wide variety of programs designed to meet other educational needs of educationally deprived children. There was also in ESEA, as originally conceived by Congress, the idea of a general antipoverty program to help poor people and poor school districts--e.g., the stated purpose of providing funds to school districts "whose ability to operate adequate educational programs is impaired by concentrations of low-income families."

The specificity of the aims of AFDC make it fairly easy to develop a measure of Need--i.e., the amount required to provide food, shelter, medical care, etc., to a child multiplied by the number of children in families who are financially unable to provide this care. The actual AFDC program accepts as the measure of Need, the individual State's definition of how much is needed per child and which families are too poor to provide this amount for their children.

General Revenue Sharing assumes a general Need based on the level of per capita income and the level

of taxes collected. That is, it is implicitly assumed that the amount a State or local government requires for general governmental functions is reflected in how heavily it is taxing its residents. The amount received under General Revenue Sharing is a direct function of the level of adjusted taxes and inversely related to per capita income squared. Population is only brought in at the upper and lower constraint levels.

ESEA uses as its primary measure of Need (1) the number of children in poverty families, (2) two-thirds of the children in non-poverty families receiving AFDC payments and (3) the number of children in institutions for neglected and delinquent children and in foster homes supported by public funds. This is directly in line with the purposes stated above. The measure of Need originally excluded the children described in (3) above but included 100 percent of the children in nonpoverty families receiving AFDC payments. AFDC uses, to measure Need, the total payments made for children in poverty families or foster homes (also see Appendix B-1, "AFDC Counts and ESEA Title I").

For CETA, the main measure of Need is the number unemployed. For States, the (expanded) CPS (Current Population Survey) estimate of unemployment can be used. Below the State level, unemployment must be estimated mostly from unemployment insurance data. A supplementary measure of Need for CETA is the number of adults in low-income families. The estimate of such adults currently used is derived from the 1970 Census of Population with no updating to reflect change since that time.

For CDBG the measures of Need are the poverty count and the number of overcrowded dwelling units. Both measures are derived from the 1970 Census of Population and Housing. The poverty

count is the number of persons in poverty families as shown in the 1970 census. An overcrowded dwelling unit is defined as one with 1.01 persons or more per room.

Measurement of Population, Capability and Effort

In the General Revenue Sharing and other formulas, total population as a measure of size enters implicitly in the use of a measure of total Need rather than per capita Need multiplied by population. Population is also used (explicitly) in the computation of per capita income which is the measure of Capability in the General Revenue Sharing and AFDC formulas.

Actually, in the CDBG formula, population is used as part of a measure of relative total Need rather than as a simple measure of the size of the area. That is, at each step, the allocation to an area is the average of its relative standing (ratio of the measure to the total for all areas in the class being allocated) with respect to population and number of overcrowded units (given weights of 1) and persons in poverty families (given a weight of 2). Since these three statistics are averaged in the formula, they must all be taken to represent measures of total area need for housing (relative, of course, to the total Need for all areas in the class). The AFDC and ESEA formulas use total population implicitly in the form of a measure of total Need (total amount of AFDC payments or number of 'educationally under-. privileged' children for the area).

Per capita income as a measure of Capability is used by General Revenue Sharing and AFDC. The General Revenue Sharing formula uses the reciprocal of per capita money income so that an area's allocation is inversely proportional to this measure of its Capability of raising the needed funds locally. The AFDC formulas use per capita income

(squared) to determine the percentage of AFDC payments to be met by State (or local) funds. This is called the State percentage and is subtracted from 100 percent to give the percentage to be reimbursed to States by Federal funds (subject to an upper and lower limit on the Federal Government's share of the AFDC costs).

In the Title I ESEA formula, per pupil expenditure is used as a measure of both Capability and Effort. Using per pupil expenditure as a measure of Effort, the formula provides for an area's share to go up proportionally to this Effort measure. However, using per pupil expenditure as a measure of Capability, there is a provision for increasing the allocation in States with low Capability--i.e., where the State expenditure per pupil is lower than 80 percent of the national average, 80 percent of the national figure is used in place of the State figure. At the other end, for States with high Capability the per pupil expenditure is reduced to 120 percent of the national figure.

Capability and Effort do not appear in the CDBG formula. As already noted, per pupil expenditure is used as a measure of Effort and of Capability in the ESEA formula. In the AFDC formula, payments made to poor families with dependent children and to foster homes are, in effect, taken as a measure of both Need and Effort. The General Revenue Sharing formula uses, as a measure of Effort, State and local tax revenues divided by aggregate personal income. This attempts to relate taxing effort to taxing capability.

Constraints and Time References

Formula constraints tend to be aimed either at obtaining a more equitable distribution of Federal funds (either between States or between localities within States) or at preventing large sudden changes in the amount a State or local area receives. Both

types of constraint represent an attempt to balance an allowance for real differences (in Need or Effort or Capability) represented by the main formula, against a concern that extreme values may represent peculiarities due to random occurrences (or temporary conditions) and defects in the formulas or the statistical data used in them.

General Revenue Sharing does not apply restrictions to the formula or data for allocations among the States but does provide for upper and lower limits on the allocation below the States level. The logic of this distinction is that (a) figures for States are probably subject to distortions for all States whereas there may be considerable variation in the

quality of the available data below the State level and (b) in the State allocation one is dealing with the entire range of non-Federal (general) government functions while local general government units may have a restricted range of functions.

AFDC places restriction on the Federal Percentage and the Federal Medical Assistance Percentage which apply to all jurisdictions. These operate to curb extremely high Federal payments to States with low per capita income as well as extremely low Federal payments to the richer States. In the regular formula, the restrictions are further buttressed by providing that the State will be reimbursed 5/6 of the first $18 per recipient paid out, regardless of the Federal Percentage (limited to a maximum of 65 percent), and will get zero reimbursement for amounts paid out in excess of $32 per recipient. For AFDC also the use of constraints can be justified on the basis of the failure of the statistics and the formula to properly reflect a balance of Need, Effort, and Capability that is equitable for all States.

The constraints imposed to prevent large sudden changes in the allocation to an area frequently take the form of exponential smoothing; i.e., using an

allocation which is a weighted average of the current computation and the allocation for the previous period. A constraint with a similar purpose (distinguishing between permanent changes and temporary aberrations) is the provision that, to be eligible for an allocation under Title II of CETA, an area must have an unemployment rate of 6.5 percent or more for three consecutive months.

Not previously mentioned are constraints on eligibility for a given program designed mainly to prevent the administrative nuisance and waste of handling a large number of extremely small grants, thus dissipating the available funds in areas where the amount allocated is too small to get an effective program going. This appears to be a relatively rare constraint but the provision of CETA Title II just cited appears to be motivated as much by this consideration as by the time series smoothing objective.

The question of short-term fluctuations in Need, and the techniques adopted to reduce their effects upon Federal allocations is closely related to the question of updating (keeping the statistics used in a formula current) and to the question of what is the appropriate time reference for a formula.

Time reference refers to the amount of updating which is appropriate to the particular program. Only one of the five programs examined in the Appendixes requires an immediate (i.e., month-to-month) time reference. This is the AFDC and, even here, since this is primarily a question of the Federal Government providing partial reimbursement to the States for money already spent, the only question is the Federal vs. State Percentage. In determining the Federal Percentage, the formulas and the data used in them are such that a redetermination once a year using per capita income figures for the preceding year should be quite adequate.

At the opposite extreme from AFDC with respect to time reference is the CDBG program. Here, the problem to be met is primarily an accumulated short-

age of adequate housing and community facilities. For example, the rate at which such housing can be planned, gotten into construction and completed, is such that there is probably a minimum of three years from initiation of a housing project to occupancy of the completed project. Only one of the components of the CDBG formula, the number of persons in poverty families, is likely to show very substantial changes over a three-year period and, even if one could obtain figures on this factor for the current year in order to recompute the CDBG entitlement of each area, changes in work already underway would not be possible; by the time housing based on the new formula is underway more current data would again be available to require a change in plans. Overcrowding has also diminished but the measure is not available for small areas on a current basis. Actually, the five-year period for transition from the old to the new housing formula is probably not excessive (it is, in fact, desirable to permit completion of work contracted on the basis of the old formula grants). At present, 1970 figures are being used for housing overcrowding and poverty in the CDBG formula along with 1973 population estimates. Some updating for future computations may be desirable but may not be as urgent for CDBG as for some other programs.

The appropriate time reference for General Revenue Sharing, CETA and ESEA is somewhat greater than for AFDC and considerably less than for CDBG. For General Revenue Sharing, figures for the preceding year probably provide a satisfactory base (from the standpoint of time reference.) for the current year's allocations. These can be provided for the GRS formula at the State level (probably with an accuracy almost as good as the 1970 figures). Below the State level, problems of providing current figures for all the GRS jurisdictions eligible becomes somewhat more questionable. Actually, it has been suggested that fluctations in GRS allocations from

one entitlement period to the next may influence unfavorably the fiscal policies of some local governments. Last year's figures are probably also satisfactory for the ESEA formula and would also be satisfactory for CETA, except for the hold harmless provisions of the program. These provisions, it is claimed, are so severe that allocations for a large part of the CETA money are based primarily on 1970 data, even where satisfactory current figures are available.

Allocation to Small Areas

All the programs mentioned here address a Need. In each program, there is a different governmental or quasi-governmental agency which is responsible for administering the funds and meeting the Need: State, county, and local governments in GRS, local education agencies in ESEA, prime sponsors in CETA, county welfare agencies in AFDC, and cities and counties in CDBG. Each program must devise a way of determining the fund level for these agencies and each program has a different method. GRS allocates to all eligible governments by formula. For sub-State areas, ESEA allocates by formula to counties. States then divide county allocations among the school districts within each county. The State procedure must follow Federal guidelines. CDBG allocates to SMSA, cities, and counties by formula. Other areas compete for funding, with total State and SMSA allocations determined by formula.

CETA has different procedures under each Title. In general, CETA is distinctive in that recipients of funding need not be preexisting governmental units: consortia of governments and agencies representing areas of substantial unemployment may apply for

funding. Once applications are accepted, the money is divided up by. formula.

In AFDC, unlike the other programs discussed, there is no ceiling on the Federal contribution. County agencies expend whatever is appropriate under State law; the reimbursement rate varies by State. Caseload data primarily determines the level of Federal contribution to each area.

CHAPTER II

Why Existing Allocation Formula Techniques Do Not Fully Achieve the Stated Objectives of Federal Programs

Problems of Choice of Formula Structure and Constraints

In view of the examples of formula creation and use found in the five case studies, it is clear that the typical allocation formula has a complex structure entailing the identification and selection of various options. For this reason, a decision to adopt a specific formula involves--at least implicitly--a series of distinct prior choices. An inappropriate decision at any of these choice points may lead to a formula which results in allocations that do not reflect congressional priorities. We realize that such choices are, as a result of the interaction of individuals and committees, often judgmental and sometimes not made in a fully logical order. Nonetheless, there are some necessary elements in any such specification process which we feel need to be made explicit as a basis for understanding problems and limitations of formula selection.

The first choice involves the definition and measurement of Need. As discussed in Chapter I, any proposal for a Federal grant-in-aid program that is to involve a formula mechanism is motivated in some fashion by a perception of a Need. A working definition and some measure of that Need must be

adopted, whether or not there is full understanding or agreement on all of the dimensions of Need. For example, in the first enactment of Title I of the Elementary and Secondary Education Act it was recognized that school districts serving large numbers of low-income children were in some need of special assistance. While there was general agreement that such school districts needed more money, there existed by no means any fully consistent statement of the nature of special burden which low-income children represented. In fact the statute related the level of funding to the number of low-income children but left it up to individual school districts to assess the requirements of their children and to plan programs accordingly.

A measure of Need that is perfectly congruent with the definition of that Need is almost never available. As a consequence the program designer must resort to some proxy indicator, and the choice of a suitable proxy is by no means trivial. Surrounding Title I, for example, there was considerable debate over the proper measure of low-income status, and the measure was in fact improved in 1974. Yet, the 1974 debate did not resolve all questions concerning the appropriate measurement of the target population or even settle its definition. Dissatisfaction with the criteria of disadvantage embodied in the present formula led the Congress to commission a study at HEW on the measure of poverty (which was completed in 1976) and a related set of studies to be carried out by the National Institute of Education of the feasibility and probable impact of using measures of educational rather than economic disadvantage for Title I ESEA fund allocations.

As noted in Chapter I, in adopting allocation formulas Congress frequently takes into account some measures of what we have termed in this report Capability and Effort. These are, if anything, more difficult to define than Need, and may involve problems of measurement as well. After the program

designer has set forth a working definition and measure of each of these elements the actual process of formula construction properly begins. At that point there is a wide range of possible allocation formulas which might be constructed as well as a variety of possible constraints and special rules.

A central question that must be answered by the program designer is in what way the resulting allocations should vary over the range of possible values of the measures of Need, Capability, and Effort, and also reflect considerations not accounted for by these concepts. In some cases the difficulty of ameliorating a social problem may be proportional to the measure of Need, so that a linear allocation formula would be appropriate. In other cases a non-linear relation between the allocation and the Need measure may be called for. If the designer wishes to take into account Capability or Effort, then the maximum and minimum allocations for a given Need must be decided in relation to the expected range of measured Capability and Effort. There may also be other desired patterns of allocation to meet policy purposes other than those reflected in the measures of Need, Capability, and Effort.

Once these issues are settled, the formula can be

constructed. This process necessarily includes both policy and technical considerations. The central technical problem is the choice of a mathematical structure which in some sense best utilizes available data to produce the desired allocation pattern. As discussed further below, there are additional issues of data limitations, of interactions between the formula and data, of the dynamic properties of the formula, of its understandability to the public, and of its computational efficiencies.

The essential elements in the choice of a mathematical structure are as follows: (1) the class of the formula (e.g., additive as in the CDBG program or multiplicative as in General Revenue Sharing); (2)

the weights or scale factors to be applied to each of the terms in the formula (e.g., giving unit weight to relative population and to overcrowding, double weight to poverty in CDBG); and (3) the specifications of constraints, if any, on either particular variables or on the resulting allocation (e.g., floors and ceilings on the cost factor and hold-harmless levels on the allocation in Title I ESEA). The statistical consequences of these choices are often not fully understood by either statisticians or program designers. Although the design sequence can be described as a set of logical choices, the sequence and timing of such choices will vary from program to program. In addition, both the valid demands of the political process and the primitive state-of-the-art of formula practice lead to choices at every stage of the program design whose full statistical and distributional implications cannot be foreseen at the time they are made.

For example, floors and ceilings or other types of constraints involve in some sense a distortion of the ideal allocation. As noted in Chapter I, the setting of such constraints is sometimes an attempt to limit annual variations in allocation levels and sometimes an attempt to modify a less-than-ideal formula by making sure that no one gets too much or too little. In either case, constraints may influence allocations more strongly than they were intended to. A striking example of this effect is seen in General Revenue Sharing under which townships with minor governmental functions are guaranteed a sizable minimum payment--a consequence that was not generally, anticipated at the, time the law was passed.

The complexity of the task of selecting a formula structure leads in practice to other problems. Every allocation formula represents a simplification of the real world. We have just pointed out that constraints distort an ideal allocation, but the very notion that ideal allocation could be described for reference purposes implicitly assumes that we are willing to

determine just how much of the fine-grain complexity of the real world should be captured in such an ideal formula. While technicians might reach some consensus on the attributes of an optimal degree of simplification, no statement of principles based on such a technical consensus would be immune from criticisms that some important aspect of reality was omitted from a formula designed according to such principles. This point serves to reinforce our recognition that formula building if it is to be successful in implementing legislative goals should not be the exclusive province of either the technician or the Politician.

An important implication of the need to accommodate both political and technical considerations is that an allocation formula should be comprehensible to all parties involved. The policymaker needs to understand more about an allocation formula than just bow much money it allocates to various jurisdictions this year. The formula should be transparent enough to support direct analysis of its distributional effects --across States and within States--at a point in time and as well as over time. The recipient--whether local official or ultimate beneficiary--should at the very least be able to verify the correctness of his allocation. For example, the General Revenue Sharing formula is extremely complicated, both in the determination of State allocations and in the division of funds within States. The procedure for allocation to States, which resulted from a compromise between House and Senate, combines two formulas to give each State the higher of two computed allocations. Because there is a fixed total appropriation for the program, the actual computation must be carried out iteratively, and only expert analysts can estimate the impact of even very simple changes in the existing formulas. Thus we see that lack of transparency in the formula for an ongoing program can be an important deterrent to meaningful attempts at reform of existing programs.

Problems Arising From the Nature of the Data Used and From Interaction of Data aid Formulas Over Time

However difficult it may be to understand and evaluate the performance of a formula at one point. in time, the task of foreseeing and assuring good performance through time is even more difficult. There seem to be three issues: (1) The formula may

require data which cannot be updated frequently, and the degree of distortion caused by the use of obsolescent data can neither be bounded closely in advance nor estimated precisely at the time current allocations are made; (2) Statistics which can be updated for formula use may slowly or suddenly depart from their historical behavior and from their assumed stable relationships with other variables; and (3) The social or economic problem to which the program is directed may evolve in such a way that the measures chosen to represent Need, Capability, and Effort may cease to be the most relevant measure available.

All of these issues are illustrated by the history of the measure of economic disadvantage used in Title I, ESEA. This measure has been and continues to be the sum of counts obtained from various sources. Census counts for 1960 were a major component in the Title I measure from 1965 until 1973, by which time they were hopelessly out of date. Annual counts of children in families receiving a high level of AFDC payments departed from their historical behavior shortly after Title I was enacted, as a result of an unprecedented increase in the AFDC caseload and of the onset of an unforeseen price inflation. While in 1965 the AFDC counts represented about ten percent of the total Title I measure, by 1974

they were sixty percent of the total measure. While some growth in the importance of the AFDC counts might have been expected in 1965, it was not anticipated that they would become the dominant component. While it could have been predicted that the fixed dollar family low-income threshold specified in Title I ($2,000), would become quite inappropriate upon the introduction of 1970 income data, Congress took no action to revise this specification until the effects of the use of the old cutoff with 1970 data were evident in the 1974 Title I allocations (see also Appendix B-1, "AFDC Counts and ESEA Title I").

Our third issue is illustrated by the rapid expansion of in-kind transfer programs, such as Food Stamps and Medicaid, whose income equivalent is not currently counted in family income statistics from the decennial census and the Current Population Survey (CPS). Depending upon the distribution of in-kind benefits, they might bias relative measures of low-income status across geographical areas. The degree to which they depart from such a uniform relationship with money income is not fully known, but the magnitude of these in-kind programs raises the possibility of serious bias.

Both here and at earlier points in this chapter we have reviewed. issues which demonstrate that data and measurement limitations may dominate all other considerations in formula design and assessment. As we have stressed before, no measure can be perfect in all respects. One of the most difficult tasks in program design is to determine in advance whether a measure will prove to be at least minimally acceptable. A recapitulation follows of the different ways in which an operational measure may fail to fulfill the objectives of the program drafter.

(1) Lack of fit between a measure and the real world phenomenon it is intended to portray.

An inappropriate measure may be chosen because of its familiarity or its intuitive appeal. Within

CETA, for example, the local unemployment rate is used both to measure the need for public employment, of which it is probably a satisfactory indicator, and to measure the need for job training, for which there may well be more appropriate though less familiar measures. The overcrowding index used in the CDBG program is a good example of a measure, the intuitive plausibility of which may exceed its suitability to the program in question. What makes the index attractive, however, is that it conveys some information about whether the inadequacy of housing leads to hardship. This possible relationship is certainly something one would want to measure in a Federal housing program. The overcrowding index, though, may be inferior as an indicator of the quality of the kind of housing generally available to the poor when compared to some possible physical measure of housing stock quality which contains no overt reference to occupancy. However, no simple measures of the physical quality of housing is available at this time. Perhaps CDBG should consider developing a more comprehensive measure of housing needs in which the overcrowding index is only one of the factors.

As the case study on General Revenue Sharing indicates, the use of per capita income as a measure of Need has been criticized despite its obvious virtues of familiarity and general plausibility.

(2) Accuracy of a measure for the geographical area it applies to.

This presents a problem for all programs which require formula allocation to small areas. The unemployment data for CETA and the poverty data for Title I, ESEA are pertinent examples. In the case of CETA, the flexible definition of labor market areas, although perhaps desirable for policy reasons, is

made less desirable because of the inadequacy of the statistics from which area Need must be calculated.

In the case of Title I ESEA the congressional intention to allocate directly to school districts was thwarted by the inadequacy of school district poverty data, and instead allocations were made to counties, with the States being responsible for subcounty allocation to school districts.

(3) Stability of a measure in relation to the frequency of updates.

Data which are expensive to gather as well as subject to considerable variability through time may not be cost-effective for allocation purposes. This is the chief obstacle to the generation of small area price deflators which could be used to adjust grant levels to local price differences.

CHAPTER III

Subcommittee Findings

In this chapter, we will present four major findings together with some illustrations.

Finding No. 1

There are very real difficulties in translating congressional intent into statistical terms.

We will illustrate this finding by reference to the Community Development Block Grant program authorized by the Housing and Community Development Act of 1974.

a. Section 101(c) of the Act states that "The primary objective of this title is the development of viable urban communities, by providing decent housing and a suitable living environment and expanding economic opportunities, principally for persons of low and moderate income."

b. The section goes on to say that the CDBG Federal assistance is for the support of com-

munity development activities directed toward certain specific objectives, including "the elimination of slums and blight and the prevention of blighting influences and the deterioration of property and ...facilities...; the elimination of conditions which are detrimental to health, safety, and public welfare, through-code enforcement, demolition,...; the conservation and expansion of the Nation's housing stock ...; the expansion and improvement of the quantity and quality of community services...; a more rational utilization of land and other natural resources...; the reduction of the isolation of income groups within communities and geographical areas...; the restoration and preservation of properties of special value for historic, architectural, or esthetic reasons."

c. As described in the CDBG case study, the allocation and distribution of funds is specified in the Act on the basis of a three-term additive formula counting population, poverty (weighted twice), and housing overcrowding--where the count for, say, a metropolitan city is entered as the numerator in each of three ratios with the denominators being the counts for all metropolitan areas. In the framework of our report this is a Needs formula with no explicit components for Capability or Effort.

d. Congress apparently felt that the extent of poverty and housing overcrowding were reasonable surrogates for its target population (persons of low and moderate income) and for the conditions it hoped to alleviate (slums, blight, inadequate services, etc.). They did not try to legislate the use of some direct measure of housing quality or service adequacy. But a paradox remains: Two communities of the same size, poverty count, and overcrowding index might have, to an impartial observer,

two quite different levels of adequacy of housing stock and services.

e. As can be seen from the above discussion, it would be very hard to construct a formula that would adequately operationalize the goals of the Act. It should be noted that Congress is expected to reconsider the CDBG formula during the 1977 session, partly in recognition of some of the problems outlined above (1.d).

f. The CDBG program illustrates the potential conflict between policy objectives and the rationalization of formula and data requirements. In this case, the broad objectives make it difficult to define and measure Need in the program formula. Congress set up CDBG to consolidate a number of categorical programs. One objective of CDBG was to allow for considerable local discretion in the specific purposes to which the allocated funds would be applied. Accordingly, a large number of program goals were recognized, and, purposely, there was no ranking of the various possible objectives.

Finding No. 2

Current administrative and statistical practices do not always deal adequately with the problems that have been identified in Chapter II.

a. A good example of "why... existing allocation formula techniques do not fully achieve

the stated objectives of Federal programs" may be found in the methods for counting Title I ESEA eligibles.

With regard to the problems arising from the nature of the data used, the law specifies a determination of the "number of children aged

five to seventeen, inclusive, from families below the poverty level on the basis of the most recent satisfactory data available from the Department of Commerce for educational agencies (or ... counties) ... utilizing the criteria for poverty ... in the 1970 Decennial Census." There is a parallel provision for counting some disadvantaged children (AFDC recipients, etc.) above the poverty level.

(1) The "most recent satisfactory data" may not in fact be recent enough to be satisfactory. Furthermore, in spite of the age of the data, no provision has been made for a reinterpretation of the counts in a way that might constitute a partial adjustment for time effects. For example, instead of the cohort aged 5 to 17 in 1970, the cohort aged 0 to 12 in 1970 (which was aged 5 to 17 in 1975) might be considered as a relevant reference group for current allocations.

(2) The argument is sometimes made that the Title I formula is partly protected from obsolescence by the inclusion of the AFDC factor which in some sense can update the eligibility counts, even if the poverty counts cannot be updated. As pointed out in Appendix B-1, the AFDC component is only about 7 percent of the total and is distributed among States and counties very differently from the poverty count--in either 1970 or, say, 1975.

b. Another example is provided by the General Revenue Sharing program, which has been operational for more than six years. Much of the criticism of the program has been focused on how well the formula structure reflects the needs of the recipient localities. The GRS program distributes funds to approxi-

mately 39,000 jurisdictions, the great majority of which are areas of population less than 2,500 in the 1970 Census of Population. Because of the complexity of dealing with different kinds of local governments, and the severely limited data available for this purpose, GRS has used a uniform procedure that treats similarly governmental units with very different sets of responsibilities.

In addition, the use, of GRS as a counter-cyclical device is hampered by considerable data lags. Despite the procedures involved for updating census money income (one of the elements of the formula mandated by the Act), based on the more current IRS wage data (used in conjunction with BLS county and State wage data) and the BEA county personal income data, there is still a lag of several years between the reference year of the data used in the formula and the year in which the allotment is made. Even if the currentness of the inputs could be improved enough to appreciably narrow the gap it could not be done without introducing other difficulties. Although improvements in the formula have been proposed, introducing other elements purported to be better indicators of Need, these other elements also can be measured only with several years' lag, and may not even be available for smaller areas or only with some sacrifice of precision.

A further criticism has been that occasional sharp fluctuations in the size of the allotment for a given area from one period to the next, caused by unusual variations in the data inputs, tend to hamper long-range planning by the recipient governments for efficient use of the revenue sharing funds. However, changing a formula structure which has been in operation and has come to be generally accepted by all levels of government could be more disruptive

than the occasional random fluctuations in allotments encountered with the present formula.

Finding No. 3

The nature of the statistical problems arising in formula programs is such that present knowledge does permit the identification of at least some interim principles. There are some existing programs for which the existing formulas or allocation rules appear to be satisfactory from a statistical standpoint.

a. One example may be found in the AFDC case study. Whether or not the resulting reimbursement levels to the individual States are completely appropriate is a matter for Congress to consider from time to time. But there are no apparent statistical bases for concluding that the resultant reimbursements are inappropriate. There is an inverse relationship between per capita income (PCI) and reimbursement rates. If this were adopted simply as a fair relationship it would be hard to argue that it is not. By that standard there would appear to be no serious problems with the current practice. If the inverse relationship were interpreted as an incentive device to get the poorer States to set up programs comparable to those of the richer (higher PCI) States with higher benefit levels, then that Federal purpose would have to be seen as not fulfilled by the matching rate rules, since the poorer States have not so responded.

b. Another example concerns the Comprehensive Employment and Training Act of 1973. The major portions of the funds allocated by formula under Title I of that Act are distributed in a manner that incorporates several

elements that are sound from a statistical standpoint:

(1) The units to which funds ire allocated, the prime sponsors, are large (100,000 or more population) and thus avoid the problems associated with the preparation of estimates for very small units.

(2) The prime sponsors are defined in terms of units of general local government. While these may be combined into various configurations, this eliminates the difficulties associated with the development of estimates for neighborhoods or other parts of cities or counties that do not have an established geographic definition.

(3) The unemployment data used in the allocation is based on annual averages. It is, therefore, not subject to seasonal influences and the distortions that they can inflict on the allocations. The use of annual averages is, in a sense, an example of the use of the best available data from a single standard source--the Current Population Survey (CPS). However, the CPS is used only for the States and 30 SMSA's and 10 central cities; a problem remains for large counties and large cities. Moreover, the formula incorporates legislative determination that while all areas need manpower services, the need is greater where the number of unemployed is higher. The distribution is therefore based on the number of employed.

(4) The problems of administering a continuing program of manpower services with a shifting financial base are recognized by providing a floor based on the preceding year's allocation below which the funding of the current year cannot fall, and a ceiling above which the allocation cannot

go. Title I, CETA avoids wide year-to-year swings in the allocations received by prime sponsors. It does this both by distributing funds largely on the basis of the previous year's allocation, and also by providing floors and ceilings, based on the previous allocation, beyond which the current allocation cannot go. This facilitates the chief objective of Title I--to provide a continuing program of manpower services--by keeping funding levels relatively constant and predictable.

c. The third example concerns the sub-county allocation system in Title I ESEA. This is a creative approach to the problem of allocation to small areas, in this case to school districts. The data used in the formula to allocate to counties--1970 poverty counts, special AFDC tabulations, and counts of neglected, delinquent, and foster children--are not currently available at the district level to the Federal Government. States therefore have been given the right to allocate county funds to the school districts in each county, using the most recent appropriate data. The Federal guidelines recommend census And AFDC data, but States may choose among a number of data series. While not without problems, the system appears to work relatively smoothly. One benefit of this system is that questions about the correctness of the data for very small districts can be raised as well as resolved locally, by people familiar with the actual conditions.

Finding No. 4

The present state-of-the-art will not permit formalization of a fully definitive or wholly acceptable

set of statistical rules for formula programs. In view of the present gaps in our knowledge there is a need

for some short-range applied research on problems of allocation statistics. For example, while the use of quadratic loss functions (minimizing the mean squared error) is well established, there appears to be a need in formula research for the use of asymmetric loss functions. At present there is little readily applicable theory and some research is needed soon on this topic as well as on related problems in approximation theory (also see Appendix B-5, "An Agenda for Basic and Applied Research Problems").

CHAPTER IV

Ways to Reduce Allocation Errors and Inequities

Introduction

It is usually easy to arrive at a consensus that the allocation of funds under any given program is inequitable. However, it is often difficult to get any agreement on the nature and location of the inequities and even more difficult to get agreement on how to correct the inequities. There are, though, some aspects of allocation formulas and the data used in them which lead to substantial discrepancies from the intent of the original legislation. This chapter addresses this type of problem.

1. Problems arising from the data used. In connection with data used for allocation, there are rather complicated trade-offs among five factors, three of which are relatively well understood (at least we think we know what they mean), namely bias, variance, and cost. The other two are the timeliness of the data (the time-frame of the data) and the appropriateness. The appropriateness can be defined as the extent to which the

concept one is using (no matter how well or poorly it is measured), approximates the thing that one really wants to measure.

a. Updating. Before discussing the interaction of the five factors, a few observations are in order on timeliness and the question of updating statistics for use in allocation formulas. As noted in Chapter I (p. 6) the appropriate time reference (timeliness) varies from program to program. In the field of government action, one can distinguish between programs to meet immediate (and very time-dependent) requirements and those designed to deal with situations which change relatively slowly over time. In the first category are those welfare and unemployment insurance programs designed to deal on an emergency basis with immediate problems. The impact of this type of problem on any given area at any specified time is largely unpredictable. Here one is dealing primarily with questions of accounting for funds after they have been spent, rather than of allocating funds to specific areas. This type of problem is best handled by providing for a central pool of Federal or State funds which is drawn upon as required locally. To the extent allocation of Federal funds is involved, the statistical problem becomes one of determining the amount of allocations appropriate to maintaining the State or local pools of funds at (legislatively) specified levels over a time period of a year or more.

   Thus, even when there is the requirement for immediate action that varies locally from month-to-month (and even week-to-week), updating of data used for allocations is not necessary more frequently than once a year. Where the basic economic and social conditions at which a program is aimed change

slowly, updating statistics every 2 or 5 or even 10 years may be adequate. In the case of programs involving massive training or building programs (highway and mass transportation programs, slum clearance, teacher training or retraining of individuals in declining industries) frequently updated figures, even for rapidly changing situations, may be of little appropriateness for fund allocation, since a large portion of the work in progress must be completed even though plans for future work may need drastic revision.

b. Trade-offs. The total population of an area is a factor in many allocation formulas and the problem of making estimates of population illustrates the trade-off among bias, variance, cost, timeliness, and appropriateness. The cheapest estimate that might be in any sense acceptable is, of course, the population of the area based on the most recent decennial census. However, for some allocations, the decennial census figures are out-of-date by the time they are published. Even if one uses the hand counts (announced locally immediately after completing the census field work) and takes the risk of major differences from the final revised figures, decennial census figures are at least 10 years old by the time the figures for the next decennial census are available.

The recent authorization of quinquennial

censuses will somewhat reduce the problem of updating population census figures; but the cost of taking a 100 percent census will almost certainly mean that the 1985 enumeration will be on a sample basis. While the sampling biases and variances of a sample census will be small for most states and for

major metropolitan areas, the sampling errors for small areas will be, at a minimum, a source of considerable controversy (e.g., claims that "my city or county was 'robbed' in GRS allocations"). Even for the largest areas there can be considerable dispute since, while the relative sampling errors will be small, the absolute errors and the absolute sums of money involved may be substantial.

For some uses, updating population figures every five years will be considered unsatisfactory; there is pressure for annual and biennial sample surveys and for the use of more current statistics derived from administrative records (birth and death registrations, income tax returns, school enrollments, etc.). The unit costs of a sample survey are high and, for a number of quite valid reasons (difficulties with privacy, confidentiality, public resistance, availability of satisfactory personnel), are increasing, in spite of improved survey techniques and generally improved overall efficiency in the conduct of sample surveys. Even well funded and well conducted sample surveys (e.g., the 1976 Survey of Income and Education) are restricted to small samples and also require the use of clustering in order to minimize travel time and other nonproductive expenditures. Small, highly clustered samples mean large sampling variances even for some relatively big areas, and also mean that many small areas will have no sample households at all.

Using administrative records to update the population involves major problems and can involve serious biases. Applying statistics from birth and death registration records to the previous census should produce reasonably good figures for areas which have had very little in-or-out-migration since the cen-

sus. For the areas with relatively heavy (net) migration in the 1950's and 1960's (e.g., California, Florida, Arizona, Nevada, Alaska, most metropolitan areas west of the Mississippi, rural areas of the South Central and West North Central States); estimates based on births and deaths tend to be improved by making an adjustment based on past migration trends--e.g., using the average population change in any area due to migration (total population change less births plus deaths) from 1960 to 1970 as an estimate of the annual change due to migration since 1970. Adjustment for past migration trends usually gives improved estimates for the areas with substantial past in- or out-migration but it does not allow for the second (and higher) order derivatives of the population change curve for an area. Such an allowance can be made by using a curvilinear regression on past migration trends but this involves either using still earlier censuses (e.g., the 1950 and the 1960 censuses) and intercensal births and deaths to estimate net migration since 1970 or obtaining estimates of intercensal populations. While the use of past migration trends (linear or curvilinear) will improve most estimates of current population based on births and deaths, it results in poorer estimates for some areas because of the biases and variances of the estimates of past migration trends as well as changes in the shape of the population growth curve since the last census. While sudden and drastic changes in the shape of the population growth curve of an area are rare, they occur (e.g., the decrease in California population growth rates between 1960 and 1970) and in these cases there may be serious biases in the population estimates in spite of the adjustment.

Similar difficulties of bias and variance occur in the use of estimates based on other administrative records. For example, population estimates derived from income tax returns do not provide for persons who did not file a tax return for the given year. Partial adjustments for these omissions could be made by using supplementary sources (e.g., W-2 files, files of welfare families) but adjustmerits (e.g., determining how many persons are represented by W-2 forms to adjust for the cases where the income recipient did not file a 1040 return) are difficult and the estimates will still be deficient for other reasons (e.g., individuals may not be shown as dependents or income recipients in any source).

The estimates can be improved by using the administrative records to estimate change in an area since the last census (rather than the current population level) and by applying this estimated change to the census figure for that area. Similarly, the percent change since the census in school enrollments can be applied to the census population to produce a current estimate. One can also use a combination of change in income tax and enrollment statistics to estimate current population by applying the regression of the census population on census-year tax returns and school enrollments to current-year tax returns and school enrollments.

For some allocations updating the census population counts may be unnecessary. However, even for these cases, there is a question of biases in the counts. The Census Bureau estimates that, even after very vigorous efforts (and very large expenditures) to

obtain 100 percent coverage in the 1960 and 1970 Censuses, there were undercounts of 2.7 and 2.5 percent. It is likely that census techniques in 1980 will have to be improved, and efforts and expenditures per capita will have to be increased even to attain the 97.5 percent coverage level of 1970.

The trade-offs of cost, bias, variance, and appropriateness are particularly evident in the area of control and estimation of census coverage error. There is, for example, the question of trying to reduce differentials in coverage among areas and subgroups. For several reasons Black, other minority, and low income groups are more difficult to enumerate completely than the rest of the population. The coverage problem is particularly acute for certain types of areas, e.g., sparsely settled rural areas and ghetto areas in large cities. Frequently improving coverage of the poorly enumerated groups and areas requires very much higher census expenditures per household, and this, in turn, raises the question of reducing expenditures elsewhere or increasing total census costs. Reducing expenditures elsewhere may mean slightly higher overall bias in order to decrease the differentials in coverage bias.

The handling of imputations in a census also provides an example of the problem of balancing variance, bias and cost. Because of imperfections in the most well-designed census, problems of imputing for known errors always arise. Thus, discrepancies between the area hand counts and the initial machine counts have existed for every census where tabulation machines or computers have been used. These may be due to errors in addition, to failure to count some census sheets or lines, to errors in the hand count, to questionnaires lost in transit to the proc-

essing center, to questionnaires misfiled and lost in the sheer mass of paper, to failure to punch or film questionnaires or groups of questionnaires, to errors in punching or optical sensing of the questionnaires, etc.

In the 1970 Census possible errors in the census counts were also signaled by the Vacancy Recheck and PEPOC (the Post-Enumeration Post Office Check). These involve checking units reported as vacant to determine whether they were, in fact, vacant, and having the local post office check the census listing for possibly missed households for those areas where a post office check was not done before the enumeration.

Possible census errors detected by discrepancies between hand and machine counts or by. a vacancy recheck or by PEPOC can be met by:

(1) ignoring the possible hand count or vacancy recheck or PEPOC results,
(2) tracing the errors and making corrections,
(3) reenumerating areas or units where errors are detected, and
(4) imputing more correct values.

All of these methods were used in the 1970 Census. Small discrepancies between hand and (initial) machine counts were ignored; some misfiled questionnaires were detected and the appropriate counts corrected; a sample of vacant units and a sample of the enumeration districts where PEPOC showed possibly missed households was reenumerated; the results from the Vacancy Recheck and PEPOC samples were used to impute corrections for the nonsample vacant units and the nonsample enumeration districts of PEPOC; and imputations were also

made where the initial machine count was well below the hand count and investigation

confirmed that the hand count was more accurate.

With respect to updating and coverage error and imputations, possible solutions involve some compromise among the five factors. Thus, in a quinquennial sample census there may be a satisfactory compromise between the low cost, low variance, and poor timeliness of using decennial population figures, and the increased cost, high variance and bias, and good timeliness of using annual sample survey estimates. Essentially making imputations based on a sample check (as was done for the Vacancy Recheck and PEPOC) is a compromise between the bias and low cost of not correcting for the known census error, and the lower bias and higher cost of trying to follow-up and (re)enumerate all of the questionable cases.

A form of compromise which seems particularly desirable for the problems of updating and adjusting for undercoverage is the use of low bias and high variance data from a small sample study to adjust higher bias but low variance estimates from a larger, scale study. Thus, for updating population data we could use the high variance and low bias of changes measured from, a small annual or biennial sample survey to correct the bias of (zero variance) statistics derived from administrative records. By substituting regression of the changes on other characteristics, we decrease the variance of the resultant estimates with some (hopefully small) increase in bias. In estimating undercoverage, we can correct biased estimates from a large sample survey by the low bias results of

small samples of administrative records (from IRS, Medicare, driver's license files, etc.) matched to the census. By using regression techniques, we can obtain considerable reduction in the biases of the estimates from the large sample source and avoid the high variances of the estimates for individual areas in the small sample study. The impact of errors on allocation is discussed in Appendix B-2, "Technical Notes on Sensitivity Analysis". Raking as a statistical adjustment procedure may be used to reduce error in data (see Appendix B-4).

c. Data Comparability. Where different areas (States, counties within a State, school districts within a county) are in competition for a share from the same pot, equity dictates that the allocation data for the competing jurisdiction be as nearly comparable as possible. Comparability is usually attained by taking the estimates for all competing jurisdictions from the same source. Thus, the population estimates for all States might come from the census, and adjustments for updating would all be computed in the same way--e.g., from the regression of data from a national sample survey on the numbers of taxpayers and dependents (determined from Federal income tax records) and current school enrollments.

The fact that comparability between competitive jurisdictions is frequently best served by taking the data for these governmental units from the same source, has been extended into a rule that data for all jurisdictions, competitive or noncompetitive, must come from the same source. Such a rule can actually lead to less rather than more comparability. It may, in fact, force the use of grossly inadequate data because the only

source available for all jurisdictions is a very inferior source. In tiered allocation systems it may be better to use a common data source at any one level but not to insist on using it at all levels. Thus, sample survey estimates of current State populations might be the best estimates for the allocation of funds to States, but, for allocating the total for a State among cities and counties, we might use estimates based on adjusting 1970 Census populations for changes in school enrollments and in the number of income tax payers and dependents.

It may even be desirable to use different, data series for allocations within different States. Thus, one State may be able to get a quite good estimate of the population of each city and county (and also of each township and city ward) in the State from the regression of census population on the number of registered voters plus school enrollments, while the voter registration and school enrollment statistics would be much inferior

for another State in projecting past intercensal population increases.

The use of different data series for different levels and for different allocations within a level is a case where the use of non-identical data actually helps to maintain comparability. A much more difficult problem is the availability of better data for a few jurisdictions in a set of competing jurisdictions. For example, one city in a State has a special census taken which shows a population increase for the area of 30 percent since the previous census, as against the estimates of population growth of five to nine percent obtained for this city and other cities and counties of the State by projecting popu-

lation trends shown by the last three decennial censuses. Is it proper to use the population figure from the special census for this particular city when no comparable figures are available for the other cities and counties of the State? One could argue that using the special census estimate gives an unfair advantage to this city since other cities or counties may have had similar or greater population growths. On the other hand, it could be argued that there are, at most, two other areas in the State that had more than nine percent population growth and that it is unfair to penalize this city because the other areas of the State had no reason to take a special census. Solutions to the problem might be:

(1) to try to find some method using already existing data which would properly reflect post-censal population growth for all areas of the State, or

(2) to execute a small sample survey to determine whether any other city or county has had unusual population growth and follow up by larger sample surveys of those jurisdictions which do show large population changes.

Problems arising from the formula. There are many alternatives in the construction of allocation formulas. For some of these alternatives (e.g., the use of an additive versus a multiplicative formula) the pros and cons are pretty evenly balanced and the choice becomes a matter of purposes to be served, the data available, and individual tastes. There are a few alternatives which are clearly inferior from both a statistical and policy standpoint. The handling of cutoffs is one of these.

a. Additive versus multiplicative formulas. In a multiplicative formula the allocation is

automatically, equally sensitive to variation in any of the factors. That is, a 10 percent change (or a 10 percent difference between two areas) in any factor will mean a 10 percent difference in the allocation (unless the formula includes a cutoff provision). In an additive formula weighting is needed to determine the relative sensitivities of the allocation to the different factors in the formula. Weights are frequently arbitrary and poor choice of weights can lead to serious dissatisfaction with the operation of an additive formula. On the other hand, if a multiplicative formula is used, a small error in one factor can throw the whole allocation seriously off. Thus, one is damned if one does, and damned if one doesn't. The choice of the formula type must, then, depend upon judgments of the accuracy of the various data to be used versus the availability of suitable weights for an additive formula. It is important to provide for constant monitoring of the allocation system so that major errors in the data can be promptly detected and corrected for multiplicative formulas, or so that a poor choice of weighting factors (or a major shift in the underlying causal system) can be promptly detected and corrected for additive formulas.

b. Cutoffs. Undesirable discontinuities may be introduced into an allocation system by cutoffs, especially by eligibility cutoffs. For example, if an area must have an unemployment rate of five percent before it can receive any funds, a very trivial error in the estimation of the unemployment rate can easily throw an area from under five percent or from over five percent into the other group. Here a very small error can make a tremendous difference and lead to continual complaints about the accuracy of the data on the part of governmental units which feel the

cutoff operates to their disadvantage.

A common solution to controversies over cutoffs is to provide alternative formulas and to permit each jurisdiction to select the formula which is most advantageous. While this

works moderately well, it has the disadvantage of making it difficult to predict in advance (and budget for) the amount required for the program if no fixed overall sum to be allocated is specified. If an overall sum is specified but each jurisdiction may choose which formula it will use in determining its share (with the computed amounts totaled over all competing jurisdictions, so that the percent of the total allocated to each jurisdiction can be determined), one gets a floating cutoff, where the amount one jurisdiction gets depends upon the decisions made by other jurisdictions.

For eligibility cutoffs it is almost always possible to devise a formula such that there is a gradual approach to zero (or to some cutoff point lower than the existing absolute cutoff). Here small errors in the data lead only to small changes in the allocation and the tendency to prolonged (and insoluble) arguments over minor errors is removed. Of course, major errors will and should continue to be the subject of controversy but one will be spared the waste of time and effort involved in the use of a formula which requires data of unattainable accuracy.

c. Sensitivity to change. In most cases it is desirable for allocation rules to be relatively insensitive to short-term fluctuations in the data but responsive to long-term changes. However, short-term and long-term are in the eye of the beholder. How short is short-

term and how long is long-term? The answer varies from one program to the other. The CDBG obviously needs at least a four or five year period even to permit contemplation of a building project or the planning of any substantial building program. What one needs is something that will not be thrown totally off the target by short-term fluctuations. On the other hand, gradual change in response to changing needs is desirable and some type of damped dynamic system (for example, an exponential smoothing type of function) is required.

CDBG appears to be the only one of the case studies which tried to use such a damped dynamic system (for bridging the transition to a drastically changed allocation system). The CDBG formula used for this purpose involves a so-called hold harmless provision. However, it should be noted that the hold harmless provisions of most allocation formulas are the reverse of damped dynamic systems. At one end, hold harmless clauses create a totally static situation, permitting an area to claim its allocation of last year (and possibly of several years previously) although conditions may have changed permanently so that a considerably reduced allocation would be quite adequate. At the other end, an area can claim a sharply increased allocation due to a temporary change in the local situation.

When responsiveness to short-term problems is desired, fixed annual allocations for every funding level usually are not appropriate. Switching of funds as needed, both from one time period to another and from one jurisdiction to another, may be required. In AFDC, funds are allocated for a year so that jurisdictions can determine approximately what to expect. The specific allocations are determined as the money is spent and can

vary from month to month.

3. Setting feasible accuracy goals. A major question is to what extent should one adjust the data to fit the accuracy requirements, and to what extent should accuracy requirements be adjusted to fit the data. Some people tend to think in terms of statistics that are literally correct and in terms of an absolute truth which must be met in fund allocations. Many law suits deal with errors in the data and with other errors which cannot possibly be avoided at a reasonable cost. We need to learn to accept the fact that the function of the statistician is not to provide error free data but to pick out those errors which are largest, and try to control them. As for the smaller errors, we must learn to live with them.

   Recognizing that errors in the data and resultant inequities in the allocations are inevitable, major attention must be given to deciding which errors need to be reduced. As mentioned above, a subject of considerable controversy is whether one should try to minimize the sum of the absolute errors or of the relative errors (or of something in between) in the data for individual areas. When sample data are used, minimizing the sum of the relative sampling errors of the individual area figures leads to allocating the

same number of sample cases to each area (e.g. to each State); minimizing the sum of the absolute errors leads to allocating a number of sample cases proportional to the total population of the area.

   A commonly used compromise between the two allocation rules mentioned above (minimizing the sum of the absolute errors vs. minimizing the sum of the relative errors) is

   a. to minimize the sum of the absolute errors by assigning cases proportional to the area population;

b. if this would give any area a relative error larger than the predetermined error limit, increase the sample for the area(s) to the level necessary to give the desired relative error; and

c. reallocate the residual sample for the areas not changed by (b) above, proportionally to area populations.

For fund allocation, this sampling design fits the logic that a big relative error for a small area leads to a serious error in the amount allocated to that area, but cannot have an appreciable effect on fund allocation to the other areas (since the amount of funds going to the area is small in any event), while, for the larger areas, even a small relative error can involve a substantial sum of money and thus lead to inequities in the allocations to all areas when the total amount to be allocated to all subdivisions is a fixed sum.

The technique of proportional allocation with the overall sample set to give a predetermined maximum relative error for an individual area has some limitations. For example, the amount budgeted for the survey may not permit a sample large enough to achieve the predetermined maximum relative error. An alternative is to use proportional sampling for larger areas but to take a sample sufficient to achieve the maximum relative error limit for the smaller areas. Further discussion of these issues may be found in Appendix B-3, "Some Considerations in Designing Samples to Obtain, Data for Use in Allocation Formulas."

Finally, there is no such thing as an ideal formula or ideal data. Therefore, one may have to sacrifice something in the formula and something in the data in order to reach a reasonable compromise between an ideal formula with poor data or a poor formula with ideal data.

CHAPTER V

Subcommittee Recommendations

In Chapter II of this report, a number of causes were identified that contributed to a phenomenon encountered in the five case studies--that "existing allocation formula techniques do not fully achieve the stated objectives of Federal programs." Our review identified problems of formula structure and constraints, problems of implicit and explicit assumptions, problems arising from the data used, and some effects of the interaction of formulas and data. In Chapter III the Subcommittee has presented some general findings on the basis of the five case studies and in Chapter IV has identified some specific ways to reduce allocation errors and inequities. On the basis of these general and specific findings, the Subcommittee has formulated the following set of recommendations to improve the Federal process for specifying and administering the formula aspects of grant-in-aid, programs, for dealing with statistical considerations in formula construction, and for relating programmatic measures to ongoing statistical series.

The Subcommittee recognized in its review of the five case studies that there were pervasive problems in the obsolescence of key data, particularly where decennial census data were required to be used, and in the choice of statistics to represent small geographic areas. The Subcommittee feels that it is quite important to recognize these elements as important problems early in the program design process so that sufficient attention can be devoted to the generation of at least partially satisfactory solutions. The, specific recommendations on these points are as follows:

RECOMMENDATION 1. Program Goals and Statistics:

That program goals be specified as clearly and

completely as possible in the statement of purpose of each grant-in-aid act and that program drafters guard against over-specification of the statistical data and procedures to be used.

Comment:

Vague specification of program goals and over specification of statistical procedures are common problems. Providing flexibility to program administrators in the choice of statistics for allocation is sometimes desirable for a variety of reasons, but in the absence of reasonably clear and complete goal statements, administrative decisions which involve use of that flexibility will necessarily be arbitrary to some degree, and may run counter to the intent of Congress. The AFDC counts in Title I, ESEA are an example of highly specified statistical procedures written into authorizing legislation. The Education Amendments of 1974 describe with some precision how to determine the number of AFDC children counted for Title I, ESEA funding, which year's poverty standard to use, which of the many poverty cutoffs (nonfarm family of four), and which month's caseload data. What is lacking is a clear statement of what the resulting total is supposed to represent.

The Subcommittee has recognized in its review of the five case studies that some Federal programs have an extensive list of specific purposes and amount to a form of special revenue sharing, or are directed toward some broad categorical objective in, say, education or community development. The Subcommittee does not expect legislative drafters to alter markedly the kind of purposes set forth in future allocation legislation, but rather to recognize the problem of translating such statements of purpose into programmatic measures. If goal statements can be made clear then there will be less necessity to build into legislation in rigid form the specification of the statistics and techniques to be used. For example, Congress might decide to specify a certain

mechanism for allocation to, say, the State level, might leave to Federal-State negotiations and administrative determination the mechanisms for making allocations to lower levels. It should be recognized that sound, flexible administration depends on clear and distinct statutory goals. When goal statements are not clear, then an administering agency which exercises discretion may be subject both to political pressure and to litigation.

RECOMMENDATION 2. Legislative-statistical Interface:

That provision be made for an active, continuous interface between legislative program drafters and the statistical community.

Comment:

This recommendation by the Subcommittee is motivated in part by a recognition that Recommendation 1 will be most difficult to achieve without

sustained professional interchange between program and statist ical staff, both executive and legislative.

RECOMMENDATION 3. Formula Performance Testing and Monitoring:

That statistical and program agencies provide to program drafters an analysis of the sensitivity over time of proposed formulas and of the statistics they incorporate so that possible effects on allocations can be anticipated. Also, that provisions be made for testing, monitoring, and assessing by program agencies of the performance of each specific formula or allocation rule prior to enactment.

Comment:

An example of the type of analysis that might be provided, is that given in the Bureau of the Census report "Coverage of Population in the 1970 Census and Some Implications for Public Programs," which describes some possible effects on the distribution of General Revenue Sharing funds of adjusting the 1970 Census of Population for the estimated under-count and for error in income reporting.

Before an allocation procedure is adopted, it should if possible be subjected to a test. In some cases this could be done by using data from prior years to determine whether or not the proposed procedure would have allocated funds for each prior year in accordance with Congressional intent. In cases where data from prior years are not available the testing would have to rely on simulation techniques. It is important that allocations be neither unduly sensitive to short-term fluctuations nor lacking in sensitivity to long-term changes in programmatic measures. Once a program is in place, a built-in monitoring mechanism is needed to provide early warning to the executive branch and the Congress that a particular formula or allocation rule may not be behaving as expected.

RECOMMENDATION 4. Undesirable Formula Practices:

That legislative drafters and program designers be advised of data problems and the existence of statistical practices, as exemplified in the five case studies. which may lead to formulas with consequences that are generally recognized as undesirable.

Comment:
CETA allows ASU's (Areas of Sustantial Unemployment) considerable freedom in drawing their own boundaries. They need not follow jurisdictional lines. While it may (or may not) be going too far to say political jurisdiction boundaries should be followed, the current procedure may be too free offer-

ing substantial opportunity for drawing boundaries in an artificial way. In addition, ASU's in order to qualify for CETA Title II funds must experience an unemployment rate of 6.5 percent or more for three consecutive months. This specific eligibility cutoff introduces the problem that small errors close to the cutoff of 6.5 percent may have serious effects on the distribution of funds. These two factors substantially complicate the data collection and may lead to possible inequities as well. An alternative might be to base the amount allocated on the difference between the unemployment rate and some lower cutoff, for example 5 percent, arranging the formula so areas above some upper limit point (say eight percent) get the allocation provided by the present formula.
The GRS program distributes funds to approximately 39,000 jurisdictions, the great majority of which are areas of population less than 2,500 in the 1970 Census of Population. For these areas the problems of obtaining intercensal estimates of population and per capita income are very serious.

RECOMMENDATION S. Needed Formula Research:

That a limited program of applied research and development be initiated to attack some critical problems and fill certain identifiable gaps in the present state-of-the-art of formula design.

Comment:

As discussed further in Appendix B-5, "An Agenda for basic and Applied Research on Allocation Formula Problems," the identification and characterization of key technical problem areas involves the following elements: equity considerations, structural aspects and the nature of the data required for the computational formula, performance criteria, presence or absence of constraints and other specification or modeling problems. Furthermore, relevant meth-

odological tools and relevant areas of substantive theory need to be brought together if we are to achieve a coherent approach to allocation problems. Some of the statistical research issues of allocation procedures can be illuminated by theoretical principles from other fields. We need to bring togeth>

Transfer interrupted!