Journal of Official Statistics, Vol. 18, No. 3, 2002, pp. 371–391

Interactions Between Survey Estimates and Federal Funding Formulas

Alan M. Zaslavsky[1] and Allen L. Schirm[2]

Formula allocations of United States federal government funds to states and localities are determined jointly by formula provisions and the data sources and estimation procedures used to derive formula inputs. Their interactions can have unanticipated consequences that are inconsistent with the policy goals of a program. Such interactions occur particularly when estimates subject to error are used with nonlinear estimation procedures or with nonlinear formula provisions such as a hold harmless or a threshold for funding. Introducing a new estimation procedure, such as model-based small area estimation, or a new data source, such as the planned American Community Survey, can also affect the allocations. Through simulations, we illustrate some of the paradoxical effects of interactions among formula provisions, data sources, and estimation procedures. We also describe features of the allocation formulas used by programs to improve elementary and secondary education and maternal and child nutrition for families in poverty.

Key words: Government statistics; census; current surveys; American Community Survey; small area estimates.

1. Introduction

In many countries, revenues are collected by the central government and then a portion is redistributed to subnational governmental units, either for general support or for specific programmatic purposes. Estimates of units' needs for funds typically are used in conjunction with a formula to determine allocations.

In the United States, government programs that allocate funds to states and localities, often for assistance to the low-income population (Citro and Kalton 2000b, Ch. 2), have typically used estimates from the decennial census in the allocation formulas. These programs have relied on the census because only the ''long form'' sample provided the required measures of income and family structure for a sufficiently large number of households in the areas for which estimates were needed. For example, the Title I education program used census estimates of poor school-age children for allocations. Recently, however, this program began using more up-to-date estimates from the Census Bureau's

[1] Department of Health Care Policy, Harvard Medical School, 180 Longwood Ave., Boston, MA 02115, U.S.A.
[2] Mathematica Policy Research, Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, U.S.A.
Acknowledgments: This article was first prepared for discussion at the Workshop on the American Community Survey, September 1999, sponsored by the Committee on National Statistics (CNSTAT) of the U.S. National Academy of Sciences. The article was revised and presented at the CNSTAT Workshop on Formulas for Allocating Program Funds, April 2000. The authors thank members of the CNSTAT Panel on Estimates of Poverty for Small Geographic Areas, especially David Betson, Tom Jabine, and Tom Louis, and CNSTAT project directors Connie Citro and Michael Cohen, for helpful discussions.

Small Area Income and Poverty Estimates (SAIPE) Program. The SAIPE estimates are derived using statistical models with data from the March Current Population Survey (CPS), the census, and administrative records (Citro and Kalton 2000a, Ch. 2-3; Citro and Kalton 2000b, Ch. 3). Statistical modeling is required because subnational samples in current surveys are small and, therefore, direct estimates are imprecise.

Estimates for Title I and other program allocations might soon be obtained from a new data source, the American Community Survey (ACS), which is scheduled to become available during the 2000-2010 decade, and to replace the census long form beginning in 2010 (Alexander 2001). The ACS has nearly identical content to the census long form, and is designed so that the ACS sample cumulated over five years provides direct estimates that are about as precise as estimates from the long form sample. If the ACS is implemented as planned, it will be a source of continuously updated estimates from a large sample of households. Such estimates could be used for allocating funds.

The introduction of a new data source or estimation method for the allocation of federal funds to states and localities can affect allocations substantially for two reasons. First, the new data source may measure a concept differently from previously used sources. For example, the CPS and the decennial census long form find different levels and distributions of poverty (Citro and Kalton 2000a, Ch. 3). Such differences may be consequences of differing survey items, modes of administration, survey protocols, and other details of survey design. Second, even if two surveys provide unbiased estimates of the same quantity, statistical characteristics of the surveys may differ. Among the relevant statistical characteristics are the distributions of errors and the frequency of the survey.

In this article, we consider the second of these issues, by drawing out some of the potential implications of introducing a new survey, such as the ACS, for allocating funds. Our intent is to address some general characteristics of federal funding formulas and the ways in which they might be affected by a shift to a new data source. We do not attempt here to assess quantitatively how using the ACS might affect allocations under a specific program or how any specific state or locality might be affected.

To this end, we first discuss some of the data sources and estimation approaches that are currently used for distributing federal program funds in the United States. We then describe generic features of funding formulas. We describe some potential anomalies inherent in applying the current formulas to sample data, and illustrate these anomalies with simulations. Finally, we argue that when data sources change, properties of the formulas change as well, and consideration should be given to modifying the formulas in light of the original objectives for which they were designed.

2. Data Sources and Estimation Approaches

Funding formulas often require estimates of the numbers of people who are eligible to receive a program benefit. For example, the number of children in certain age ranges who are in low-income families is required for distributing Title I education aid and calculating state grants under the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). The number of children who are in low-income families and the number of children who are in low-income families and have no health insurance are required for allocating funds under the State Children's Health Insurance Program

(SCHIP) initiative. The fraction of a population that falls into the eligible category may also be important for determining where need is concentrated. Hence, estimates of the total population in a broad category (such as the number of children), the number falling into an eligibility category within that population (such as the number of poor children), and the fraction of the population falling into the eligibility category (such as the poverty rate among children) are all potentially needed for allocating funds.

Estimates of the total population are derived from the most recent decennial census, updated to the present year. These demographic estimates are subject to errors, which are larger for relatively small areas and toward the end of the postcensal decade. Still, analyses by the Census Bureau suggest that errors from this source are smaller than those due to estimation of eligibility counts and rates (Citro and Kalton 2000a, Ch. 8).

Estimates of the eligible population are based on the decennial census, survey data, and possibly auxiliary data sources (Citro and Kalton 2000b, Ch. 2). Estimation procedures may be simple and direct, or complex. For example, before the 1997-1998 school year, Title I education funds were distributed to states based on the previous decennial census, and hence allocations were updated only once each decade, apart from relatively minor adjustments due to school district boundary changes and updating of the small part of the counts (e.g., children in institutions for neglected and delinquent children) based on noncensus data. Since then, however, state and county estimates of children in poverty have been estimated using a complex Empirical Bayes model fitted to CPS data, in which decennial census estimates appear as a covariate along with income tax poverty and non-filing rates and numbers of food stamp recipients. Even the CPS data that are inputs to the model are not simply annual estimates, but instead are *cumulated* over a three-year period for the county-level small-area estimation model. (Cumulation means combining data or averaging estimates over a period of time that includes the reference date of the required estimates.) CPS data are sparse for all but the largest states and counties, and the models that were used only imperfectly fit the data. Nonetheless, assessments by the Census Bureau and by a National Academy of Sciences panel concluded that the model-based estimates were on the whole superior to those obtained by simply carrying forward rates or shares from the previous decennial census (Citro and Kalton 2000a, Ch. 6). (For small domains such as small counties and school districts, sampling error in census long form estimates may be substantial, perhaps even larger than model error.) Numbers of WIC eligibles by state are calculated using a similar, although more complex, model.

Among the most important perceived advantages of the ACS is that it will provide a relatively dense sample in each year, bridging the gap between the current census long form with its dense but temporally infrequent sample, and the CPS and other current surveys which are collected almost continuously but with relatively sparse samples. This offers the possibility of developing current estimates using simple models or cumulation procedures. Depending on the size of the target area (and the sampling rate applied there), ACS estimates may be based on simple cumulation of one to five years of data.

Aside from the purely statistical advantages of such an approach, it may also achieve superior public acceptability because of its apparently greater directness. (Fay and Thompson (1993) made such an argument in reference to an unusually high-profile estimation problem, adjustment of census population counts for states.) Direct estimates are defined as those based only on data collected within the domain for which the estimates are being made; indirect estimates are those that use data for other domains as well. Domains may be defined cross-sectionally (as geographical areas or demographic subgroups of the population), temporally, or both. Simple indirect estimators may average over spatial domains (for example, combining several school districts in a county to estimate a single poverty rate that will be used for all of them) or over time (cumulation over years). More complex indirect estimators include the full range of small area estimation models (Ghosh and Rao 1994), such as synthetic estimation, regression estimation, and hierarchical Bayes models.

The cumulation procedures proposed for the ACS are at an intermediate level of directness between those used in Title I estimation before and after the shift to model-based estimates. Geographically they are direct, but temporally they are indirect. From a purely statistical point of view, both forms of indirectness raise similar issues of model error. Temporal indirectness of the form proposed for the ACS, however, can hardly be criticized if it replaces the even more indirect procedure of estimating the present from a single previous year (the decennial census year) with no current data.

3. Funding Formulas

3.1. Commonly used features of funding formulas

Formulas for the distribution of federal funds to states and substate units can be quite complex (Citro and Kalton 2000b, Ch. 2). A single program may distribute parts of its funds according to several different formulas. Nonetheless, a few typical features of these formulas are relevant to this study.

As noted earlier, funding formulas often involve the distribution of funds in proportion to a measure of need, such as the number of members of a subpopulation who are in poverty by some standard. Generally, the total ''pie'' to be divided is determined by the appropriation for the program, although the level of the appropriation may itself be affected by Congress's perception of total need. Consequently, funding formulas have an aspect of indirectness, in the sense that an increase in allocation to one unit implies a decrease somewhere else, although the effect of each unit's allocation on other units is generally small.

Proportional allocation of funds may be modified by *''hold harmless'' provisions* and *thresholds*. A hold harmless provision limits the amount by which the allocation to a unit can drop from one year to the next. With a 100% hold harmless, no unit's allocation is allowed to decrease. With an 80% hold harmless, no unit's allocation may decrease by more than 20% in any year. The hold harmless level may vary from year to year as part of the appropriations process. The hold harmless level may also depend on some other characteristic of the unit, such as its poverty rate. The rationale for a hold harmless provision is that it moderates fluctuations in the allocation to each governmental unit, softening the effect of cuts on a unit that has budgeted services in anticipation of allocations similar to those from a previous year. The asymmetry of a hold harmless might be motivated by the greater political sensitivity of unexpected cuts in aid as opposed to unexpected increases. With a high hold harmless level and static or declining total appropriations, allocations may be essentially frozen regardless of shifts in the distribution of need indicated by more recent data. With growing budgets, the effect of a hold harmless provision is ameliorated, if the provision is stated in terms of absolute amounts (as is typical), rather than shares of the total amount distributed. For example, if the total budget grows by 5%, a 100% hold harmless allows a unit's share to fall by almost 5%.

A threshold is a minimum level below which a unit is not entitled to receive funds from a program (or a component of a program). A threshold may be an absolute count (e.g., a minimum number of children in poverty) or a rate (e.g., a minimum poverty rate). A threshold on counts operates to prevent dispersal of funds across small units in which the scale of the local program would be too small to administer effectively or efficiently. A threshold on rates directs funds to units where the relative burden of need is greatest and consequently the governmental unit is least able to meet it with its own resources.

The allocation provisions described above are illustrated by two important programs: the WIC nutrition program and the Title I compensatory education program.

3.2. Fund allocation based on state estimates: The case of WIC

WIC is a federal grant program for states that is administered by the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA). The program, which is currently funded at about 4 billion USD per year, provides nutrition and health assistance services for low-income childbearing women, infants, and children. The current rule for allocating WIC food funds to states became effective on October 1, 1999, and specifies that if there is sufficient funding, each state receives a grant equal to its final prior year grant. Thus, there is 100 per cent hold harmless. (If there is insufficient funding to give all states their prior year grants, each state's grant is reduced pro rata.) After prior year grants have been provided, up to 80 per cent of remaining funds are allocated as inflation adjustments. Then, all remaining funds are allocated based on each state's estimated ''fair share,'' that is, its share of the estimated national population of individuals' income eligible for the program. Thus, a state with one per cent of the eligible individuals has a fair share of one per cent of total available food funds, and the dollar amount that is one per cent of the total is the fair share target funding level. States whose prior year grants adjusted for inflation are less than their fair share targets receive ''growth funds.'' The amount of growth funds received by an ''under fair share'' state is directly proportional to the difference between the prior year grant adjusted for inflation and the fair share. States with prior year grants adjusted for inflation in excess of their fair share targets do not receive growth funds (unless all the ''under fair share'' states decline to accept the full amount of growth funds available).

States' fair shares are calculated from estimates of the numbers of infants and children in families with incomes at or below 185 per cent of poverty, the income eligibility threshold for WIC. Beginning with fiscal year 1995, state allocations have been determined from model-based estimates obtained using CPS, decennial census, and administrative records data (Schirm and Long 1995; Schirm 2000). In prior years (under somewhat different allocation rules), state grants were calculated from decennial census estimates. Estimates from the 1980 Census were used from the early 1980s until fiscal year 1994, when 1990 Census estimates were used.

3.3. Fund allocation based on substate estimates: The case of Title I

Title I of the Elementary and Secondary Education Act provides federal funds to school districts for compensatory education for disadvantaged children. To date, Congress has appropriated funds for two types of Title I grants — basic grants and concentration grants, which totaled about seven billion USD and one billion USD, respectively, in fiscal year 1997. Title I funds were allocated to school districts through a two-stage process, through the 1998-1999 school year. The Department of Education allocated funds to counties, and states suballocated funds to school districts within each county. Direct allocations to school districts began with the 1999-2000 school year. Allocations are based on the estimated numbers and percentages of school-aged children who are poor. The rules for allocating funds are complex and include both hold harmless provisions and eligibility thresholds. For example, a variable hold harmless rate pertains to basic grants. A school district is guaranteed at least 95 per cent of its prior year grant if at least 30 per cent of its school-aged children are poor. The guarantee falls to 90 per cent if the percentage poor is between 15 and 30 and to 85 per cent if the percentage poor is below 15. To receive basic grant funds, a school district must have at least ten eligible children who constitute at least two per cent of the district's population aged 5 to 17. To receive concentration grant funds, a district must have at least 6,500 eligible children or have at least 15 per cent of children aged 5 to 17 be eligible. Further complicating the allocation process, Title I grants also depend on other factors, such as state average per-pupil expenditures.

Model-based estimates of the numbers and percentages of school-aged children who are poor in states and counties were first used to allocate Title I funds for the 1997-1998 school year. The Census Bureau developed these estimates from CPS, decennial census, and administrative records data. In prior years, estimates from the decennial census were used to allocate Title I funds. More recently, the Census Bureau developed model-based estimates for school districts that have been evaluated (Citro and Kalton 2000a) and were used in allocating funds directly to school districts for the 1999-2000 school year.

4. Interactions Among Data Sources, Estimation Procedures, and Allocation Formulas: General Findings

Data sources, estimation procedures, and allocation formulas each play a role in the successive steps of calculating fund allocations. In practice, the distinction between the roles played by the estimation procedure that generates the inputs to the funding formula and the formula itself can be a legalistic formality when the same calculations can be positioned in either the estimator or the formula. For example, the law may specify that allocations are to be based on a three-year moving average, while the estimation procedure obtains each year's estimate from a single year's data. The same effect is obtained, however, if the formula uses a single year's estimate but the estimate for that year is calculated (for purely statistical reasons) as a three-year moving average. For another example, a formula may specify that a school district's eligibility for a category of funds shall depend on the poverty rate in the district, but if estimates are calculated only for counties and then carried down to the districts, the effect is the same as if eligibility were calculated at the county level. In that case, developing a capability to estimate poverty rates by district effectively changes the formula. On the other hand, some formula provisions do not have natural counterparts in estimation procedures; a hold harmless provision is a common example.

Keeping this relation between estimation procedures and formulas in mind, we consider in the next section the effect of various choices of formula and estimator under various scenarios for sampling error (determined in part by the size of the unit) and year-to-year patterns in the population value (number or rate) for the target group (e.g., children in poverty). Before setting out detailed scenarios, we note several facts.

First, reliance on decennial census data implies that much of the time the data will be seriously out of date. Because of the time it takes to process long-form data, these are about two years old by the time they are tabulated, and the reference year of the data is the year previous to the year in which they are collected. Therefore, by the time the next census's data become available, data from the previous census will have been used to allocate funds up to 13 or even more years past the reference year. Analyses of CPS data for Title I allocations suggested that substantial shifts in the geographical distribution of poverty can take place in periods of three or four years (Citro et al. 1997), a finding that should not be surprising to students of regional business trends. Consequently, reliance on census data implies unresponsiveness to important short-term regional trends in poverty.

Second, even in terms of long-run averages, reliance on decennial census data is problematical because it only gives a few widely separated snapshots. For example, over a 30-year period only three censuses take place, and it would not be surprising if some units happen to have poverty rates at all three censuses that are substantially below their average rates over the 30-year period. Such units would not receive their fair share of allocations, even averaged over the 30-year period. Similarly, a unit could fall below a threshold in a single year that happens to be a census year, and hence lose the entitlement to funding that it might have obtained if the census had occurred in any other year. In effect, the estimates suffer from small temporal sample size. This problem cannot be solved by adding sample in the census years. Instead, more frequent measurements must be taken.

Third, the effect of hold harmless provisions depends on both the frequency with which new data become available and the frequency of allocations. For example, after new decennial census data become available, allocation shares could be determined only once, or they could be calculated annually applying a hold harmless each year, so that a unit whose share has fallen would move to its new share through a series of annual steps. With decennial adjustments of allocations and a fairly high hold harmless level, it may take several decades for a unit with a single spike in its poverty rate to receive an allocation appropriate to its more typical level. With annual adjustments, even with a hold harmless level very close to 100%, the cumulative change in allocations over a decade is likely to be larger; for example, ten decreases of 7% are about equivalent to a single decrease of 50%. In practice, hold harmless levels are decided by legislation or regulation. Consequently, the actual effect of changing the schedule of allocations is unpredictable, because policy-makers may be influenced by the change in the schedule to set a different hold harmless level than if allocations were adjusted only after each decennial census. (We show in Section 5 that sampling error further complicates the effect of hold harmless.)

Fourth, if annual samples are independent, or almost so as in the ACS, then variances can be reduced by cumulation, i.e., calculation of a moving average. Assuming uncorrelated sampling error with equal variances in each year, using a three-year equally weighted moving average multiplies variances by a factor of one-third (.333). Less obviously, an exponentially weighted moving average using three years of data with weights proportional to $.7^0=1$, $.7^1=.7$, and $.7^2=.49$ (at lags 0, 1, 2 years) multiplies variances by a factor of .361, very close to the reduction obtained by equal weighting, while giving larger weight to the most recent data. (The weighting factor of .7 is a compromise value that reduces the weight on data from two years back substantially, to half that of the most recent year, but does not too greatly affect variances.) These results on cumulation do not apply to the CPS, because of the positive correlation between annual estimates induced by its rotation group design. Although this design can be exploited to obtain improved estimates of changes, simple cumulation will not reduce variance as much as when samples are independent.

Fifth, holding estimation and allocation procedures and annual appropriations constant over time, a linear estimation procedure with unbiased inputs (e.g., a weighted moving average of direct estimates, with fixed weights for each lag) combined with a fixed linear allocation formula gives allocations that tend to agree, on the average over a long time period, with the correct allocations. This follows from the fact that every year is given equal total weight (appearing at each relevant lag) except the years close to the beginning or the end of the interval; thus, with unbiased direct estimates and linear estimators and formulas, sampling error averages out over a period of time. The premises of this argument are not entirely realistic. Annual appropriations for a program are not constant (in current or constant dollars). Hence, it is inevitable that some units will have the good fortune (or political influence) to be entitled to their largest shares of the pie in the years in which the pie is largest. Such a unit will receive an aggregate share over the period that is larger than the average of its annual shares; conversely, another unit will receive a smaller aggregate share. Furthermore, it is not evident that ''unbiased'' aggregates in this sense are a particularly desirable property from the standpoint of fair or efficient allocation, when needs change from year to year. Nonetheless, this result suggests that some of the complexities of the interaction between the estimation procedures and formula arise because one or both is nonlinear.

5. Interactions Among Data Sources, Estimation Procedures, and Allocation Formulas: Illustrations

5.1. Factors defining the simulation scenarios

We now consider some of the more complex interactions among the elements of the allocation process by developing several illustrative scenarios. We assume that allocations are based on a single variable, which may be interpreted as a standardized poverty rate, set on a scale (for simplicity of presentation) where a typical value is about 1.

We ignore the dependence among units' allocations. In practice, each unit's allocation is affected by the allocations to other units, because the units must share a prespecified total appropriation, but this is not important to these illustrations, in which we focus on the *differential* effects on different units. In Section 6 we consider more rigorously how this form of dependency among units affects our results.

We simulate annual allocations over a four-year period. Each scenario is defined by four elements, drawn from a set of alternatives.

(a) *Population process.* The poverty rate is either constant (CONS) at one of several rates, trending upward from .75 to 1.25 (UP) over a 4-year period, or trending downward from 1.25 to .75 (DOWN).

(b) *Sampling standard error (SE) of estimates*. This takes one of four nonzero values, .1, .25, .5 and 1. These may be regarded as corresponding to a moderately large unit, two mid-sized units, and a small unit, defined in terms of sample size. We also consider a unit with no sampling variance, representing a very large unit, as a standard of comparison. We assume that sampling error is normally distributed with mean zero. (This is a reasonable approximation for small values of the sampling standard error, but not for SE = 1, for which normality would imply a substantial probability of a negative estimate of the rate.)

(c) *Estimation method*. We consider three methods: single year estimate (SINGLE), three-year moving average with equal weights (MA3), and three-year moving average with weights proportional to $.7^0 = 1$, $.7^1 = .7$, and $.7^2 = .49$ (MAE3).

(d) *Allocation Formula*. There are four possibilities. The first formula (PROP) has proportional allocations. In fact, a unit's allocation is simply equal to its estimated poverty rate. The second formula (HH) has proportional allocations subject to an 80% hold harmless, meaning that a unit's allocation is the maximum of the current estimated poverty rate and 80% of the last allocation. The third formula (THRESH) has proportional allocations subject to a threshold. Thus, a unit's allocation is equal to its estimated poverty rate if the rate is above a threshold of 1 and zero if it is below 1. The fourth formula (HH-THRESH) has proportional allocations subject to both a threshold and a hold harmless provision, implying that a unit's allocation is equal to the maximum of the unit's current estimated poverty rate (or 0, if the current rate is less than 1) and 80% of the last allocation. Under the second and fourth formulas, we assume that the hold harmless provision does not affect allocations in the first year.

Rather than simulating all possible combinations of these elements, we focus on a few combinations to define scenarios that illustrate specific points. In many of our simulations, we emphasize the effect of sampling variability on a unit's expected allocation under a particular scenario. Because sampling variability is so much affected by the size of the unit, this approach focuses attention on possible inequities to large or small units that are otherwise similar, i.e., the tendency to receive smaller than deserved allocations simply because of size. Furthermore, because sampling variability is affected substantially by decisions on the design of data collection (particularly sample allocation) that are made on technical and cost grounds, it should be known if such decisions affect relative allocations. The variability of the allocations from year to year could also be important, and is considered in Section 5.6. Furthermore, we assume that the allocation methods and data sources remain constant from year to year; our results suggest effects that might also be expected to appear in a more dynamic environment.

5.2. *Effect of sampling variability when there is a threshold*

Table 1 illustrates the effect of sampling variability when there is a threshold. For examining this effect, we assume that there is no trend, and each year is estimated independently. In other words, the population process is CONS, with constant true poverty rates 1.3, 1.1, .9, or .7; the estimation method is SINGLE; and the allocation formula is THRESH. The values for the sampling standard error are displayed in the first column of Table 1. The other entries in the table are expected values of the allocations (averaging

Table 1. Simulated allocations with a threshold and varying amounts of sampling error

True rate	1.3	1.1	0.9	0.7
Correct allocation given true rate	1.3	1.1	0.0	0.0
Standard error	Expected allocation			
SE = 0.00 (exact)	1.30	1.10	0.00	0.00
SE = 0.10	1.30	0.95	0.17	0.00
SE = 0.25	1.20	0.81	0.40	0.13
SE = 0.50	1.11	0.84	0.57	0.36
SE = 1.00	1.19	0.99	0.82	0.65

over the sampling distribution of the estimates). For each value of truth and standard error in our simulations, we simulate annual poverty rate estimates by adding to the true rate some random noise (draws from the normal distribution with the standard deviation that we have specified and mean equal to zero). Then, we calculate allocations over a four-year period using the chosen formula. We repeat these steps many (20,000) times and average the allocations. Because each year is independent in the simulations conducted for Table 1, it suffices to simulate a single year.

Note that with exact information (no sampling variance), each unit receives its proportional allocation if above the threshold, and nothing if below, as required by the allocation formula. However, with increasing sampling variance the below-threshold units have increasing probabilities of estimates above the threshold and therefore an increasing expected allocation. This effect, of course, appears at smaller standard errors in units where the true rate is just below the threshold, as shown by comparing the last two columns of the table. The situation for above-threshold units is more complex. With modest amounts of sampling variability, the probability that the sample estimate falls below the threshold, causing the unit to lose all of its funding for the year, becomes large enough to reduce the unit's expected allocation. When sampling variability becomes sufficiently (perhaps unrealistically) large, however, the expected allocation begins to increase again, because the positive errors (which are in theory unbounded) begin to compensate for the negative errors (which are bounded because the allocation is never negative). This increase in expectation is accompanied by a drastic increase in variance, as eligibility for any funding approaches a coin toss (assuming, again unrealistically, that the sampling distribution is symmetrical).

Reading down any column, we see how changing sampling variance affects a unit's expected allocation at each value of the truth. Particularly for true poverty rates close to the threshold, the differences down the column can be large. It is difficult to imagine a rationale for giving a unit a larger expected allocation because a decision that was made about sample design for a survey caused the unit's poverty rate to be estimated less precisely.

As sampling error increases, the sharp cutoff envisioned in the formula is replaced with an increasingly smooth (ultimately almost linear) relation between the true poverty rate and the expected allocation. It is arguable that sharp thresholds in funding formulas are not entirely sensible, and that a smoother transition would give more stability and less importance to very small shifts near the threshold. However, smoothing expected allocations around the threshold through sampling noise is a poor way to achieve this objective. For units with substantial sampling variability, the threshold magnifies annual variability in allocations relative to a smooth transition, even though the expected allocation over time is smoothed. Furthermore, the amount of smoothing around the transition is dependent on the design for each unit, and the cutoff at the threshold is sharpest for units with small sampling variability.

5.3. *Effect of sampling variability when there is a hold harmless provision*

Figure 1 shows the effect of sampling variability when there is a hold harmless at 80% and the underlying population poverty rate is constant at 1 (HH, CONS). Each panel pertains to a different estimator (SINGLE, MA3, MAE3). The solid line in each panel shows the ‘‘correct’’ allocation (based on the true value 1) and the dotted lines show the expected allocations with annual SE = .1 (triangle), .25 (+), and .5 (X). In these simulations, we draw the estimated poverty rates independently in each year (simulating independent sampling). Nonetheless, the calculated allocation is affected, through the hold harmless provision, by the allocation in the previous year.

Expected allocations in the first year are all equal to 1, because we assume no effect of hold harmless in the first year. In successive years the expectation climbs because the allocation is ‘‘ratcheted up.’’ That is, although sampling error can raise the allocation, perhaps substantially, in a year, the hold harmless always keeps the allocation from falling very far downward the next year, regardless of how low the estimated poverty rate is. Comparing the lines in a given panel, we find that the ratcheting effect is greatest for the smallest units (i.e., the units with the largest standard errors). Thus, like the bias from a threshold, the upward bias from hold harmless is size-related. Comparing the three panels, we find that use of a moving average estimator of the poverty rate greatly mitigates this ratcheting effect, more than would be expected simply due to the reduction in variance (from tripling the sample size relative to single-year estimation). With a three-year moving average, the standard error of the estimates for the scenario with annual SE = .5 is reduced to $.5/\sqrt{3} = .289$, but the bias in Year 4 is reduced to .029 (estimated by simulation), much less than the bias of .057 that is found with single-year estimates with SE = .25. This reduction in bias is a consequence of the fact that the three-year moving average estimates for consecutive years use data from two of the same years (and one different year at each end), so the series of estimates is positively autocorrelated (i.e., a year with a positive estimation error will tend to be followed by another year with a positive error). Hence the moving average estimates are smoother over time than independent annual estimates with the same standard error, and big jumps in estimates that trigger the hold harmless provision are less likely to occur. (See Section 5.6 for illustrations of this greater smoothness.) This illustrates that a linear smoothing procedure can give some of the stability that is sought with a hold harmless provision, without the size-related bias that hold harmless can engender.

The combined effect of a hold harmless and a threshold is even more drastic than the effect of either alone. Table 2 is comparable to Table 1, except the simulations for

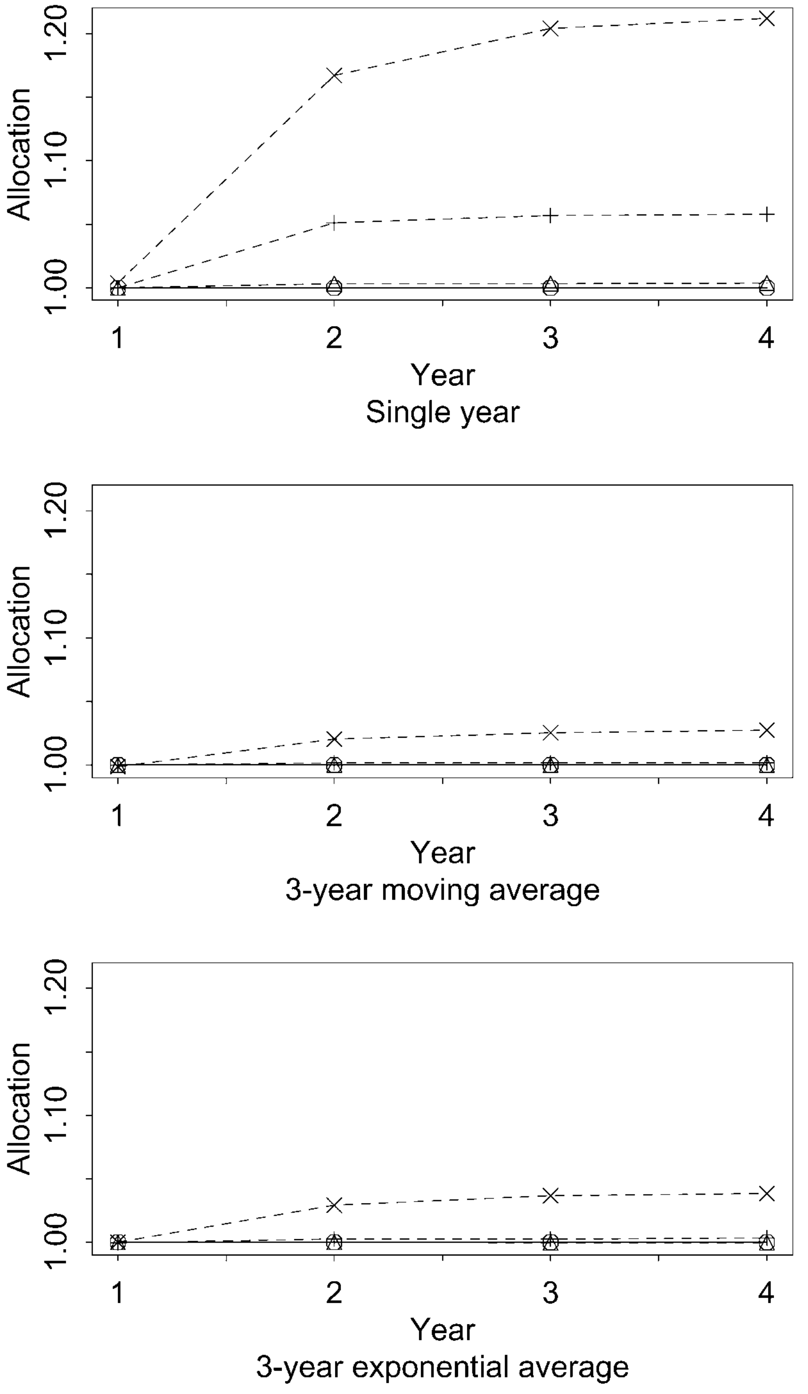

Fig. 1. Effects of sampling variability with a constant poverty rate (= 1.00) and a hold harmless provision. Correct allocation (solid line) and expected allocations (dashed lines) with annual SE = .1 (triangle), SE = .25 (+), and SE = .5 (X).

Table 2. Simulated allocations with a threshold and an 80% hold harmless provision, and varying amounts of sampling error

True rate	1.3	1.1	0.9	0.7
Correct allocation given true rate	1.3	1.1	0.0	0.0
Standard error		Expected allocation		
SE = 0.00 (exact)	1.30	1.10	0.00	0.00
SE = 0.10	1.30	1.09	0.41	0.00
SE = 0.25	1.34	1.12	0.78	0.33
SE = 0.50	1.47	1.27	1.04	0.77

Table 2 assume that there is an 80% hold harmless as well as a threshold. According to Table 2, the expected allocations, calculated for Year 4 when the effects of hold harmless have approached steady state, are extremely sensitive to sampling variances. A unit that has a true poverty rate that is just below the threshold (set at 1) but a large measurement standard error has a very high-expected allocation relative to what it would have received if there were no measurement error. This occurs because once a unit has an estimated poverty rate above the threshold and receives funding, the hold harmless delays the unit's drift down toward zero funding even if its estimates are below the threshold for the following several years.

5.4. *Effects of various linear estimation methods when there is a trend*

Figure 2 shows a hypothetical downward trend (solid line) in the population poverty rate and the expected allocations with three estimation methods (dotted lines): single-year (SINGLE = triangles), three-year moving average (MA3 = +), and exponentially weighted moving average (MAE3=X). The sampling standard error is not relevant to the calculation of expected allocations in this case: the estimators and formula are linear,

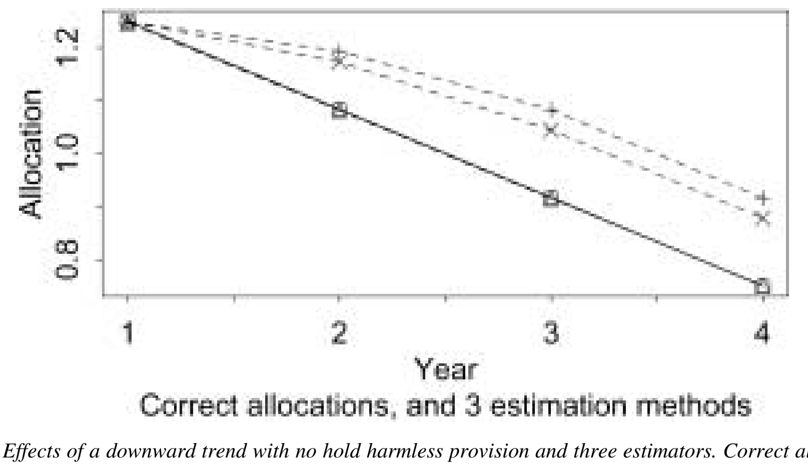

Fig. 2. Effects of a downward trend with no hold harmless provision and three estimators. Correct allocation (solid line) and expected allocations (dashed lines) with single-year (SINGLE = triangles), three-year moving average (MA3 = +), and exponentially weighted moving average (MAE3 = X) estimates.

so adding sampling variability does not affect expectations. As anticipated, the single-year estimates track (in expectation) the correct allocations, but the moving averages trail them. The exponentially weighted average, because it weights more recent years more heavily, trails slightly less far behind than the equally weighted average. This illustrates the bias-variance trade-off inherent in modeling. Note that as long as ''what goes up must come down,'' the upward bias when poverty rates are falling is balanced by a downward bias when poverty rates are rising. The optimal weighting method (number of years and weights on each lag) depends on sampling variances, the magnitude and pattern of process variability over time, and the importance attached to timeliness and accuracy of estimates.

5.5. *Effect of a hold harmless provision when there is a trend*

These scenarios are similar to those in Section 5.4. but add a hold harmless provision, which makes the sampling standard error relevant. Each panel in Figure 3 pertains to a different estimator (SINGLE, MA3, or MAE3). The solid line in each panel shows the ''correct'' allocation (based on the true value 1) and the dotted lines show the expected allocations with annual SE = .1 (triangle), .25 (+), and .5 (X). The effects are a combination of those seen in Sections 5.3. and 5.4.: units with large standard errors tend to be ratcheted upward more than units with small standard errors, and moving averages lag behind the trend and reduce the ratcheting effect.

Figure 4 shows expected allocations when there is an upward trend in poverty rates. Here, the bias due to hold harmless has been mitigated: with increasing rates, the hold harmless is less likely to have an effect.

5.6. *Hold harmless versus moving average as methods for moderating downward jumps*

In these scenarios, estimates fluctuate around a mean of 1 with SE = .5. These fluctuations represent the sum of sampling error and uncorrelated year-to-year variability in the population poverty rate. We compare three approaches to reducing the magnitude of downward jumps in allocations from year to year. In the first, an 80% hold harmless is applied to single-year estimates with SE = .5. The second is like the first except that we assume that the standard error is reduced to $.5/\sqrt{3}$. (If variability is entirely due to sampling error, this reduction in the SE could be obtained by multiplying sample size by 3.) The third approach assumes that a formula without a hold harmless provision is applied to a three-year moving average estimate (which has $SE = .5/\sqrt{3}$, the same standard error as for the second approach). For evaluating the alternative approaches, we look at the changes in allocation from Year 3 to Year 4, when the effect of the hold harmless has almost reached steady state. Table 3 shows the fraction of changes that go in the downward direction, the mean of those changes, and the mean of the changes in the upward direction. As expected, allocations based on moving average estimates are equally likely to go up or down, because there is no hold harmless. When there is a hold harmless, the asymmetry of the hold harmless leads to more downward than upward shifts: because the downward shifts are limited in magnitude, there must be more of them. Another way of explaining this effect is that the upward bias of the hold harmless with large SE means that the current allocation tends to be higher than the long-run mean in the absence of hold harmless (i.e., the true poverty rate), and therefore more often than not will be adjusted downward

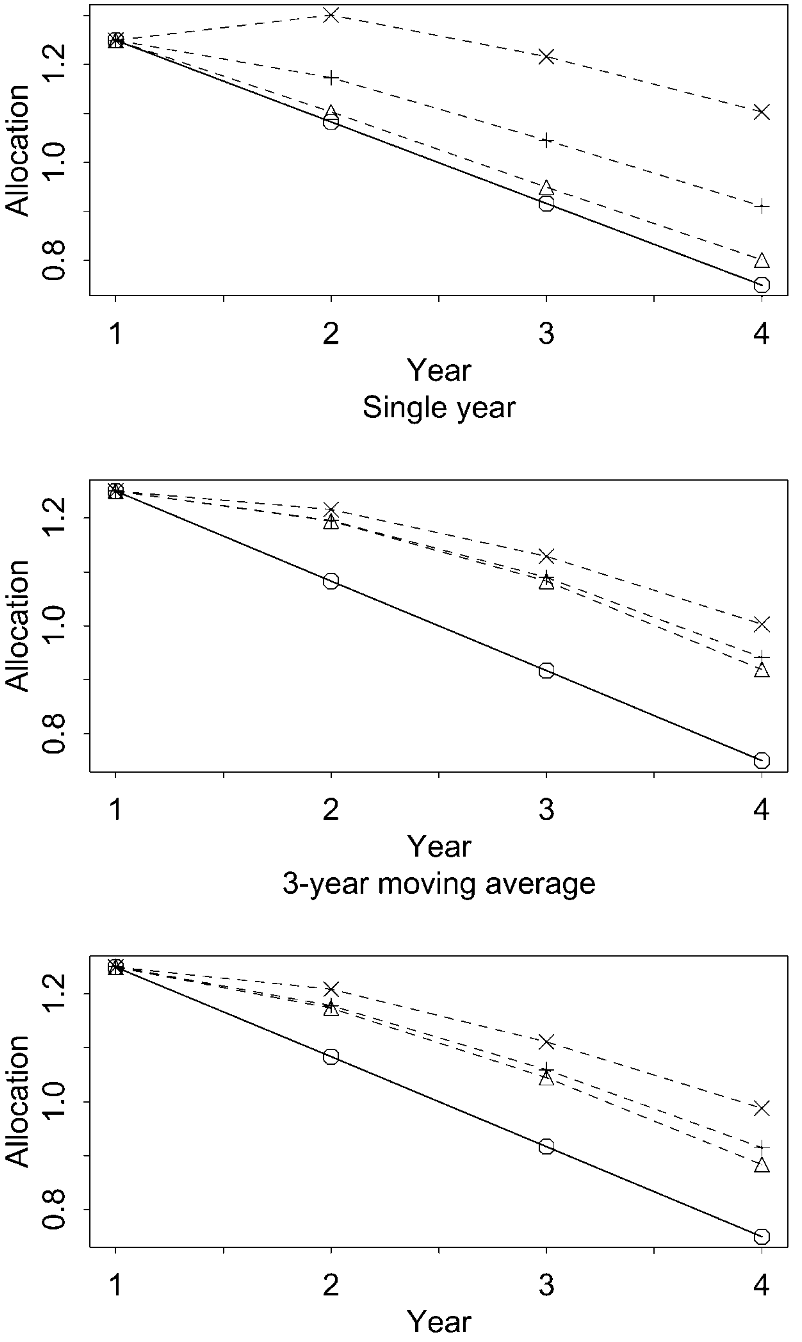

Fig. 3. Effects of a downward trend with a hold harmless provision and three estimators. Correct allocation (solid line) and expected allocations (dashed lines) with annual SE = .1 (triangle), SE = .25 (+), and SE = .5 (X).

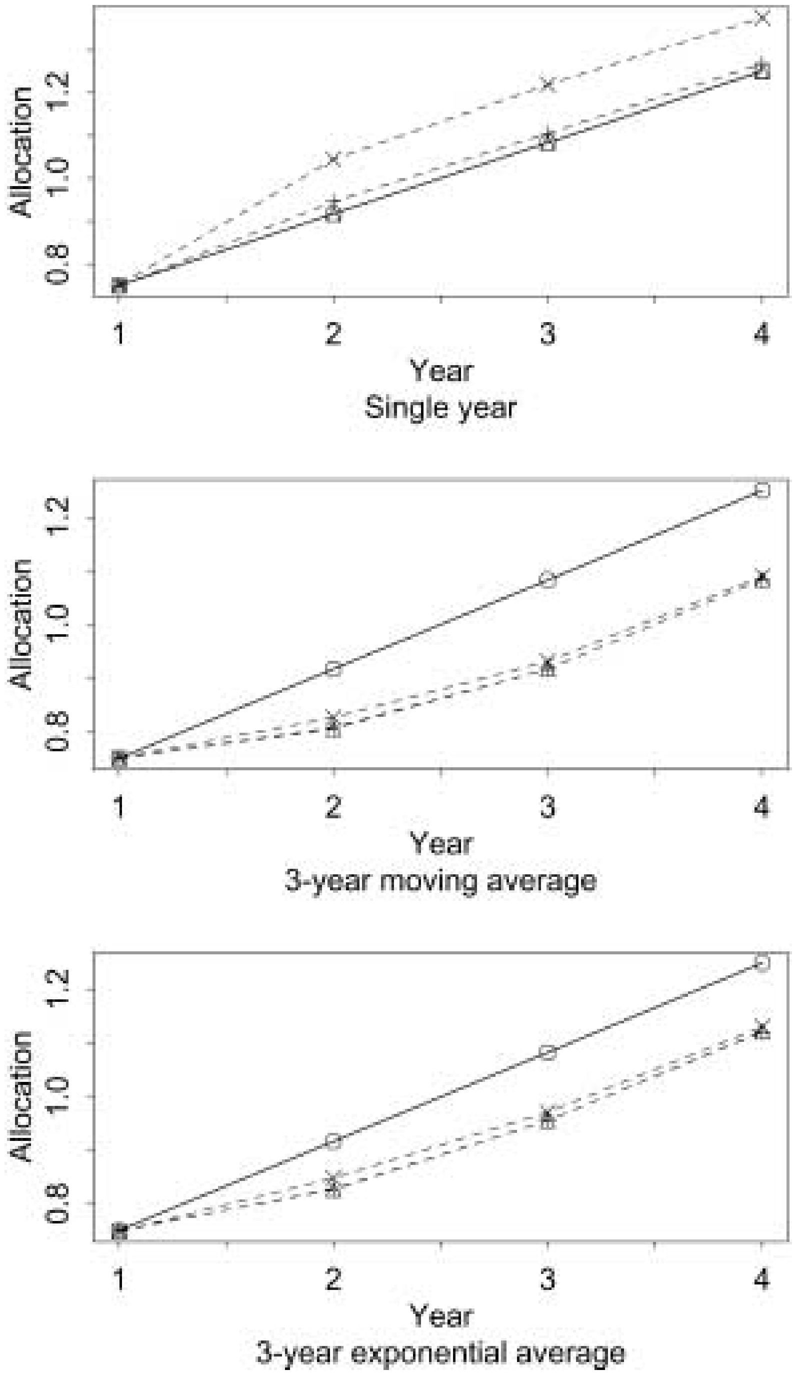

Fig. 4. Effects of an upward trend with a hold harmless provision and three estimators. Correct allocation (solid line) and expected allocations (dashed lines) with annual SE = .1 (triangle), SE = .25 (+), and SE = .5 (X).

Table 3. Effect on magnitude of annual changes of a hold harmless compared to a moving average

	Estimation and allocation approach		
	HH	HH(3)	MA3, no HH
Fraction down	0.624	0.576	0.500
Mean down	0.236	0.189	0.188
Mean up	0.415	0.261	0.188

because the probability is larger than half that the estimated poverty rate will be less than the current allocation.

Comparing the mean magnitude of the steps, we find that in the realistic comparison of the first and third columns, both the downward and upward steps engendered by the allocation formula with a hold harmless are larger on the average than those engendered by a moving average estimator with a proportional allocation formula. Even the second column (representing a somewhat unrealistic scenario, since we assume that an expansion of sample size could be afforded) has downward changes that are no smaller than those obtained with a moving average. This suggests that use of a moving average can be as effective as a hold harmless in moderating downward swings in allocations. The cost of using a moving average, however, is that it is less responsive than a single-year estimate to upward jumps in the rate; such sensitivity might be valued if one of the purposes of the allocation formula is to be responsive to rapidly rising needs.

6. Effects of Global Budgets

The preceding simulations have been based on the assumption that each unit's allocation is independent of the allocations received by all other units. Often, this assumption is unrealistic. A common situation is that in which there is a fixed global budget for a program, so that the funding of each unit is dependent on the ''demand'' for funding by each of the other units. This might appear to be the case for programs such as the Title I education program. We must note, however, that the assumption of a fixed global budget may also be an oversimplification, since Congress may respond to an increased demand for funds (due to increasing poverty rates) by increasing the total amount available for distribution. Congress may also increase the total amount available for distribution in response to an allocation that reduces funds to some units by more than it collectively can tolerate, even if poverty rates have not increased on average. For the analysis of this section, nonetheless, we assume a fixed global budget.

The key intuition required in addressing the effects of the interactions among allocations to different units is that such interactions are mediated through some parameters of the fund allocation formula that are implicit functions of aggregate demand and available funds. For example, suppose that a globally budgeted amount is distributed among units in proportion to the number of individuals who fall under a criterion of need. Then if the population eligible for aid were to be overestimated for some unit (holding estimates for other units constant), the amount distributed per eligible person (the key parameter of this funding formula) would be driven down and this would affect the allocations for other units. In general, if the number of units is large, the aggregated magnitude of the effects on allocations due to applying a nonlinear formula with imprecise data may be close to its expectation, simply because it is the average of contributions from a large number of units. Hence it may be highly predictable from mathematical calculations or simulations of bias such as those illustrated in the earlier sections of this article. The total effect of sampling error may then be calculated by estimating the effect of these biases on the formula parameter and consequently the expected effect on the estimate for the single unit of interest.

We now restate this argument more formally. Let $f(x_i,\theta)$ be the formula allocation for unit i which has a measurable characteristic x_i related to need, if the overall formula parameter is θ. The parameter θ might be calculated in the process of applying a formula in which θ is not specified; for example, if a fixed budget is distributed over a variable pool of recipients, the amount per recipient depends on the number of recipients. For simplicity of presentation we assume that f is nondecreasing in both x_i and θ, i.e., needier units receive more than they would if they were less needy and increasing the formula parameter increases (or leaves constant) the amount allocated to each unit. Simple illustrations are the following:

(i) $f(x_i, \theta) = x_i \theta$, simple proportional allocation, where x_i is the number in need in the unit. In this formula, θ is simply the amount allocated per needy person.

(ii) $f(x_i, \theta) = w_i h(x_i) \theta$, where w_i is a measure of size (e.g., total population), and $h(x_i)$ is a possibly nonlinear function of a rate. (For example, the function $h(x_i) = 0$ for $x_i < c$, $h(x_i) = x_i$ otherwise, represents a rate threshold for receiving an allocation.) We regard w_i as a fixed quantity, which does not need to be included in the formula explicitly. Section 5.2 represents a special case of this class of formulas.

(iii) $f(x_i, \theta) = a w_i x_i$ for $x_i > -\theta$, 0 otherwise, with a a predetermined constant. Suppose again that x_i represents a rate. Under this formula, the neediest units, defined as those exceeding a threshold rate of need $-\theta$, receive a predetermined allocation a per needy person, while those below the threshold receive nothing. (We define the threshold as $-\theta$ to maintain the condition that f is increasing in θ.) The threshold rate ''floats'' in the sense that it is determined by the level at which the budget is exhausted, rather than being fixed in advance.

If x_i is estimated from a sample, the allocation to unit i is $f(x_i + \epsilon_i, \theta)$, where ϵ_i is measurement (sampling) error. The sampling distribution of ϵ_i depends on x_i and some characteristics s_i of the measurement, including its sampling standard error and perhaps other properties. Finally, suppose that the expected allocation for a unit, taking the expectation over the distribution of ϵ_i given s_i, is $f_s(x_i, s_i, \theta)$. This is the quantity that was studied through the simulations in Section 5; in particular, we were concerned about the sensitivity of $f_s(x_i, s_i, \theta)$ to s_i.

Given a fixed global budget A, the value of θ used in the allocation is determined by the relation $\Sigma_i f(x_i+\epsilon_i, \theta) = A$. If the number of units is fairly large, we may approximate the sum by its expectation, $\Sigma_i f_s(x_i, s_i, \theta) = A$. Hence the expected allocation to unit i, $f_s(x_i, s_i, \theta)$, is affected by the sampling properties for the measurement in that unit and by the effect of sampling properties averaged over other units.

It is difficult to draw any fully general conclusions about the effect of sampling error on allocations to each unit. It is possible to draw fairly general conclusions, however, for allocation formulas of the forms (i) and (ii) above, where θ appears as proportionality constant in the formula. In that case, the ratio of allocations for any two units is free of θ; furthermore, the ratio of the ratio of expectations to the ratio of correct allocations is also free of w_i. The latter ratio (for comparison of two domains labeled *i, j*) is given by

$$(f_s(x_i, s_i, \theta)/(f_s(x_j, s_j, \theta))/(f(x_i, \theta)/f(x_j, \theta)) = (h_s(x_i, s_i)/h(x_i))/(h_s(x_j, s_j)/h(x_j))$$

where h_s is defined analogously to f_s. The proportional bias $h_s(x_i, s_i)/ h(x_i)$, and the way it is affected by sampling properties s_i, is precisely what the previous simulations studied. Hence we conclude that for a large class of formulas, the results we have obtained for single areas apply straightforwardly to comparisons of the relative effect of sampling error in different units. We anticipate that in many situations that do not quite fit the structure of (ii), fairly similar results would nonetheless apply, i.e., units whose sampling properties augment their expected allocations the most with fixed values of θ are also advantaged when they must share a global budget with other units.

7. Conclusions

From a legal and formal standpoint, modification of the estimation procedure and modification of the formula are two entirely different enterprises. There are good reasons from the standpoint of the division of labor among the agencies of government to maintain this distinction. In actuality, though, the formula, estimation procedure, and data sources are parts of a coherent whole. As illustrated in Section 4, the distinction between the estimation procedure and the formula is often entirely arbitrary, an expression of the same calculation with different labels. Consequently, it would be shortsighted to give attention to estimation and data collection while ignoring the formulas. When new data sources become available, our goal should not be simply to devise an estimation procedure that replicates allocations that were obtained with outmoded data sources. Indeed, the procedures that were used with older sources might reflect the limitations of those data rather than an intention to obtain a specific outcome.

As the illustrations suggest, interactions among the properties of the data, estimation methods, and funding formulas may produce unanticipated and sometimes undesirable effects. The long-term effects of linear estimators and formulas are fairly predictable. The allocations obtained when estimators or formulas contain nonlinearities, however, may be greatly affected, even on the average and in the long run, by sampling variances. This is problematic, because it almost inevitably leads to situations in which larger or smaller units tend systematically to get more than their proportional shares, other factors (poverty rates) being held constant. Furthermore, although decisions about allocation of sample should be made on technical grounds related to optimizing the overall accuracy of the survey, such decisions have implications for outcomes for specific units when the outcomes are sensitive to variances. Such a link between methodological choices and outcomes puts the data-collection and estimation agencies of government in an untenable position. Similar issues are discussed in the context of estimating penalties imposed on states for error detected by audits of welfare and food stamp payments, in a set of articles by Kramer (1990), Hansen and Tepping (1990), Fairley, Izenman, and Bagchi (1990), and Puma and Hoaglin (1990).

Widely used nonlinear allocation procedures include hold harmless provisions and thresholds. These can be replaced to some extent by linear estimation and allocation procedures that accomplish some of the same goals but have less paradoxical properties, specifically less sensitivity of long-run allocations to sample design, so their use should be reconsidered. On the other hand, some nonlinear and indirect procedures, such as empirical Bayes estimation, can be shown to produce estimates with improved accuracy relative to linear and direct estimators. Therefore, the indirect estimators are likely to be useful when high-precision direct estimators are not available. These indirect estimators tend to have sampling characteristics (such as variation from year to year) that are less dependent on sample size than those of direct estimators; on the other hand, they may be affected by model biases that tend to persist over time. How the properties of indirect estimators interact with allocation formulas needs to be better understood as indirect estimators become more widely used.

Funding formulas are often ingenious ad-hockeries, hammered out from a political process based on compromise. Although notions of equitable and efficient allocation of resources are implicit in them, they do not, by themselves, define those notions. It is the responsibility of those who generate data and implement the formulas and best understand how they work together in practice to consider the ways that new procedures and data change the formulas' effect and to suggest revisions to formulas that best serve their original intent.

8. References

Alexander, C. H. (2001). The American Community Survey for Economic Analysis. Washington, DC, U.S. Bureau of the Census.

Citro, C. F., Cohen, M. L., et al. (eds.) (1997). Small Area Estimates of School-age Children in Poverty. Interim Report I: Evaluation of 1993 County Estimates for Title I Allocations. Washington DC, National Academy Press.

Citro, C. F. and Kalton, G. (eds.) (2000a). Small Area Estimates of Children in Poverty: Evaluation of Current Methodology. Washington, DC, National Academy Press.

Citro, C. F. and Kalton, G. (eds.) (2000b). Small Area Estimates of Children in Poverty: Priorities for 2000 and Beyond. Washington, DC, National Academy Press.

Fairley, W.B., Izenman, A.J., and Bagchi, P. (1990). Inference for Welfare Quality Control Programs. Journal of the American Statistical Association, 85, 874–890.

Fay, R. E. and Thompson, J. (1993). The 1990 Post Enumeration Survey: Statistical Lessons, in Hindsight. Proceedings of the Bureau of the Census Annual Research Conference. Washington, DC, U.S. Bureau of the Census.

Ghosh, M. and Rao, J. N. K. (1994). Small Area Estimation: An Appraisal. Statistical Science 9, 55–76.

Hansen, M.H. and Tepping, B.H. (1990). Regression Estimates in Federal Welfare Quality Control Programs. Journal of the American Statistical Association, 85, 856–863, and 871–873.

Kramer, F.D. (1990). Statistics and Policy in Welfare Control: A Basis for Understanding and Assessing Competing Views. Journal of the American Statistical Association, 85, 850–855.

Puma, M.J. and Hoaglin, D.C. (1990). Food Stamp Payment Error Rates: Can

State-Specific Performance Standards Be Developed? Journal of the American Statistical Association, 85, 891–899.

Schirm, A. L. (2000). The Evolution of the Method for Deriving Estimates to Allocate WIC Funds. Presented at Workshop on Formulas for Allocating Program Funds, National Academy of Sciences, Washington, DC.

Schirm, A. L. and Long, C. (1995). Fund Allocation and Small Area Estimation in the WIC Program. Proceedings of the American Statistical Association, Section on Survey Research Methods, Alexandria, VA, 555–560.