Pergamon

0038-0121(95)00022-4

Socio-Econ. Plann. Sci. Vol. 30, No. 2, pp. 103–119, 1996

0038-0121/96 $15.00 + 0.00

Optimal Consolidation of Municipalities: An Analysis of Alternative Designs

MALACHY CAREY[1], ASHOK SRINIVASAN[2] and ROBERT P. STRAUSS[3]

[1]Faculty of Business and Management, University of Ulster, Northern Ireland, U.K.

[2]Krannert School of Management, Purdue University, West Lafayette, IN, U.S.A.

[3]H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA, U.S.A.

Abstract—In this paper, we present an analytic framework for the geographic aggregation of municipalities into larger and more populous municipal districts in order to reduce the costs of providing public services. We first develop a simple model for estimating the cost of providing local government services, and then introduce the notion of the "optimal" size of a municipality. The model allows us to determine the optimal number of districts for a fixed population size in a given geographic area, and to arrive at the extent of cost savings possible in the absence of additional considerations. We statistically estimate a municipal cost function and use this to construct a least cost consolidation plan subject to certain constraints on how far the socioeconomic characteristics of the new, consolidated districts can differ from those of their constituent parts prior to consolidation. The constraints make the combinatorial problem of consolidation computationally more tractable. More important, the constraints reflect the fact that the plan would have to be accepted (and voted on) by the municipalities to be consolidated, and there is evidence that citizens are much less likely to accept a consolidation plan that greatly disturbs the existing tax and service levels. In this formulation of the consolidation problem, there is thus a tradeoff between financial efficiency (the economic benefit of consolidation) on the one hand, and political feasibility (the socioeconomic constraints) on the other. We illustrate the use of the model by developing a least cost consolidation plan for the municipalities in Allegheny County, PA, and compare this plan with two alternative consolidation schemes proposed unsuccessfully in the past. Our consolidation scheme scores better than the other two alternatives on costs and on satisfying the socioeconomic constraints. Also, it is reassuring that is has strong parallels with some aspects of the other two schemes, since these reflect existing school district and cooperative schemes that have evolved over time. Copyright © 1996 Elsevier Science Ltd

INTRODUCTION

Consolidation has been periodically proposed as a means for local governments to cope with slow growing or declining revenues. The major benefit of consolidation is typically argued to be gains from scale economies. A number of factors influence the extent to which larger governments may be able to provide services at lower cost *per capita*. Among them are the capital intensiveness of the production process (e.g. water, garbage, sewage), the extent of factor indivisibility, the availability and cost of scale-related technologies, the delivery distance and related costs, discounts available for purchasing in quantity, and the random nature of service calls for services such as fire and police [1, 2]. Small size generally means that individual jurisdictions cannot benefit from possible economies of scale in the provision of various public services, though small size can also mean that the provision of such services, while more expensive, is more responsive to residents' needs. (See Ref. [11] for a summary of empirical studies.)

Though there will always be a debate between those who favor larger or smaller units of local government, we present a model to help focus this discussion. Given a region (a metropolitan area) having a large number of local government districts, we show how to construct combinations of these which would minimize (optimize) the costs of providing local public services, and, at the same time, satisfy certain constraints that reflect political acceptability. We certainly do not claim to resolve all issues concerning what is "optimal", or feasible, and various limitations and qualifications are mentioned throughout the paper.

Optimal consolidation is closely related to optimal political redistricting [8, 10, 17], to defining geographic boundaries for health regions [25], to allocating and utilizing park and recreation facilities [15] and to delineating school district boundaries [12]. All of these problems are concerned with clustering a large set of smaller zones into a smaller set of larger contiguous zones [16]. However, they differ substantially in terms of the criteria used for such clustering. For example, in political redistricting a typical objective is to ensure that district sizes (populations) do not differ by more than a certain amount, and each district consists of a compact contiguous area. In defining health care regions, the objective may be to maximize the proportion of the state's population that receives health care in its region of residence [25]. The municipality consolidation problem, on the other hand, is usually driven by economies of scale in the provision of public services.

In this paper, we are concerned mainly with voluntary consolidation, since most municipal consolidations in the U.S. are voluntary: of the 34 consolidations in the U.S. in the 1980s, all were voluntary (see Ref. [14]). However, the approach in this paper is also applicable to involuntary consolidations, since it would be quite reasonable for the relevant decision makers to adopt objectives and constraints similar to those presented here. We are not concerned with annexations in the U.S. since this implies annexing unincorporated areas and involves only one or very few municipalities, so that the problems considered in the present paper do not arise. Incidentally, in almost all of the 34 consolidations in the 1980s the component municipalities had populations of less than 5000, which is similar to those considered in this paper.

We develop a simple model of estimating the cost of providing local government services, and introduce the notion of the "optimal" size of a municipality. Municipal cost functions obtained from this model are then used to construct an optimal consolidation plan subject to socio-economic constraints. These constraints make much more tractable the computational problem of finding the best sets of districts to combine. More importantly, the constraints are designed to rule out combining districts that differ greatly in certain economic characteristics (their tax base, income levels etc.). The reason for this is that such differences make it more difficult or even impossible to obtain consent to voluntary consolidation. In most states (including Pennsylvania), proposed consolidation schemes have to be voted on by the population of the component municipalities prior to implementation. These constraints can also be derived from the literature on local public choice (see, for example Refs [4–7, 18] and the subsection below concerning socio-economic constraints).

We illustrate use of the model by developing an optimal consolidation plan for Allegheny County, Pennsylvania, and by comparing this plan with alternative consolidation schemes that have been historically suggested by municipal reform groups. The alternatives considered are based on the traditional approaches to addressing the municipal fragmentation problem, both nationally and in Allegheny County. These are: (1) county-wide or metropolitan government; (2) consolidating those groups of municipalities that already have significant cooperative arrangements in providing services; and (3) consolidating on the basis of school district boundaries.

The remainder of this paper is organized as follows. In the next section, we set out a simple model of the costs of administering a district, and in the following section set out a model for generating an "optimal" (least cost) consolidation scheme. The subsequent section applies these models to data for Allegheny County, PA. The resulting optimal consolidation scheme is compared in the following section to two alternative consolidation schemes, namely: (i) using school district boundaries; and (ii) using Council of Government (COG) boundaries. A final section offers a summary and conclusions.

AN ECONOMIC MODEL OF LEAST-COST MUNICIPALITY SIZE

We introduce here a simple model for estimating local government service costs so that we can estimate the service cost for any proposed new district. We also introduce the notion of the "optimal size" of a municipality. These ideas are discussed further and utilized in later sections. In this section, we assume that a region can be divided up arbitrarily into any number of districts of any size, with no restrictions on what areas can be combined. This allows a very simple calculation of the optimal number of districts. In the following sections we restrict the redistricting so that the region can only be divided into combinations of currently existing municipalities. We also introduce restrictions on which municipalities can be combined.

Municipal costs and optimal size of a municipality

We assume that the population P of the region is independent of the number or size of the districts formed within the region. With population P given, we can minimize the total cost of service in the region by minimizing the *average* cost of service *per person* in each district in the region. Suppose the region is divided into districts, each of population size p, so that the number of such districts is $n = P/p$. Let the total cost of serving any one district be $f(p)$. Any variables, other than the district population p, which affect the district cost c can be assumed to be already evaluated and treated as constants in this function [see also the section below on "Some econometric evidence on $f(p)$"]. The reason for this is that we are here concerned only with how costs vary with district population, since we are considering combining districts, which will immediately change their population. The average cost of service per person in the region is thus $c = f(p)/p$, and the total cost of serving all districts is $C = Pc = nf(p)$. Assuming $f(p)$ is differentiable, the minimum of c (and of C) is characterized by:

$$dc/dp = 0 = f'(p^*)/p^* - f(p^*)/p^{*2}.$$

Hence $f'(p^*) = f(p^*)/p^*$, where p^* denotes the optimal (cost minimizing) value of p. This corresponds to a well-known result in economics, namely, that the minimum average cost occurs where the average cost $f(p)/p$ equals the marginal cost $f'(p)$. The minimum of the continuous average cost function $[c = f(p)/p]$ occurs at population p^* which yields an "optimal" (least cost) number of districts $n = P/p^*$. But this number is likely to be non-integer, hence not feasible. However, if the continuous average cost function is convex we can find the optimal integer number of districts by adjusting the district population up or down to give the nearest integer n—round in both directions and then check which gives the lower cost.

Implications for consolidation

If the average cost curve $f(p)/p$ is "U" shaped, then the population corresponding to the minimum of the "U" is the optimal (least-cost) population. If this minimum occurs at a population several times less than the population of the region, then it is optimal to divide the region up into several districts. It is worth considering how fixed overhead costs and varying marginal costs can generate a "U" shaped average cost curve (see Refs [19] and [21]). [Note that minimizing the total cost for the region as a whole $[Pf(p)/p]$ does not imply minimizing the total cost $f(p)$ for each district taken *separately*. Minimizing for each district separately would yield more (smaller) districts than minimizing for the region as a whole.]

Case 1: Fixed overhead cost and linearly increasing marginal cost per capita, for each district. Let there be a fixed initial cost, $a_1 > 0$, for providing any level of service and let the cost of serving each additional person be $mc = a_2 + a_3 p$. Then the cost of serving a district is:

$$f(p) = a_1 + \int_0^p (a_2 + a_3 p)\, dp = a_1 + a_2 p + a_3 p^2/2, \quad (1)$$

and the cost of serving all $n = P/p$ districts is thus $C = nf(p)$, i.e.

$$C = (a_1/p + a_2 + a_3 p/2)P. \quad (2)$$

The cost-minimizing population size p^* for each district is given by letting:

$$dC/dp = 0 = -Pa_1/p^2 + Pa_3/2;$$

hence,

$$p^* = (2a_1/a_3)^{\frac{1}{2}}. \quad (3)$$

For a minimum, of course, $d^2C/dp^2 \geqslant 0$, that is, the marginal cost must not decrease when the population increases. Equation (3) implies that the optimal population of each district will be "larger" if the fixed cost of serving the district is larger, and will be "larger" if the (marginal) cost a_3 of serving each additional person or population unit is smaller. The optimal number of districts is obtained by computing (P/p^*) and rounding up or down to an integer, rounding in both directions to see which gives the lower total cost. Rounding gives the optimal solution since the cost function is convex $(d^2C/dp^2 \geqslant 0)$. If $p^* > P$, then the optimum is to form the region into a single district, to population size P.

Case 2: Fixed overhead cost and constant marginal cost per capita, for each district (i.e. $a_3 = 0$). This implies that average costs *per capita* decline indefinitely, in which case the optimal solution is a single district consisting of the whole region. In the notation of Case 1, the total cost for a district is $f(p) = (a_1 + a_2 p)$; hence the total cost for the region is $C = n(a_1 + a_2 p) = (na_1 + a_2 P)$ which is minimized by setting $n = 1$.

Case 3: Zero overhead cost and constant marginal cost per capita, for each district (i.e. $a_1 = 0 = a_3$). Here, the total cost for a district is $f(p) = a_2 p$; hence, the total cost for the region is $C = (a_2 p)n = a_2 P$. In this case there is no unique optimal size for each district: we can divide the region into any number of districts without affecting the total cost.

LEAST COST CONSOLIDATION WITH SOCIO-ECONOMIC CONSTRAINTS

A consolidation plan for a region consists of a set of districts chosen so as to exactly cover the total area of the region without any overlapping.† In other words, in any consolidation plan, each municipality (or other sub-area) must be included in one and only one of the newly-formed districts. Since there is usually a very large number of consolidation plans that would satisfy this condition or definition, it is natural and desirable to choose, out of all these possible plans, one which also optimizes some objective, and satisfies any additional constraints. Thus, we define an *optimal consolidation plan* as one which:

(a) satisfies any constraints on forming districts, e.g. topological constraints (such as contiguity and compactness) and socio-economic constraints (such as avoiding causing large changes in the tax base within the district); and

(b) yields the lowest (or highest) possible value of some objective (e.g. the cost of providing local government services).

A *feasible* consolidation plan is one which satisfies the constraints (a), but does not necessarily optimize (b).

Following Ref. [8], the generation of a consolidation plan can be decomposed into two stages, paralleling (a) and (b) above:

Stage 1. Generate a list of "candidate" districts, satisfying the conditions/constraints of (a) above.

Stage 2. From the set of candidate districts generated in Stage 1, select a set of districts (a consolidation plan) such that: (i) each municipality is included in one and only one district; and (ii) the plan achieves objective (b). In Stage 1, each sub-district (municipality) may be included in many different candidate districts, but Stage 2(i) ensures that only *one* of these candidate districts will appear in the consolidation plan.

The output from Stage 1 is a list of $j = 1, \ldots, J$ districts, each having a list of (at most m) members. This can be stated as a set of 0–1 numbers:

$$a_{ij} = \begin{cases} 1 \text{ if municipality } i \text{ is a member of district } j, \\ 0 \text{ if municipality } i \text{ is not a member of district } j. \end{cases}$$

Using these 0–1 indicators (a_{ij}'s), Stage 2 above can be stated as follows. Let x_j be a 0–1 integer variable such that,

$$x_j = \begin{cases} 1 \text{ if district } j \text{ is chosen to be part of the optimal plan,} \\ 0 \text{ otherwise.} \end{cases}$$

Then, the total cost of providing local government services, for all districts in the consolidation plan is $\Sigma_{j=1}^{J} c_j x_j$, where c_j is the estimated annual cost of providing local government services for (candidate) district j. From the previous section, $c_j = f(p_j)$, where p_j is the population of district j. [Estimates for $c_j = f(p_j)$ are given in the next section.]

†In Pennsylvania, various classes of cities, boroughs and townships may exist within a county geographic area; however, they do not overlap.

The requirement that municipality i be included in one and only one district can be stated as $\Sigma_{j=1}^{J} a_{ij} x_j = 1$, so that the consolidation plan that minimizes the total cost of local government services is obtained (see Ref. [8]) by solving,

$$\text{P1:}\quad \text{minimize } w = \sum_{j=1}^{J} c_j x_j, \tag{A1}$$

subject to:

$$\sum_{j=1}^{J} a_{ij} x_j = 1 \quad i = 1, \ldots, m, \tag{A2}$$

$$x_j = 0 \text{ or } 1, \quad j = 1, \ldots, J. \tag{A3}$$

AN APPLICATION TO ALLEGHENY COUNTY, PENNSYLVANIA

We will apply the above district consolidation approach to Allegheny County in Pennsylvania, using data from 1980. (Incidentally, the population and distribution of population in the county have not changed much since then.) The county had a population of 1,023,825 in 1980, excluding Pittsburgh which has a population of 423,938. Allegheny County is made up of 130 non-overlapping municipalities, or 128 excluding Pittsburgh and Neville Island; hence, the average size of the municipalities is small. Only two other urban county areas in the U.S. of comparable population size have, on average, fewer residents per municipality: Cook County, Illinois and Harris County, Texas (see Ref. [26]). Of 130 municipalities in Allegheny County, 90% of them (or 117) had a 1980 population of 5500 or less. Overall, Allegheny County had 1.45 million residents. Some statistics on Allegheny County are given in the Appendix.

There have been several attempts to create a metropolitan government in Allegheny County that would have merged all municipalities and the county into a single unit. Constitutional amendments in 1928 and 1933 provided procedures for the consolidation of the municipalities in Allegheny County (see Article XV, of the Pennsylvania Constitution). While the conventional wisdom in Allegheny County is that metropolitan government is not politically feasible, there is a growing awareness that it is increasingly difficult for many local governments to function effectively in the face of declining Federal aid (e.g. the elimination of General Revenue Sharing), and a stagnant or declining real property base.

In the remainder of this section, we estimate parameters of the municipal cost function $f(p)$ (discussed in a previous section) for Allegheny County using 1980 Census of Population data, and provide an estimate of the maximum expenditure savings possible and of the optimal number of new districts.

Some econometric evidence on f(p)

Derivation of the optimal district size depends on the empirically estimated district cost function $c = f(p)$. To estimate this we obtained a 1980 cross-section of the total expenditure per municipality and used this as a proxy for total operating costs. The corresponding population p is from the 1980 Census of Population. We used this data and ordinary least squares to estimate $c = a_1 + a_1 p + a_3 p^2$, as in equation (1) above, across all municipalities in Allegheny County. This yielded:

$$\text{(E)}\quad c = \$148{,}020 + 143.87p + 0.001013p^2 \quad (R^2 = 0.99)$$

All coefficients were statistically significant at the 95% confidence level.

Using this estimated equation, and following Case 1 from the subsection on "Implications for consolidation" above, the total cost for the whole country is $C = f(p)n = f(p)(P/p)$, or:

$$C = P(148{,}020/p + 143.87 + 0.001013p).$$

Minimizing this cost with respect to p yields an optimal sized municipality of $p = (148{,}020/0.001013)^{1/2} \approx 12{,}000$ population, with a municipality cost of $f(p) = \$2021$ millions in 1980 dollars. [Recall that we chose p so as to minimize $nf(p)$, and *not* $f(p)$. Indeed, the above municipality cost function $f(p)$ is upward sloping and hence has a minimum only at p = zero.] This cost function C is "U" shaped, declining steeply when p is small (say under 5000) and increasing slowly for populations over 12,000. Thus, most cost savings from consolidation are likely to be achieved from combining a large number of small districts, e.g. those under 5000.

We examined other specifications for district costs. For example, dropping the p^2 term yielded the ordinary least squares regression equation $c = -\$2{,}957{,}650 + 557.43p$. These coefficients were statistically significant though not as significant as with the p^2 present. Further, this regression equation is not plausible as it implies that small districts have negative costs, and hence that a large proportion of municipalities in the county have negative costs, which they do not. As noted earlier, a large proportion of districts in Allegheny county are small, and do not, of course, have negative costs. Introducing other socio-economic variables into the above linear-in-population equation created problems of poor prediction, as well as instability in the coefficients due to multicollinearity. We also introduced various socio-economic variables (e.g. income or aged population), as well as population, into the quadratic-in-population equation (E). These additional variables were usually not significant, and in some cases (as expected) highly collinear with population. Also, it is of particular interest here that introducing these other variables did not greatly change the predicted optimal size district. For example, when we introduced the number of persons over age 65 and aggregate family income (as well as p and p^2) as explanatory variables, the optimal size population rose slightly to $p = 13{,}500$. When we introduced the number of persons over age 65 and the aggregate tax base (as well as p and p^2) as explanatory variables, the estimated cost equation yielded an optimal population size of $p = 11{,}000$. These are comparable to the least cost population of approx. 12,000 for the cost equation (E).

With respect to the quadratic specification, we observe that in each of the regression equations we ran with p and p^2 as variables, the coefficient of p^2 remained near 0.001. This stability gives us an added reason for retaining the p^2 term. Finally, it is worth noting that even it there were significant explanatory variables missing from the district cost equation, these would not necessarily affect the estimated coefficients of p and p^2. If the missing variables are uncorrelated with population and the regression equation is linear in the parameters and separable in population, then ordinary least squares still yields unbiased estimates of the coefficients (see Ref. [24]).

Level of service

If there are differences in levels of service in each districts, one would expect the total cost for each district to depend on the level or standard of service provided, as well as on the district size (population). That is, $c = g(p, s)$ where s is the level of service *per capita* in the district and p is the district population. But estimates of the level or standard of service are difficult to obtain. If we do not have data on the level of service in each district, and estimate $c = f(p)$ rather than $c = g(p, s)$, what effect does this have on the estimated cost equation? This is discussed in econometrics under specification error or omitted variables (e.g. [13], and the original article on the topic [23]). Suppose that the true relationship between a dependent variable and several independent variables is linear in its parameters, e.g. $y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + u$, where u is a random disturbance term and, as usual in linear regression, it is assumed that $E(u) = 0$. Also suppose that one or more of the independent variables are omitted when estimating the parameters of the equation using ordinary least squares (OLS). If the omitted variables are not correlated with any of the included variables, then the estimated coefficients for the included variables will be unbiased (e.g. see Ref [13], p. 169). Note that this result does not require that the equation be linear in the variables, e.g. the variable may be defined as a prespecified function of x_1 or x_2, say $x_3 = x_1^2$.

To relate this to the present context, suppose that the true cost equation is of the form:

$$g(p, s) = a_1 + a_2 p + a_3 p^2 + b_1 s + b_2 h_2(s) + b_3 h_3(s),$$

where $h_2(s)$ and $h_3(s)$ are prespecified nonlinear functions of s, e.g. $h_2(s) = s^2$ or $h_2(s) = s^{\frac{1}{2}}$ or $h_2(s) = 1/s$. This equation is linear in the parameters to be estimated (a_1, a_2, a_3, b_1, b_2, b_3). Hence, if s, $h_2(s)$ and $h_3(s)$ are uncorrelated with p and p^2, estimating the parameters of $f(p) = a_1 + a_2 p + a_3 p^2$ by OLS will give unbiased estimates. It seems reasonable to assume as an approximation that the current levels of service s are uncorrelated with p and p^2, in which case nonlinear functions of s [e.g. $h_2(s)$ and $h_3(s)$] are uncorrelated with p and p^2. This result can easily be generalized to allow for more than one service level variable.

The standard error of the estimates will be larger when estimating $f(p)$ rather than $g(p, s)$, since, for any population level in $f(p)$, there will be a range of possible service levels and hence a wider range of possible expenditures. However, as we have a large number of observations (128 municipalities) and only three parameters to estimate in equation (*E*), we have 125 df. Hence, the larger standard deviation is not a problem so that (see above) the estimated parameters of (*E*) are statistically very significant.

To measure levels of standards of service we could use such items as availability or response times of police, ambulance, fire service, etc. provided by the municipalities, or measures of health care, or sanitation etc. However, for most of the municipalities such data are simply not available. We contacted and interviewed many municipal officials and found that records were not kept or were not available. Note that we cannot use *per capita* expenditure as a measure of level of service, since this contradicts the basic idea in the paper, that *per capita* costs or expenditures reflect (dis)economies of scale.

It is worth briefly discussing how we would use district cost equations $c = g(p, s)$ if they were available. Recall that in this paper we use the district cost equations only to predict the total costs of potential new consolidated districts, and for existing districts. To compute district costs from $c = g(p, s)$ we would first have to specify a level of service for each district. There are basically two ways to do this:

Approach 1. Specify a service level for each district as close as possible to the actual observed service level for the district. This has the advantage that it causes least disturbance to the existing choices and preferences of the total population. It maintains interjurisdictional equilibrium.

Approach 2. Experiment with service levels different from the actual or observed level for each district. For example, let all districts have the same service level, equal to the mean or minimum or maximum of the actual or observed service levels. Though this may be an interesting exercise, it raises issues and problems which go beyond the scope of the present paper.

An estimate of the maximum possible net saving of expenditures

In order to put the cost savings that result from various consolidation schemes in perspective, we compute here an upper bound on the net saving that might be possible through consolidation. In a previous section we found that a district size of 12,000 yielding a minimum of the "U" shaped short-run *per capita* cost function. Letting all districts have populations of exactly 12,000 would (if it were practicable): (a) minimize the predicted total expenditure for the county; and, hence, (b) maximize the predicted total expenditure saving, relative to current expenditure. These two quantities [(a) and (b)] are computed as follows. Substituting the 12,000 population into the total expenditure equation for each district yields:

$$e^* = 149{,}020.0 + 143.87(12{,}000) + 0.001013(12{,}000)^2 = \$2{,}021{,}313 \approx \$2{,}000{,}000.$$

The number of such optimal sized districts is thus,

$$n^* = (\text{county population})/(\text{population of optimal sized district})$$

$$= 1{,}023{,}825/12{,}000 = 85.317 \approx 85,$$

where the predicted total expenditures for the county, based on districts of size 12,000 is thus $n^* \cdot e^* = (85) \cdot (\$2{,}021{,}313) = \$172.3$ million. The *observed* total municipal expenditure for the county (for the year whose data was used in the above estimation) is \$185,121,609. Thus, the estimated total expenditure *saving* for the county equals, (\$185,121,609–\$172,300,000) = \$13,821,609 or approx. \$13.8 million (7.45%).

Note that this is the maximum possible predicted expenditure saving, based on (a) assuming that municipalities are divisible; (b) using the *per capita* cost equation above; and (c) given the above current county population. Any other consolidation scheme using districts with populations not exactly equal to 12,000 will yield predicted total expenditure savings of *less* than the above maximum of \$13.8 million. In particular, this applies to each of the three consolidation schemes discussed below since they do not allow municipalities to be split, and introduce additional criteria beyond minimizing costs *C*.

Spatial and socio-economic constraints on consolidation alternatives

We consider here in more detail the spatial and socio-economic criteria that we utilize empirically to constrain the choice of municipalities to combine into new districts. We first require that the municipalities to be combined into a new district should be contiguous. For each municipality we drew up a list of contiguous municipalities. If two municipalities touch at only one point, or are separated by rivers or other barriers, we can consider them not contiguous. The main example of this is that municipalities are considered not contiguous it they are separated by any of the three main rivers that run through the County: the Allegheny and Monongahela rivers join in the middle of Allegheny County to form the Ohio River. This constraint allows us to partition the county into three regions, and yields a natural decomposition of the problem into three smaller sub-models, which is computationally much more tractable.

We next require that the municipalities to be combined into one district should satisfy certain socio-economic constraints (concerning tax base, income levels and public expenditure). These constraints may be viewed as empirical measures of political feasibility. Wide disparities in, for example, median family income or tax base, not only imply that two municipalities are unlikely to merge voluntarily, but also that to try to do so would probably fail electorally.

These constraints can also be derived from the literature on local public goods, as follows. In this literature, it is assumed that we observe a distribution of the population within a region to be in equilibrium *vis-à-vis* their preferences for local public goods and the public service offerings of competing municipalities. That is, we presume that the population has sorted itself among municipalities on the basis of their preferences, including those for public services and their ability and willingness to finance such services (see for example Refs [6] and [9]). Combination to achieve cost savings is potentially attractive to a household, if a municipality can be combined with another similar municipality having similar preferences for public services, since then the same basket of public services can be obtained at a lower cost for both municipal populations.

A review of the local public goods literature indicates four empirical factors other than population size that are likely to affect the taste for public services and the ability and willingness to finance them, which drives this sorting out process. These are: (1) the median family income for each municipality; (2) the *per capita* tax base for each municipality; (3) the *per capita* spending for each municipality; and (4) an indicator of the extent to which the municipality contains an elderly population. We discuss each of these briefly in turn.

A necessary condition for equilibrium between local jurisdictions that is often stated in the literature is the so-called stratification condition, that each community is formed mainly by individuals with incomes in a single interval (see Refs [5, 6, 27]. Some evidence of stratification is provided by Williams *et al.* [28] and Wood [29]. This condition ensures that a newly formed district will not have substantial income differentials among its constituent municipalities. We operationalize this by requiring that the median family income in each of the constituent municipalities of a newly formed district is within 50% above or below the median family income of the district as a whole (Test 1).

The second factor, *per capita* property tax base, indicates the ability of each municipality to finance public goods. We ensure that newly formed districts have sufficient tax capacity, by setting a lower bound on this tax capacity (taxbase) as Test 2.

The third factor, *per capita* expenditures, is an indicator of the level of public goods provided. We wish to ensure that there will not be a substantial increase in expenditure burden for any particular municipality.

The last factor concerns preferences for one type of public service—services to the elderly. Age is an important determinant of the type of public services desired (see Ref. [22]). We wish to ensure that this does not vary excessively among municipalities that are combined into a single district. This is accomplished by setting as an upper bound the proportion of population over 65 yr old in a newly formed district.

The specific numerical implementation of the four socio-economic tests, using the data for Allegheny County, is as follows:

Test 1: Is the median family income for *each* municipality in the district $\leqslant 50\%$ above or below the weighted average of the median family incomes for the district as a whole?

Test 2: Is the "tax base" *per capita* ⩾\$2400? (This is the first quartile of the distribution of "tax bases" for all the municipalities in Allegheny County).

Test 3: Is the estimated expenditure *per capita* for the district *after* consolidation ⩽50% above the actual expenditure *per capita* incurred by any municipality in the district *before* consolidation?

Test 4: Is the percentage of the population over 65 years old ⩽17.5%? (This is the third quartile of the distribution of the population over age 65 in Allegheny County).

Generating the candidate districts and solving the optimization model

As set out in an earlier section, we find a least cost consolidation plan by a two stage process. In Stage 1 we generate a set of "candidate" districts, and in Stage 2 we find the subset of these which yields a minimum cost and ensures that each municipality is included in one and only one district. Special algorithms for solving this problem are available: see Ref. [8], and more recent work on solving set partitioning problems [3], of which the current problem is a special case.

To handle Stage 1, we developed an algorithm which is analogous to that in Ref. [8]. We did not use the algorithms from Refs [8] or [3], since we were here not very concerned with computing time, and wished to experiment with some types of constraints we found easier to code-up from 'scratch', rather than seeking to adapt and code-up existing algorithms. (Also, we didn't have a code for Refs [8] or [3] available.) Our algorithm starts, as in Ref. [8], with an arbitrary municipality and keeps adding and deleting municipalities, each time noting the district formed. If adding a contiguous municipality would exceed a population cut-off (see below) or breach a certain measure of "spread" or compactness (see below) then we do not include it.

The population cutoff referred to above is as follows. If candidate districts have a relatively high estimated cost we know that they are very unlikely to be included in the eventual optimal solution. The estimated cost is given by substituting population in equation (E). We chose a cutoff cost corresponding to a population of 50,000, which is more than three times larger than the population that yields a minimum cost in equation (E). Hence, when generating potential candidate districts, we stopped adding municipalities to a district if this would make its population exceed 50,000. We note that in the optimal consolidation plans (from Stage 2), the consolidated districts all had populations very much smaller than this cut-off value. This suggests that imposing the cost cut-off did not significantly affect the eventual solution, though it did reduce our search time.

The measure of spread or compactness referred to above is as follows. If municipality a is contiguous to municipality b we say a and b are one step apart. If municipality a is contiguous to municipality b which, in turn, is contiguous to municipality c, we say that a and c are at most two steps apart, and so on. Our measure of spread is that no two municipalities in a district should be more than n steps apart. We experimented with different values of n and chose $n = 5$. To illustrate the effect of this, suppose municipalities are like squares on a chess board. Then the spread restriction allows any contiguous combination which can be fitted in a 5×5 square, hence allowing up to 25 municipalities to be combined in a district. The largest number of municipalities in any district selected in Stage 2 below was 9. Hence, it seems that the spread constraint was unlikely to have much affect on the maximum number of municipalities in districts in the Stage 2 solution. However, the spread constraint does prevent districts consisting of long strings of municipalities.

Generating the set of candidate districts (Stage 1) for set partitioning problems is often computationally very time consuming as there may be an enormous number of potential candidate districts to be considered. In the present problem, the number of potential districts to be considered was reduced by the above population and spread cutoff constraints. It was also greatly reduced by the fact (stated earlier) that the three main rivers in the County act as natural boundaries. Thus we did not generate any potential candidate districts crossing these boundaries.

We generated potential candidate districts for the northern, eastern and southern sectors of the County respectively, and checked to see which of these districts satisfied all four socio-economic feasibility tests set out in the previous section. It turned out that some municipalities were not included in any districts, which satisfied the socio-economic tests. These municipalities were thus included as a separate candidate district, even though they did not satisfy all the tests. Including these districts, we found 264, 304 and 218 candidate districts for the northern, eastern and southern

Table 1. Sizes of integer programs

Sector	Northern	Southern	Eastern
No. of variables in the integer program	264	305	218
No. of constraints in the integer program	48	45	35

sectors, respectively. These are the candidate districts, which go forward to Stage 2 of the solution process.

Stage 2 of the solution process is to find the optimal subset of candidate districts; that is, solve program P1. This combinatorial problem is again greatly simplified by the fact that the County divides into three separate sectors, so that P1 decomposes into three smaller more tractable problems. In P1, the number of integer variables is equal to the number of candidates districts, and the number of constraints equals the number of municipalities. Table 1 shows the sizes of the resulting integer programs.

We solved these integer programs using the 0–1 integer programming facility of a standard linear programming package (LINDO [20] on a Vax 11/780). If the number of potential candidate districts (Stage 1 above) and/or the number of candidate districts (in Stage 2) had been very much larger, then the above solution approach would not have been tractable. In that case we could have adapted the algorithms already referred to, in Refs [8] and [3], as these are known to easily solve very much larger problems.

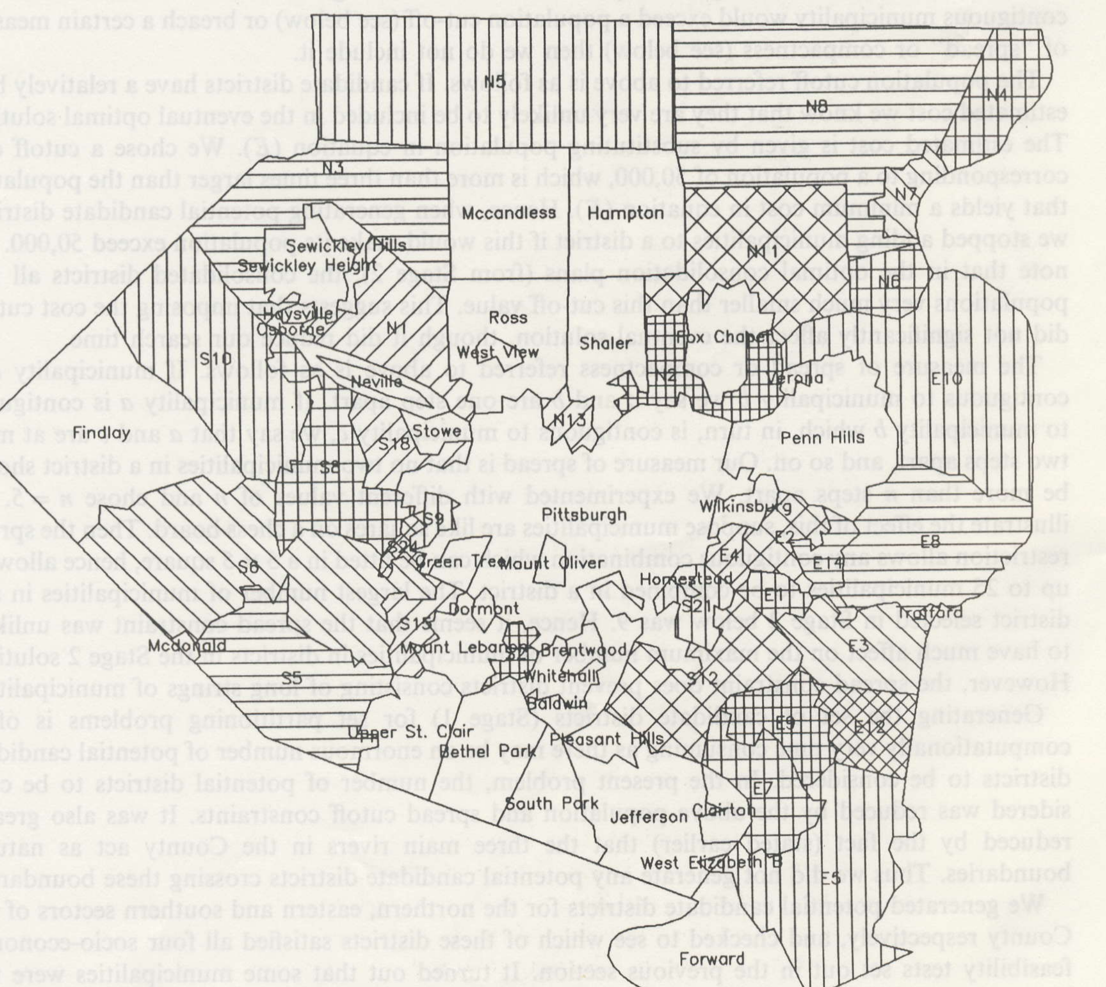

Fig. 1. The "Optimal" Scheme.

On solving Stage 2, we found that the 128 municipalities combined into 65 districts (see Fig. 1). In the next section, we discuss the characteristics of this consolidation scheme, and, to put it in perspective, we define two alternative consolidation schemes and compare the above scheme with each of these.

ANALYSIS OF THREE CONSOLIDATION SCHEMES

In this section, we report the empirical implications of three consolidation schemes, which combine in different ways, 128 out of Allegheny County's 130 municipalities. The City of Pittsburgh and Neville Island have been excluded from the consolidation experiments, the former because it would require deconsolidation into separate parts, and the latter because of geography. Deconsolidation is beyond the scope of this paper. The schemes which we consider are:

The "Optimal" Scheme. This is the scheme obtained in the previous section.

The School District Scheme. Combine municipalities into jurisdictions with the same boundaries as the school districts in Allegheny County. Since there are 45 school districts this yields the 45 jurisdictions. (With small geographic exceptions, school district boundaries in Allegheny County follow municipal boundaries or aggregations of municipalities).

The Council of Government (COG) Scheme. Combine municipalities into districts with the same boundaries as the Council of Governments (COG) in Allegheny County. These are agencies that provide sharing of services and thus allow the municipalities to benefit from economies of scale at an intermediate level of geographic aggregation, while allowing them to remain politically independent jurisdictions.

The COGs have no power to levy taxes, and have no elected representatives from the citizenry. They are voluntary organizations which were formed after World War II, when the move to consolidate all the municipalities of Allegheny County into one metropolitan government failed. (There are eight COGs in the county, but not all municipalities currently belong to a COG, in the sense of participating in the various COG cooperative arrangements currently in place. The COGs currently include 85% of the municipalities, excluding the City of Pittsburgh. In order that the COGs yield a consolidation scheme, we assigned each of the remaining municipalities to the geographically nearest COG).

Comparison of municipal composition and populations of the districts in the three schemes

In each of the consolidation scenarios discussed below, the term "municipalities" refers to the present 128 municipalities and their geographic boundaries while the term "district" refers to combinations of these municipalities under any of the three consolidation schemes. Under the Optimal Scheme, 65 districts are formed from the 128 municipalities. Under the School District Scheme, 45 districts are formed and under the COG Scheme eight are formed. Maps of the districts resulting from these schemes are shown in Figs 1–3, respectively.

Under the Optimal and the School District Schemes, some districts consist of single municipalities, while under the COG Scheme no municipalities other than Pittsburgh and Neville Island, remain independent. The shaded districts in Fig. 1 are the ones that have been formed by consolidating smaller municipalities, while the districts that are not shaded consist of single municipalities. In the Optimal Scheme, 34 out of the 65 districts consist of single municipalities, while in the School District Scheme 11 out of the 45 consist of single municipalities.

It is of interest to compare the consolidation pattern from the Optimal Scheme and the School District Scheme in terms of the municipalities that remain independent or combine with others. Of 11 municipalities that are independent school districts, eight remain independent in the Optimal Scheme. It is notable that two of the consolidated school districts correspond exactly to two of the consolidated districts in the Optimal Scheme: Moon Area school district to S10, and Gateway school district to E8. Also, a number of optimal districts correspond closely to school districts. For example, E7 corresponds to the South Allegheny school district except for Borough of Port Vue, and E10 corresponds to the Plum school district without the Borough of Oakmont. Also, in 10 of the 31 consolidated districts (32%) in the Optimal Scheme, all the constituent municipalities belong to the same school district.

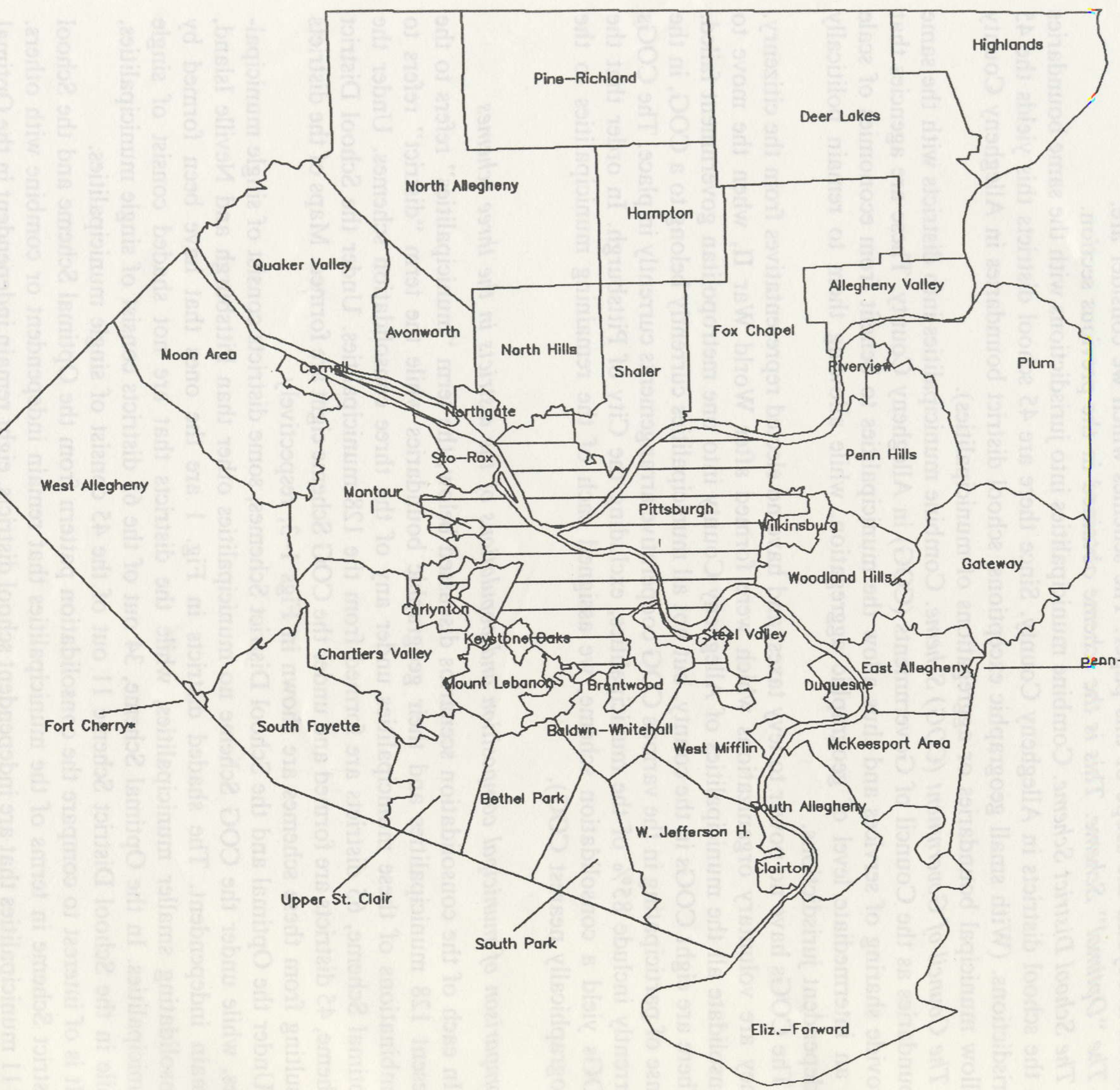

Fig. 2. The School District Scheme.

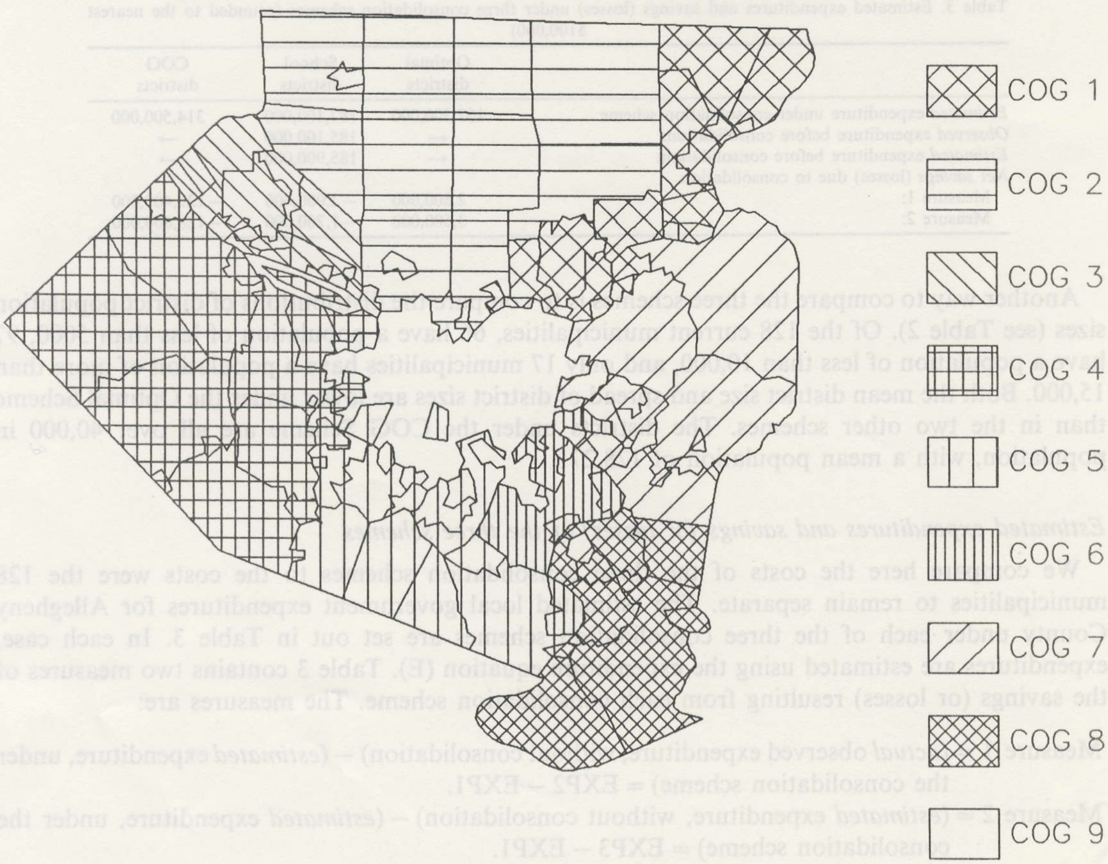

Fig. 3. The Council of Government Scheme.

We next compare the Optimal Scheme to the COG Scheme in terms of the membership of the various municipalities in the new districts. As mentioned earlier, all municipalities in Allegheny County are assigned to a Council of Government. Though participation in a COG is currently voluntary, 78% of the municipalities (100) in the county do actually participate in the COG in their area in the sense of financially contributing to the COGs or attending COG meetings. It is reasonable to view participation in a COG as an indicator of the willingness to cooperate with other participants in order to benefit from the economics of scale in providing such services as water, sewage, fire, etc. Under the Optimal Scheme, about two-thirds (45 out of 65) of the consolidated districts consist of municipalities that belong to, and participate in, the same COG.

Table 2. Size Distribution of districts under the three consolidation schemes

Population	Optimal districts	School districts	COG districts	Current municipalities
⩽5000	12	3	0	66
5001–10,000	12	3	0	31
10,001–15,000	8	12	0	14
15,001–20,000	16	6	0	3
20,001–25,000	6	5	0	4
25,001–30,000	2	5	0	3
30,001–35,000	4	3	0	5
35,001–40,000	4	2	0	1
>40,000	1	6	8	1
Total No. of districts	65	45	8	128
Mean population	15,767	23,218	128,271	8069
Standard deviation	11,687	14,218	60,406	9298
25th percentile	7862	12,782	57,863	2150
Median population	15,206	19,385	95,982	4754
75th percentile	22,266	32,339	203,826	9707

Table 3. Estimated expenditures and savings (losses) under three consolidation schemes (rounded to the nearest $100,000)

	Optimal districts	School districts	COG districts
Estimated expenditure under consolidation scheme:	182,300,000	187,100,000	314,500,000
Observed expenditure before consolidation:	←	185,100,000	→
Estimated expenditure before consolidation:	←	185,900,000	→
Net savings (losses) due to consolidation			
Measure 1:	2,800,000	−2,000,000	−129,400,000
Measure 2:	3,600,000	−1,200,000	−128,600,000

Another way to compare the three schemes is to compare the distributions of district population sizes (see Table 2). Of the 128 current municipalities, 66 have a population of less than 5000, 97 have a population of less than 10,000, and only 17 municipalities have a population of more than 15,000. Both the mean district size and spread of district sizes are lower under the Optimal Scheme than in the two other schemes. The districts under the COG Scheme are all over 40,000 in population, with a mean population of 128,271.

Estimated expenditures and savings (or losses) of the three schemes

We compare here the costs of the three consolidation schemes to the costs were the 128 municipalities to remain separate. The estimated local government expenditures for Allegheny County under each of the three consolidation schemes are set out in Table 3. In each case, expenditures are estimated using the expenditure equation (E). Table 3 contains two measures of the savings (or losses) resulting from each consolidation scheme. The measures are:

Measure 1 = (*actual* observed expenditure, without consolidation) − (*estimated* expenditure, under the consolidation scheme) = EXP2 − EXP1.

Measure 2 = (*estimated* expenditure, without consolidation) − (*estimated* expenditure, under the consolidation scheme) = EXP3 − EXP1.

Table 3 shows that only the Optimal Scheme yields an estimated net savings: $2.8 and $3.6 million, respectively. The School District Scheme yields a small estimated net expenditure loss of $2.0 or $1.2 million, respectively, while the COG Scheme yields a very large estimated net expenditure loss of $129.4 and $128.6 million, respectively. Thus, on the basis of estimated net cost savings, the Optimal Scheme is superior to the other two schemes, although it is not dramatically superior to the School District Scheme.

Comparison of the three schemes by socio-economic criteria

In a previous section we identified four socio-economic criteria considered to be particularly important to the economic and political feasibility of a consolidation scheme. In generating the optimal consolidation scheme, we used these four criteria to eliminate potential candidate districts. In the School and COG schemes these four socio-economic constraints were not imposed, since we simply followed existing school district and COG boundaries. To evaluate the three consolidation schemes, it is important to compare the extent to which they satisfy these four test criteria. [Recall that, even in the Optimal Scheme, a few districts failed one or more

Table 4. Socio-economic tests failed by three consolidation schemes

	Optimal districts	School districts	COG districts	Current municipalities
No. of districts in scheme	65	45	8	128
No. of districts failing Test 1 (income)	0	5	4	NA
No. of districts failing Test 2 (tax base)	3	1	0	12
No. of districts failing Test 3 (expenditure)	0	7	7	NA
No. of districts failure Test 4 (% elderly)	3	8	0	32
No. of districts failing at least one test	5	18	7	40
These (failing) districts comprise:				
(a) municipalities	5	69	112	40
(b) population	11,114	369,839	796,637	220,543
(c) % of population	1.08	36.09	77.73	21.52

Table 5. Optimal scheme: small municipalities in the northern sector

Municipality	Population
Sewickley Heights	899
Sewickley Hills	419
Haysville	117
Osborne	529

socio-economic tests. This is because, for a few small municipalities (see below), no district satisfied all socio-economic tests; hence, they were included as separate single municipality districts.]

The results are set out in Table 4. It can be seen that, as expected, the Optimal Scheme satisfies the four socio-economic tests best. Only the five very small districts fail any of the tests, and these five districts comprise only 1.1% of the county population. In the case of the School District Scheme, 18 districts (comprising 69 municipalities and 36.1% of the county population) fail one or more of the four tests. In the case of the COG Schemes, seven of the eight COGs (comprising 112 of the 128 municipalities and 77.7% of the county population) fail one or more of the four tests. The relatively poor performance of the COG Scheme in the socio-economic tests may explain why so many municipalities choose not to participate in the COG in which they are geographically located.

It is interesting to note that, even without any consolidation, the existing set of 128 municipalities satisfy Tests 2 and 4 better than do either the School District or COG Schemes (see the final column of Table 4). We are unable to compare the other two tests in this way, since they apply only to districts and not to municipalities.

Returning to the Optimal Scheme, we note that a number of small municipalities remained independent. In the northern sector, four municipalities with populations less than 5000 remain independent. These are shown in Table 5.

However, dropping the third socio-economic constraints (Test 3) would greatly decrease the number of such small independent municipalities. If Test 3 were dropped (see Table 5) then Sewickley Heights would combine with Sewickley Hills and Haysville would combine with Osborne. Further, both of these groupings would then combine with either Sewickley or Allepo to form a single district out of five municipalities. This involves withdrawing Sewickley or Allepo from a neighboring consolidation, but it would still be "optimal" to do since it would reduce total costs and would not violate any of the other three socio-economic criteria.

CONCLUDING REMARKS

In this paper we presented a framework for developing and examining municipal consolidation schemes. We developed a model to determine the "optimal" (or least cost) size of a district and a rationale for U-shaped marginal functions for costs of local government. Further, we set out a model to determine an optimal consolidation plan subject to socio-economic constraints, and compared the results of this consolidation scheme to two plausible alternatives.

The socio-economic constraints reflect the fact that the consolidation plan would have to be accepted, and voted on by each of the municipalities to be consolidated. Experience elsewhere, and the literature on local public goods, indicates that electors are much less likely to support a consolidation plan if it may substantially change (for the worse) their existing local government tax levels or services. The socio-economic constraints reflect this, by ruling out consolidation schemes that could substantially change these characteristics.

The methodology presented in this paper could be used to analyze another widespread regional problem, the optimal size of school districts in the face of declining enrollment. We applied our models to Allegheny County, PA, which has the third largest number of independent jurisdictions of all counties in the U.S. Our results include the following. First, we found that our Optimal Scheme: (a) gave much lower costs; and (b) satisfied socio-economic feasibility constraints much better than the other two possible schemes (the School District and COG Schemes). Second, as well as performing better on (a) and (b), the Optimal Scheme has the attraction that the districts formed turn out to have strong parallels with the districts formed in the other two schemes. This further attests to the large extent to which the model captures political acceptability. We have estimated the cost of satisfying socio-economic constraints. The need to satisfy such constraints decreases the potential cost savings, and lessens considerably the financial attractiveness of consolidation. Hence, it is no surprise that various attempts at consolidation, especially to go to metropolitan or county-wide schemes, have failed in many parts of the U.S.

Acknowledgements—The authors wish to thank Marcus Berliant for his comments on an earlier version of this paper, and the Center for Public Financial Management, Carnegie Mellon University for financial and computational support. Responsibility for the opinions and any errors rests with the authors.

REFERENCES

1. B. Bunch and R. P. Strauss. The fiscal position of municipalities in the steel valley council of governments. Center for Public Financial Management, Carnegie Mellon University, Pittsburgh, PA (1987).
2. B. Bunch and R. P. Strauss. Municipal consolidation: an analysis of the financial benefits for small fiscally distressed municipalities. *Urban Affs. Q.* **27**, 615–629 (1992).
3. T. J. Chan and C. A. Yano. A Multiplier adjustment approach for the set partitioning problem. *Opers Res.* **40**, S40–S47 (1992).
4. G. Cassidy, D. Epple and T. Romer. Redistribution by local government in a monocentric urban area. *Reg. Sci. Urb. Econ.* (1989).
5. B. C. Ellickson. Jurisdictional fragmentation and residential choice. *Am. Econ. Rev. Papers Proc.* **61**, 334–339 (1971).
6. D. Epple, R. Filimon and T. Romer. Equilibrium among local jurisdictions: towards an integrated treatment of voting and residential choice. *J. Publ. Econ.* **24**, 282–308 (1984).
7. D. Epple and T. Romer. On the flexibility of municipal boundaries. *J. Urban Econ.* **26**, 307–319 (1989).
8. R. S. Garfinkel and G. L. Nemhauser. Optimal political redistricting by implicit enumeration techniques. *Mgmt Sci.* **16**, B495–B508 (1970).
9. B. W. Hamilton. Property taxes and the Tiebout hypothesis: some empirical evidence. In *Fiscal Zoning and Land Use Controls* (Edited by E. S. Mills and W. E. Oates). Heath, Lexington, MA (1975).
10. S. Hess, J. Weaver, H. Siegfeld, J. Whelan and P. Zitlau. Nonpartisan political redistricting by compute. *Opers Res.* **13**, 993–1006 (1965).
11. W. Z. Hirsch. *The Economics of State and Local Governments*. McGraw-Hill, New York (1970).
12. C. A. Holloway, D. A. Wehrung, M. P. Zeitlin and R. T. Nelson. An interactive procedure for the school boundary problem with declining enrollment. *Opers Res.* **23**, (1975).
13. J. Johnston. *Econometric Methods*, Second Edn. McGraw-Hill, New York (1972).
14. J. Miller. *Annexations and Boundary Changes in the 1980's and 1990–91*. U.S. Bureau of the Census. Washington, DC (1993).
15. L. A. Neidell and R. D. Teach. Allocation of parks and recreation facilities. *Interfaces* **10**, (1980).
16. D. A. Plane. Redistricting reformulated: a maximum interaction/minimum separation objective. *Socio-Econ. Plann. Sci.* **16**, 241–244 (1982).
17. I. M. L. Robertson. The delimitation of local government electoral areas in Scotland: a semi-automated approach. *J. Operl Res. Soc.* **33**, 517–525 (1982).
18. S. Rose-Ackerman. Market models of local government: exit, voting and the land market. *J. Urb. Econ.* **6**, 319–337 (1979).
19. W. J. Scanlon and R. P. Strauss. The geographic heterogeneity of public expenditure functions. *Rev. Econ. Statist.* **14**, 191–194 (1972).
20. L. Schrage. *LINDO User's Manual*. Scientific Press (1978).
21. R. P. Strauss. The impact of block grants on local expenditures and property tax rates. *J. Publ. Econ.* **3**, 269–284 (1974).
22. R. P. Strauss and J. D. Hughes. A new approach to the demand for public goods. *J. Publ. Econ.* **6**, 191–204 (1976).
23. H. Theil. Specification error and the estimation of economic relationships. *Rev. Int. Statist. Inst.* **25**, 41–51 (1957).
24. H. Theil. *Principles of Econometrics*. Wiley, New York (1971).
25. W. J. Thomas. Techniques for defining geographic boundaries for health regions. *Socio-Econ. Plann. Sci.* **13**, 321–326 (1979).
26. U.S. Bureau of the Census. *1980 Census of Governmental Organizations*. Washington, DC (1982).
27. F. Westhoff. Existence of equilibrium in economics with a local public good. *J. Econ. Theory* **14**, 84–112 (1977).
28. O. P. Williams, H. Herman, C. S. Liebman and T. R. Dye. *Suburban Differences and Metropolitan Politics*. University of Pennsylvania Press (1965).
29. R. C. Wood. *Suburbia*. Houghton Miflin Company (1958).

Appendix opposite

APPENDIX

I. Number of Local Governments in Allegheny County, PA

Year	Total governments	County	Cities	Townships	Special districts	School districts
1942	255	1	73	53	10	118
1962	358	1	84	42	123	108
1982	323	1	86	42	150	45

Source: U.S. Census Bureau, Census of Governments, Governmental Organization, 1942, 1962, 1982.

II. Statistics on Socio-economic Criteria

	Mean	SD	Min	Max
Median family income (U.S.$)	23,172.14	7329	11,130	62,585
Tax base *per capita* (U.S.$)	4451	2500	1,594.92	17,682
Mean expenditure *per capita* (U.S.$)	195.39	102	52.39	813
Proportion of Population greater than 65 yr old (%)	13.84	4.5	1.88	25