ARTICLES

Combining Sample and Census Data in Small Area Estimates: Iterative Proportional Fitting with Standard Software*

Ludi Simpson and Mark Tranmer
Cathie Marsh Centre for Census and Survey Research, University of Manchester

The combination of detailed sample data with less detailed but fully enumerated marginal subtotals is the focus of a wide range of research. In this article we advocate careful modeling of sample data, followed by Iterative Proportional Fitting (IPF). The modeling aims to estimate accurately the interaction or odds ratios of complex tables, which is information not contained in the marginal subtotals. IPF ensures consistency with the subtotals. We advance this work in three practical ways. First, we show that detailed small-area estimates of both counts and proportional distributions usually gain accuracy by combining data for larger areas containing the small areas, and we illustrate the multilevel framework to achieve these estimates. Second, we find that a general classification or socioeconomic typology of the small areas is even more associated with the within-area interactions than is membership of the larger area. Third, we show how the Statistical Package for the Social Sciences (SPSS) can be used for IPF in any number of dimensions and with any structure of constraining marginal subtotals. Throughout, we use an example taken from the 1991 U.K. Census. These data allow us to evaluate various methods combining 100 percent tabulations and the Samples of Anonymised Records. **Key Words: Small areas, iterative proportional fitting, multilevel models**

Introduction

The common problem we address is how best to estimate a detailed relationship represented by a multidimensional table when relatively unreliable information for the cells of that table is supplemented by subtotals that are more reliable. This situation often arises in demographic and geographical studies and has often been tackled using Iterative Proportional Fitting (IPF). For example, when updating commuting analyses, detailed matrices of flows between origins and destinations will be approximate for years after their collection; they can be made consistent with the changing pattern of total commuters or migrants in each area that is more closely monitored (Bruton 1985). Similarly, migration data from the U.S. Census is insufficiently detailed for multidimensional demographic models, but the detailed data can be estimated from combining alternative sources with the subtotals that are published from the census (Rogers, Willekens, and Raymer 2003). The sex and age structure for small populations may be made more certain by making initial estimates consistent with more reliable data, both for the total population of the smaller areas and for the age composition of larger areas (Myers 1992; Rees 1994; Simpson 1998). Sample survey data insufficient for the accurate estimation of local relationships may be supplemented with ancillary data from a census or administrative sources that give accurate subtotals (Skinner 1991; Wong 1992; Johnston and Pattie 1993; Longford 1999).

We have previously discussed the extension of census data released for subnational areas by using the additional detailed relationships available only from sample census data for the same areas (Tranmer et al. 2005). That article explains and evaluates several statistical approaches, including multilevel modeling in order to incorporate relevant data from areas other than that being estimated. Here we extend the problem to the use of sample data for areas larger than the small areas of interest. We are interested in see-

 The work for this paper and its writing were supported respectively by the U.K. Economic and Social Research Council awards R000223703, "Combining aggregate and micro-data to extend census tables for local areas" and RR000271214, "Local demography and race." Paul Norman provided helpful comments on a draft.

The Professional Geographer, 57(2) 2005, pages 222–234
Initial submission, July 2003; final acceptance, December 2003.
Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, and 9600 Garsington Road, Oxford OX4 2DQ, U.K.

ing how estimates for small areas behave when scaled to known marginal subtotals. We wish to answer the following questions: Do approaches found to give the best results for large areas also give the best results for smaller areas? How good are the estimates from each approach? Moreover, this article goes into considerable detail to explain how the common multidimensional scaling procedure Iterative Proportional Fitting (IPF) may be applied using appropriate SPSS commands. We explain how IPF may be applied across any number of dimensions. Hence, this article describes a novel approach that a researcher may use to carry out IPF analysis using standard statistical software. In addition, we derive some useful empirical results relevant to small-area estimation using this procedure, allowing us to recommend the best approach to use.

We also review the derivation of small-area estimates of multidimensional tables from sample data, including the advantages of a multilevel modeling framework for such estimates. We use IPF to constrain those initial estimates to less-detailed aggregate data, the marginal subtotal. Using data for England and Wales from the 1991 Census of Great Britain described later in the article, our results show that the marginal constraints usually improve the accuracy of the estimates by ensuring consistency with information reliably known about each area. The level of accuracy is dependent on an appropriate estimation of the interaction, which is not informed by the constraining marginal data.

When the sample data are sparse within the area to be estimated, or not identified at all for that area, good estimates can nonetheless be achieved when the area is associated with a set of similar areas for which the sample is large enough to be accurate. We estimate electoral ward areas of population 1,000 to 15,000 using marginal subtotals for each ward and detailed relationships drawn from a national sample. The sample has geographical identifiers only for large areas containing the ward but also classifies the type of ward. We will show in detail how the IPF procedure can be flexibly implemented in SPSS, thus overcoming the limits of previous literature on the procedure that handles two-dimensional tables in special software. IPF for multiway tables for 10,000 areas is achieved in less than ten seconds in SPSS, and this aspect is likely to be of value in a great variety of applications.

Methods

We seek the best method of estimating spatial heterogeneity; a relationship varies between local relationships areas, but we do not have full information on that relationship. In his book *A Solution to the Ecological Fallacy*, King (1997, pp. 12 ff, section 1.2; tables 1.1, 1.2) uses an example based on votes by each ethnic group to illustrate the ecological analysis problem he wishes to solve. Using two tables, he first describes the aggregate data he has at district level and then at precinct level. In each of these two example tables, King puts a question mark in each cell to indicate he has no information on the cell counts; the data he has are only the aggregate margins of the table. This differs from our situation, where we have some information about the cell counts from a sample, as well as aggregate table margins from the 100 percent census counts. This situation, involving aggregate data plus a sample of individual data for the same area that may be linked at some geographical area level, is the one we consider here, involving the use of IPF. King uses other techniques, based on the method of bounds rather than IPF, to develop his solution, and his work also assumes aggregate data only, as opposed to our situation where aggregate data based on all people is combined with a sample of individual data.

We begin with the main example used in this paper, the estimation of a two-dimensional table of car ownership by tenure, for which examples are given in Table 1A. This cross-classification is available for all areas from the released Census data, which allows us to examine the accuracy of a variety of strategies for estimating the table were it not available. This will be helpful for choosing appropriate strategies for estimating tables that are truly not available from census output.

In both the national and local data, car ownership varies between tenure categories: the lack of a car is generally more evident for those in socially rented accommodation (rented from the local authority or housing association). Thus, an assumption of independence of tenure and car ownership will give poor estimates of the true relationship. However, there is spatial heterogeneity. The true relationship varies between areas: the lack of car ownership for socially rented

Table 1 *Car Ownership and Tenure, National and Two Districts (A) True Values*

England and Wales		**Privately rented**	**Socially rented**	**All tenures**	**Odds ratio: Socially rented/owner-occupiers**
No car	4,456,806	1,367,440	6,046,785	11,871,031	
	12.8%	35.4%	59.0%	24.2%	
Car	30,389,314	2,499,284	4,206,901	37,095,499	
All people	34,846,120	3,866,724	10,253,686	48,966,530	9.8
Bradford	**Owner-occupiers**	**Privately rented**	**Socially rented**	**All tenures**	
No car	75,386	16,839	56,304	148,529	
	21.9%	52.4%	75.7%	33.0%	
Car	268,524	15,313	18,085	301,922	
All people	343,910	32,152	74,389	450,451	11.1
Isle of Wight	**Owner-occupiers**	**Privately rented**	**Socially rented**	**All tenures**	
No car	14,256	3,775	6,822	24,853	
	14.7%	38.3%	54.2%	20.8%	
Car	82,895	6,093	5,760	94,748	
All people	97,151	9,868	12,582	119,601	6.9

(B) National Interaction Constrained with IPF to District Margins

Bradford	**Owner-occupiers**	**Privately rented**	**Socially rented**	**All tenures**	**Odds ratio: Socially rented/owner-occupiers**
No car	76,934	16,658	54,937	148,529	
	22.4%	51.8%	73.9%	33.0%	
Car	266,976	15,494	19,452	301,922	
All people	343,910	32,152	74,389	450,451	9.8
Isle of Wight	**Owner-occupiers**	**Privately rented**	**Socially rented**	**All tenures**	
No car	13,456	3,700	7,697	24,853	
	13.9%	37.5%	61.2%	20.8%	
Car	83,695	6,168	4,885	94,748	
All people	97,151	9,868	12,582	119,601	9.8

Note. IPF = Iterative Proportional Fitting.

households is more striking for Bradford than for the Isle of Wight. The odds ratios reflect this difference, being higher for Bradford. The odds ratios of Table 1 are calculated from the numbers in the cells as follows: the ratio of non-car owners to car owners for those in socially rented accommodation, divided by the same ratio for owner-occupiers. For example, using the data in Table 1 for England and Wales, we can derive the odds ratio for car ownership for owner-occupiers, compared with social renters as:

$$(30,389,314/4,456,806)/$$
$$(4,205,901/6,046,785) = 9.8$$

In other words, owner-occupiers are 9.8 times more likely to own a car than social renters.

Were the internal cells for the subnational areas not known, the national odds ratios for car ownership could be used in their place, but they would be inaccurate. Constraining the national table of car ownership by tenure to the marginal subtotals for car ownership and for tenure for each local area improves the estimates of car ownership for each area. This is achieved in Table 1B by IPF, known elsewhere as *raking*, *rim-weighting*, or *structure-preserving estimation*. The IPF procedure is simple in concept, scaling the initial estimates alternately to the tenure category subtotals and the car-ownership subtotals until the scaled values no longer change significantly. Thus, the marginal subtotals of Table 1B are the same as those of Table 1A, but the odds ratios are those of the national table. The mathematics of IPF is specified in Appendix 1 and illustrated by explicitly deriving Table 1B. David Wong (1992) experiments with IPF to estimate the odds ratios (interactions) of varying magnitude. He shows that the resulting estimates are more reliable to the

extent that the sample data accurately estimates the true odds ratio. Ron Johnston and Charles Pattie (1993) show the close relationship between IPF and other statistical approaches.

In the example of Table 1, all data are derived from the full census outputs. More generally, the problem is to find an accurate estimate of the local relationship before constraining to the local marginal subtotals. Information for the relationship will often be available from sample data that give information not only for the nation but for areas within it. However, the sample within the area of interest may be so small that sampling error outweighs the advantages of the direct estimate from that local sample. Tranmer et al. (2005) evaluated a number of different strategies for estimating the relationship between car ownership and tenure for 253 areas within England and Wales, combining, in various ways, sample individual data and tabular aggregates published from the 1991 census. We found that direct estimates from the sample for each local area were much more accurate than assuming the national estimate applied in every local area, but that best results were gained from a multilevel model. The multilevel estimates are based on a multilevel modeling approach, which has been well documented in the recent statistical literature (see, e.g., Goldstein 1995; Snijders and Bosker 1999). The application of this approach to small-area estimation is described in detail in Tranmer et al. (2005). Multilevel estimation is used simply to weight the local direct sample estimate with the national estimate; the local estimate is weighted proportionally to the sample it is based on such that more reliable local estimates are used with little change while less reliable local estimates are "shrunk" toward the national estimate.

As our measure of overall accuracy of the estimates from each approach, we used the root mean square error (RMSE). This is derived by taking the difference between the actual and estimated frequencies for each table cell in each area, squaring the difference, and summing over all cells and all areas. Squaring each difference gives greater weight to large errors. This sum is then divided by (number of cells * number of areas) to give the mean square error. The square root is then taken so that the measure can be interpreted as an average difference in the true and estimated frequency count. Thus, the RMSE will give a good general measure of the accuracy of each approach, taking into account all of the table cells and areas. The most accurate approach will lead to the smallest mean square error; if the approach perfectly predicted the true values, the RMSE would be zero. In other words, there would be no difference between the true and estimated frequency counts (or proportions based on these counts). The same measure is used when assessing the accuracy of the proportion of car owners in each tenure category.

$$RMSE = \sqrt{\left(\sum (estimate - actual)^2\right) / n}$$

For large areas (Tranmer et al. 2005), the RMSE of the proportion of car owners was reduced from 0.109 for the national proportion applied to all localities to 0.022 for the direct estimates from local samples alone and, again, to 0.020 for the multilevel model. In every model, constraint to the local known marginal subtotals (using IPF as above) further improved the estimates, reducing the average error to 0.016 for the constrained multilevel model.

In this article, we estimate the same relationship between car ownership and tenure for all 9,363 wards of England and Wales. The Census sample individual data do not identify the ward from which each record was taken. We compare two strategies for estimating the ward relationship between car ownership and tenure before applying IPF to make that estimated relationship consistent with marginal subtotals known for each ward from 100 percent tabulations from the census. First, we take the estimate of that relationship in a larger area containing the ward, as if it applied to all its wards. Second, we estimate from the sample individual data the relationship for each *type* of ward, irrespective of its location in England and Wales.

We can summarize our research design as follows:

(a) We estimate the relationship between two variables, car ownership and tenure. We derive the proportion of car ownership in each tenure category from four sources:

   1. Assumption of no relationship; independence between car ownership and tenure;
   2. National data for England and Wales;
   3. 253 large subnational areas, estimated directly from sample data;

4. Groups of 9,363 small subnational areas, estimated directly from sample data.

(b) We extend the third and fourth cases by using multilevel modeling to weight the direct sample estimate and the national estimate, relying more on the national estimate when the sample for the direct estimate is small.
(c) The tenure-specific proportions of car ownership from each of the four sources and the two multilevel extensions are applied to the tenure totals in each area known from tabulated census data. The counts and proportions of car ownership within each tenure category are computed both for the 253 large subnational areas and for the 9,363 small areas. The exception is that the fourth source of proportions, groups of small subnational areas, cannot be applied to the large subnational areas.
(d) We then use IPF to scale further the estimated counts to agree also with the car ownership totals in each area and derive new proportions from the new counts.

We evaluate and compare each set of estimates using the RMSE of the estimates across all small areas.

Data

All data are extracted from the 1991 Census, referring to England and Wales. The hierarchy of areas is as follows:

- 9,363 Electoral wards, average population 5,000.
- 253 SAR areas of minimum population 120,000, usually a single local authority district or an amalgam of neighboring districts. These are the areas identified on the individual Sample of Anonymised Records.
- 10 regions. For example, Yorkshire and Humberside, Inner London. These are identified on the household Sample of Anonymised Records.

The 1991 Census and its Samples of Anonymised Records (SARs) is described in detail by Angela Dale and Cathie Marsh (1993).

The marginal totals and the true cross-tabulated values are extracted from the tabulation of car ownership and tenure in table L20 of the 1991 Census Local Base Statistics. Car ownership indicates whether or not a household has access to a car. Those renting from a local authority or a housing association are classified as socially rented, while those renting privately include properties rented with a job or business; the remaining properties are owner occupied.

The sample data are also taken from the 1991 Census. The 2 percent individual SAR identifies the SAR area from which each record has been extracted. The 1 percent household SAR, not overlapping with the 2 percent sample, has no local geographical code but identifies the region and the type of ward from which each record has been extracted.

The classification of wards has been made by the Office for National Statistics (ONS), through multidimensional scaling and cluster analysis of thirty-seven socioeconomic, demographic, household composition, and housing and employment characteristics (Wallace and Denham 1996). The 1 percent SAR uses the ONS classification of wards into fourteen groups, labeled, for example, "Suburbia (Group 1)" or "Deprived industrial areas (Group 14)." Within each government region, groups with few wards have been amalgamated on the 1 percent SAR to ensure confidentiality, resulting in twenty-seven categories (Dale and Openshaw 1997). One of these categories refers only to Scotland; thus, there are twenty-six categories of the ward classification in England and Wales.

Results

The strong relationship between car ownership and tenure has already been noted. An assumption of independence between the two (the same car ownership proportion in each tenure category) produces large errors averaging over 0.2 in the proportion of car owners in the 9,363 local wards (Table 2). IPF does not change these estimates. A better strategy is to apply the national (England and Wales) relationship equally to every area. This reduces the average error considerably, but it remains above 0.15. Table 2 shows the average error for both the number in the six cells of the tenure by car ownership table and for the proportion of car owners in the three tenure categories. Each will be of interest for different applications, but, in fact, the pattern of errors is similar in the two cases. Comparative results for 253 SAR areas from Tranmer et al. (2005) are included in Table 2.

Table 2 *Average Root Mean Square Error of Estimates of the Local Relationship Between Car Ownership and Tenure*

Nature of area: Constrained to known marginal subtotals:	Average error of the six numbers in tenure x car-ownership cells				Average error of the three proportions of car owners			
	253 SAR areas		9,363 Wards		253 SAR areas		9,363 Wards	
	Tenure	Tenure and car	Tenure	Tenure and car	Tenure	Tenure and car	Tenure	Tenure and car
Source of relationship:								
1. Independent margins	13,490	13,490	381	381	0.213	0.213	0.209	0.209
2. England and Wales	5,873	1,324	258	69	0.109	0.036	0.158	0.067
2% SAR, 253 SAR areas								
3. Direct SAR area sample	738	529	184	62	0.022	0.019	0.110	0.066
3a. Multilevel model	686	467	184	61	0.020	0.016	0.109	0.065
1% SAR, 26 ward types								
4. Direct Ward type sample			146	57			0.093	0.063
4a. Multilevel model			146	56			0.093	0.063

Note. SAR = Sample of Anonymised Records.

The average error is shown when constraining the results to only the tenure marginal subtotals and after using IPF to constrain to both car and tenure marginal subtotals—known for each area from the 100 percent Census tabulations. It is clear that, in every case, IPF considerably improves the accuracy of the results. When applying the national relationship of car ownership and tenure to all areas, the additional use of IPF reduces the average error in numbers in each cell from 258 to 69 and reduces the average error in the proportion of car owners by more than half, from 0.158 to 0.067.

There are further gains when using the SAR area relationship to estimate all the wards within the SAR area. Before IPF, the average error is reduced from 0.158 to 0.110, and slightly further to 0.109 when the best multilevel SAR area estimates are used. However, after incorporating the ward marginal subtotals, the gain over using the national relationship is very little, 0.065 compared to 0.067. The marginal subtotals capture much of the variation of car ownership and tenure between wards, but the odds ratios still vary within a SAR area almost as much as they do within England and Wales.

Therefore, the lower panel of Table 2 shows the second approach. The sample information on car ownership and tenure is grouped not within SAR areas but across all wards of the same type, using the ONS classification of wards available on the 1 percent SAR. The classification has just twenty-six categories compared to the 253 SAR areas. Estimates based upon it provide a similar error to that for estimates that use each SAR area's estimated odds ratio for all wards within it. The average error is slightly reduced, to 0.063. The multilevel model provides no extra power as the sample in each ward classification is usually very large.

The importance of local area information is abundantly clear. The 253 SAR areas contain heterogeneous wards within them, which are equally well estimated by associating them with similar wards than with their containing SAR area. This is the case even when the ward classification takes only twenty-six categories and the estimates of these categories are derived from a 1 percent sample rather than the 2 percent sample used for the SAR area models.

The errors reported in Table 2 are averages over all wards in England and Wales. As we saw in Table 1, the odds ratios themselves vary between SAR areas—in that case, between Bradford and the Isle of Wight. Figures 1 and 2 explore the variation in odds ratios between all SAR areas, between all wards, and, in the latter case, between wards of different ONS classification. The odds ratio shown is the contrast between social renters and owner-occupiers, when considering the odds of non-car owners to car owners. The true, recorded, figures from the Census 100 percent tabulations are used for these odds ratios, which vary much more between wards (standard deviation 10.0) than between districts (standard deviation 2.1). The variation in the odds ratio between ONS classification types is considerable, explaining why its use was able to reduce the error in estimation.

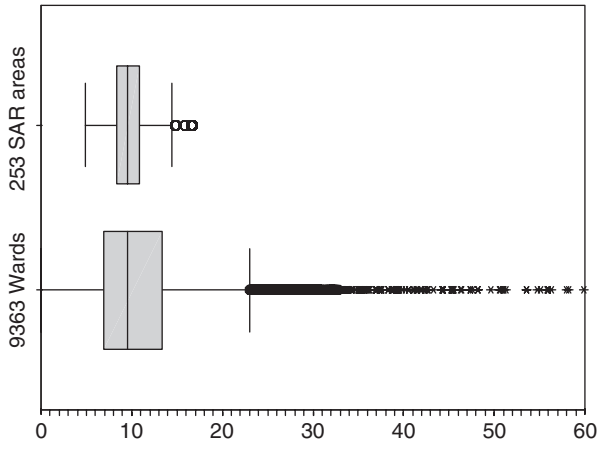

Figure 1 *Odds ratios for car ownership, contracting social renters with owner-occupiers: SAR areas and wards. SAR = Sample of Anonymised Records.*

It should be noted that the estimates for wards are not as good as those for SAR areas given in Table 2. The average error in the best estimates is 0.063, compared to the 0.016 achieved for SAR areas. This is caused by the lack of sample information about each specific ward. However, it is doubtful that a ward identifier on the SARs would improve these estimates, as the sample size in each ward would be extremely small for all but the largest wards.

The errors reported in Table 2 are averages over all wards in Britain. Not all wards will show the same gain from each method, and IPF is not guaranteed to improve the estimates in every

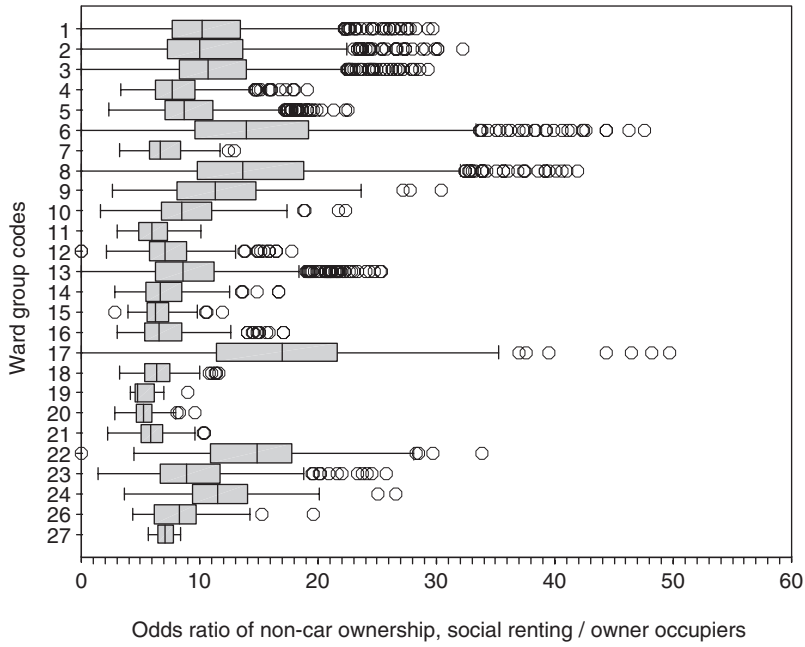

Figure 2 *Odds ratios for car ownership, contrasting social renters with owner-occupiers: ward groups. Note. Ward group code 25 is missing as it refers to wards in Scotland, not included in this study.*

case. Proportions are unlikely to be estimated well when the marginal subtotals are very small, and cell counts are more likely to have larger errors when the marginal subtotals are themselves relatively large.

IPF Implemented in SPSS

Introduction

IPF, although it was recommended as early as 1940 (Deming and Stephan), has been included in standard statistical texts for many decades (Bishop, Fienberg, and Holland 1975), and is straightforward to compute arithmetically (see Appendix 1), is nonetheless not available as a versatile explicit routine to handle differently sized tables. Paul Norman (1999) has made available an Excel routine that implements IPF for any two-dimensional array, but, generally, researchers, as a one-off task, have written their own software suited to their particular research needs.

However, within SPSS (1999), there are two procedures that achieve just the estimation that is required and that are not limited to two dimensions. HILOGLINEAR and GENLOG are loglinear modeling procedures that approximate tabular data. The logarithm of each cell count predicted by the model is a linear function of parameters that are estimated from the data. For the current article, the importance is twofold. First, in both procedures, the predicted values for a model are consistent with the observed data for the marginal subtotals specified in the model. In our context, we provide observed data that are consistent with the marginal subtotals that we wish to fix from local information. Second, a cell weight subcommand allows the weight of each cell to be taken from a different variable; the motivation of this subcommand in the documentation is the use of zeros as cell weights in order to impose structural zeros on the model. The only mention of adjusting tables to fit margins in SPSS documentation is the comment on HILOGLINEAR to the effect that "You can also use CWEIGHT to adjust tables to fit new margins" (SPSS 1999, 536). Default initial values of 1 are used in the estimation routine (SPSS 2003), and SPSS replaces these by the value of the cell weight variable before starting the routine. We use the CWEIGHT command and its equivalent CSTRUCTURE in GENLOG to supply the estimated cell values that we wish to be adjusted to the marginal subtotals.

These two SPSS loglinear procedures apply to tables of any dimensions, termed *factors* in SPSS documentation, and any number of categories within each factor. The table of tenure by car ownership for each of the 9,363 wards of England and Wales can therefore be considered as a three-way table of dimensions $3 \times 2 \times 9{,}363$. Our initial estimates fill this entire table (though many of its cells are equal when each ward from a SAR area or a ward type is given the same initial estimates). The marginal subtotals known for each ward then comprise two two-way tables of size $3 \times 9{,}363$ (tenure by ward) and $2 \times 9{,}363$ (car ownership by ward). This approach is more efficient than the equivalent estimation of 9,363 separate 3×2 tables, each with marginal 1×3 and 1×2 marginal subtotal tables.

Preparation of Data

Data should be arranged with a single record for each cell of the whole table to be estimated. The variables should include each of the factors that represent a dimension of the table, a variable containing the initial values, and a variable to represent the marginal subtotals (see Figure 3, where the three factors are ward, car, and tenure). The initial estimates may be in the form of proportions of one of the marginal factors

Ward	car	tenure	initial	margins
1	1	1	.30	183.95
1	2	1	.70	478.05
1	1	2	.54	50.30
1	2	2	.46	130.70
1	1	3	.71	79.75
1	2	3	.29	207.25
2	1	1	.30	286.31
2	2	1	.70	509.69
2	1	2	.54	87.05
2	2	2	.46	154.95
2	1	3	.71	140.64
2	2	3	.29	250.36
3	1	1	.30	85.90
3	2	1	.70	87.10
3	1	2	.54	276.56
3	2	2	.46	280.44
3	1	3	.71	135.55
3	2	3	.29	137.45

Figure 3 *Example SPSS data file prepared for Iterative Proportional Fitting (IPF).*

rather than counts; IPF maintains only the odds ratios from the initial estimates, scaling them to the marginal subtotal counts. In the example in Figure 3 the initial estimates are the proportions of car owners and others within each tenure category. They have been estimated for the larger area, which contains each of the three wards shown, and are therefore the same for each ward. All variables may be named for convenience by the user.

The "margins" variable must contain any set of values that sum to the marginal subtotals. One means of achieving this variable is to compute the cells of the table under the assumption that the margins are independent of each other. Many spatial data sets including U.K. census data are held as a single record for each area, as in Figure 4, making such a calculation straightforward. In general, for the cell referring to category i of factor 1, category j of factor 2, and so on, to category k of the final factor N, multiply the overall Total by the proportion of the Total in each category. Here Margin1_i stands for the marginal total in category i, and so on:

$$\begin{aligned}\text{Margins}_{ij\ldots k} = &\text{Total} * (\text{Margin1}_i/\text{Total}) \\ &* (\text{Margin2}_j/\text{Total}) \\ &* \ldots * (\text{MarginN}_k/\text{Total})\end{aligned}$$

For example, the variable margins for the first record of Figure 3—car ownership for owner-occupiers in ward 1—was computed from the data in Figure 4 as follows:

$$1130*(314/1130)*(662/1130) = 183.95$$

In this way, six extra variables are created for each record of Figure 4, each representing the variable margins for one of the six car vs. tenure categories.

The task of manipulating the six variables derived from Figure 4 into the single variable "margins" for Figure 3 was achieved in this example by saving each variable with its ward, tenure, and car identifiers into six separate temporary files; the SPSS command ADD FILES was used to concatenate the six files, and MATCH FILES was used to add these data to the estimates as in Figure 3.

Implementation and Comparison of GENLOG and HILOGLINEAR

The records as in Figure 3 must be weighted by the variable representing marginal subtotals. These weights are interpreted by SPSS as observed counts in the table. Estimation of a loglinear model uses these observed data to compute expected values that add to the marginal subtotals. Specification of a model defines which marginal subtotals are to be taken from the observed data, while expected values for the remaining cells of the table are estimated consistent with these marginal subtotals. The default estimation assumes no interaction (equivalently, independence) for any terms not in the model. The use of further cell weights within the loglinear procedure, equal to the initial estimates, ensures that the expected values instead retain the odds ratios represented by these initial estimates for terms not specified in the model. The SPSS terminology can be confusing in this application. The records are weighted by the margins variable, while the cell weights within the statistical command GENLOG or HILOGLINEAR take the initial values.

In GENLOG, the initial values are given in the /CSTRUCTURE subcommand, margins are specified in the /DESIGN subcommand, and the expected values are saved with variable name as specified under the /SAVE = PRED subcommand. The SPSS syntax for the file above would be as follows:

```
WEIGHT BY margins.
GENLOG tenure car ward
/CSTRUCTURE = initial /DESIGN ward
* tenure ward * car
/PRINT = NONE /PLOT = NONE
/CRITERIA = CIN(95) ITERATE(20)
            CONVERGE(.001) DELTA(.5)
/SAVE = PRED (afteripf ).
WEIGHT OFF.
```

Ward	Car=1	Car=2	Tenure=1	Tenure=2	Tenure=3	total
1	314	816	662	181	287	1130
2	514	915	796	242	391	1429
3	498	505	173	557	273	1003

Figure 4 *Example census subtotals from which margins in Figure 3 are derived.*

The same results are achieved with procedure HILOGLINEAR as follows:

```
WEIGHT BY margins.
HILOGLINEAR ward (1, 3) nocar (0,1) tenure (1,3)
/CWEIGHT = initial
/PRINT = FREQ /PLOT = NONE
/CRITERIA = ITERATE(20)
            CONVERGE(.001) DELTA(.5)
/DESIGN ward * tenure ward * car.
WEIGHT OFF.
```

While the output values are the same, there are considerable differences between the two procedures, which tend to favor GENLOG. In particular, GENLOG allows expected values (termed *predicted values*) to be saved, while HILOGLINEAR only prints expected values to the text output file in a format that is not easily exported for further use. Additionally, GENLOG accepts variables with string values and numeric values with nonconsecutive values, while these must be converted to consecutive integer values by RECODE or AUTORECODE commands before using HILOGLINEAR.

However, because GENLOG is comprehensive (allowing nonhierarchical models and estimating parameters for nonsaturated models, aspects irrelevant to this article), it uses the iterative Newton-Raphson algorithm for maximum likelihood estimation. This produces the same results as IPF but is much less efficient since it requires more random memory and takes longer to reach a solution. Use of the command SPLIT FILE to repeat analyses for each local area separately and the command SET MXMEMORY to increase the memory accessed by SPSS enabled GENLOG to run the analyses in this article very quickly (9,363 tenure by ward estimations for eight separate initial estimates, in under one minute). To prevent unnecessary output for each local area, results can be suppressed. With this strategy, HILOGLINEAR should rarely be necessary.

In some cases, the marginal subtotals themselves may contain a zero count; this cannot be handled by GENLOG, though HILOGLINEAR, correctly, sets all cells within this marginal category to zero. Before running GENLOG, such cells should be set to a small count, say 0.1, and set back to zero after running GENLOG. Finally, the occurrence of zero initial estimates of proportions or counts is treated as structural zeroes. This may be unreasonable in the applied context if there are large marginal subtotals. It is best dealt with by adding a small value, say 0.001 to these cells, which maintains their small value relative to other cells but allows non-zero estimates. The syntax incorporating these various enhancements to implement IPF for many areas is as follows:

```
COMPUTE margins1 = margins.
IF (margins = 0) margins1 = 0.1.
COMPUTE initial1 = initial.
IF (initial = 0) initial1 = 0.001.
SET MXMEMORY = 56000.
SORT CASES BY ward tenure car.
SPLIT FILE BY ward.
WEIGHT BY margins1.
SET RESULTS OFF.
GENLOG tenure car
/CSTRUCTURE =
   initial1 /DESIGN tenure car
/PRINT = NONE /PLOT = NONE
/CRITERIA =
   CIN(95) ITERATE(20)
   CONVERGE(.001) DELTA(.0)
/SAVE = PRED (afteripf ).
SET RESULTS ON.
WEIGHT OFF.
SPLIT FILE OFF.
```

In addition to output options, both GENLOG and HILOGLINEAR allow adjustments to default criteria for the estimation. The maximum number of iterations, the convergence criterion (how close the final results must be to the previous iteration), and the delta value added to observed counts before estimation may all be changed.

Discussion

IPF can now be implemented generally without the need for special programs. The advantages of a multilevel modeling framework for achieving reliable initial estimates have been clarified.

Complex tables with small samples in each cell will not be well estimated by the procedures used here. However, the multilevel approach allows further modeling to account for variation between areas in order to improve small area estimates before the use of IPF. The use of the

ONS classification has shown the potential for such modeling. However, at the time of writing, the U.K. Samples of Anonymised Records from the 2001 Census are unlikely to contain either a subregional identifier or a ward classification. Estimation of the type illustrated here may have to take place within the census offices or in supervised census data laboratories. As a separate development, integration of IPF within the estimation procedure would usefully provide estimates of reliability of the estimates after IPF, which are not provided here.

For census data, the specific tabulations required may be commissioned from the census database. These exact results are preferable to the estimation suggested in this article, but may be slow, expensive, or otherwise impractical to achieve. In the U.K., commissioned tables are not realistically available for censuses prior to 2001.

Thus, both for the census and a wide variety of other applications, the need will remain to combine less reliable detailed information with more reliable aggregate subtotals. The strategies advocated in this paper are expected to be of general and practical use.

Appendix 1. Iterative Proportional Fitting (IPF) Defined and Exemplified

We first define the general IPF algorithm following Bishop et al. (1975, section 3.5.2), and then apply it in the two-dimensional case illustrated by Table 1 of the paper.

Table A1 *Illustration of Iterative Proportional Fitting to Derive Table 1B from Table 1A*

Step 0: Initial values (England and Wales) and marginal subtotals (Bradford) from Table 1(a)

	Owner-occupiers	Privately rented	Socially rented	All tenures	Marginal sub-totals (Bradford)
No car	4,456,806	1,367,440	6,046,785	11,871,031	148,529
Car	30,389,314	2,499,284	4,206,901	37,095,499	301,922
All people	34,846,120	3,866,724	10,253,686	48,966,530	
Marginal sub-totals (Bradford)	343,910	32,152	74,389		450,451

Step 1: Scale to tenure margin

Ratio of required margin to margin at previous step: 0.009869 | 0.008315 | 0.007255

	Owner-occupiers	Privately rented	Socially rented	All tenures
No car	43,986	11,370	43,869	99,225
Car	299,924	20,782	30,520	351,226
All people	343,910	32,152	74,389	

Step 2: Scale to car margin

Ratio of required margin to margin at previous step		Owner-occupiers	Privately rented	Socially rented	All tenures
1.496893	No car	65,842	17,020	65,667	148,529
0.859623	Car	257,822	17,864	26,236	301,922
	All people	323,664	34,885	91,903	
	Maximum cell deviation from previous cycle (Step 0)				30,131,492

Step 3: Scale to tenure margin

Ratio of required margin to margin at previous step: 1.062553 | 0.921668 | 0.809433

	Owner-occupiers	Privately rented	Socially rented	All tenures
No car	69,961	15,687	53,153	138,801
Car	273,949	16,465	21,236	311,650
All people	343,910	32,152	74,389	

Step 4: Scale to car margin

Ratio of required margin to margin at previous step		Owner-occupiers	Privately rented	Socially rented	All tenures
1.070089	No car	74,864	16,786	56,878	148,529
0.968784	Car	265,397	15,951	20,573	301,922
	All people	340,262	32,738	77,452	
	Maximum cell deviation from previous cycle (Step 2)				9022.1654

Table A1. *(Contd.)*

Step 18: Scale to car margin					
Ratio of required margin to margin at previous step		**Owner-occupiers**	**Privately rented**	**Socially rented**	**All tenures**
1.000000	No car	76,934	16,658	54,937	148,529
1.000000	Car	266,976	15,494	19,452	301,922
	All people	343,910	32,152	74,389	
	Maximum cell deviation from previous cycle (Step 16)				0.0500

Take a multidimensional table with initial values $m_\theta^{(0)}$ of cell θ and a set of restrictions $q = 1, \ldots, s$ on marginal subtotals of the table, with value $X_{\theta q}$ corresponding to cell θ. The marginal subtotal computed from the initial values is $m_{\theta q}^{(0)}$.

Scale the initial values to the first marginal subtotals to derive an estimate at step 1:

$$m_\theta^{(1)} = m_\theta^{(0)} * (X_{\theta 1} / m_{\theta 1}^{(0)}).$$

Repeat the scaling to each marginal subtotal $2, \ldots, s$ to complete one cycle of s steps:

$$m_\theta^{(s)} = m_\theta^{(s-1)} * (X_{\theta s} / m_{\theta s-1}^{(s-1)}).$$

In general at the tth step where $t-q$ is a multiple of s, we have:

$$m_\theta^{(t)} = m_\theta^{(t-1)} * (X_{\theta t-1} / m_{\theta q}^{(t-1)}).$$

At the end of the rth cycle, $t = rs$, compute the maximum convergence from the previous cycle:

$$\delta^r = \max_\theta (\mathrm{abs}(m_\theta^{(rs)} - m_\theta^{(rs-s)})).$$

Accept estimate $m_\theta^{(rs)}$ when δ^r is smaller than some fixed convergence criterion, for example 0.001. Since the procedure is proven to converge when the marginal subtotals are consistent with each other (e.g., add to the same overall total), the choice of convergence criterion only affects the number of cycles that will be needed before the criterion is met.

For an example, Table A1 takes the national pattern of tenure and car ownership from Table 1A, and makes it consistent with the marginal subtotals for Bradford, also taken from Table 1A, as was done to achieve the upper half of 1B. Here there are two steps to each cycle (with two sets of marginal subtotals, $s = 2$). At each step, the table cells are multiplied by the ratio of one of the required margins from the Bradford data to that margin from the previous step. At each step the cell values are brought progressively more into line with the required margins, as evidenced by the reduction in maximum deviation after each cycle. The first two cycles are shown in full in Table A1. The eighteenth step is also shown, by which time the deviation from the previous cycle is less than 0.1 and the cell values are those shown in Table 1B. Thus, nine iterations were required for this table to converge before one could be sure that integer values of the estimates would not change further.

Literature Cited

Bishop, Y. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

Bruton, M. J. 1985. *Introduction to transportation planning*. London: Hutchinson.

Dale, A., and C. Marsh, eds. 1993. *The 1991 Census user's guide*. London: HMSO.

Dale, A., and S. Openshaw. 1997. *Adding area-based classifications to the samples of anonymised records (SAR) from the 1991 Census*. Manchester, England: Centre for Census and Survey Research.

Goldstein, H. 1995. *Multilevel statistical models*. New York: Halsted Press.

Johnston, R., and C. Pattie. 1993. Entropy maximising and the IPF procedure. *The Professional Geographer* 45:317–22.

King, G. 1997. A solution to the ecological inference problem. Princeton, NJ: Princeton University Press.

Longford, N. T. 1999. Multivariate shrinkage estimation of small area means and proportions. *Journal of the Royal Statistical Society A* 162(2): 227–45.

Myers, D. 1992. *Analysis with local census data*. San Diego, CA: Academic Press.

Norman, P. 1999. *Putting iterative proportional fitting on the researcher's desk*. Leeds, U.K.: University of Leeds, Geography Department.

Rees, P. 1994. Estimating and projecting the populations of urban communities. *Environment and Planning Series A* 26:1671–97.

Rogers, A., F. Willekens, and J. Raymer. 2003. Imposing age and spatial structures on inadequate migration-flow datasets. *The Professional Geographer* 55:56–69.

Simpson, S. 1998. Apportionment using counts of patients and electors. In *Making local population*

statistics: a guide for practitioners: Estimating with confidence, ed. S. Simpson, 61–68. Middlesbrough, U.K.: LARIA.

Skinner, C. 1991. The use of synthetic estimation techniques to produce small area estimates. *OPCS New Methodology Series NM18*. London: Office for Population Censuses and Surveys.

Snijders, T., and R. Bosker. 1999. *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. Beverly Hills, CA: Sage.

SPSS. 1999. *SPSS 10.0 syntax reference guide: Base, regression models, advanced models*. Chicago: SPSS.

———. 2003. *Hiloglinear*. http://www.spss.com/tech/stat/Algorithm/11.5/hiloglinear.pdf

Tranmer, M., L. Simpson, D. Steel, A. Pickles, and G. Davies. 2005. *Extending Census tables for local areas by combining aggregate and individual level data*. Working Paper, Centre for Census and Survey Research, University of Manchester.

Wallace, M., and C. Denham. 1996. *The ONS classification of wards*, SMPS 60. London: HMSO.

Wong, D. 1992. The reliability of using the Iterative Proportional Fitting procedure. *The Professional Geographer* 44:340–48.

LUDI SIMPSON is Reader in Social Statistics in the Cathie Marsh Centre for Census and Survey Research at the University of Manchester, M13 9PL, United Kingdom. E-mail: ludi.simpson@manchester.ac.uk. His research interests include local demography, race and census practice.

MARK TRANMER is a lecturer in statistics in the Cathie Marsh Centre for Census and Survey Research at the University of Manchester, M13 9PL, United Kingdom. E-mail: mark.tranmer@manchester.ac.uk. His research interests include multilevel modelling, data integration, and area effects.