[Regression Models for Adjusting the 1980 Census]: Comment

Gad Nathan


Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28198602%291%3A1%3C34%3A%5BFAT1%3E2.0.CO%3B2-T

*Statistical Science* is currently published by Institute of Mathematical Statistics.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ims.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.
that very good statistical work is necessary if it is to carry the day in court, especially under peer review, as practiced in courts.

Second, the nature of the critique is exceedingly instructive, especially if taken as a guide to self-critique. We see plainly the wisdom of stating the assumptions that justify a proposed analysis else someone may state them for us. Then it will be helpful to ponder those assumptions and their implications ("In effect, New York was taking the position that bias in the census—the undercount—was well related to the three explanatory variables, but bias in PEP was not.") Perhaps we can find, as they did, empirical checks on some of the assumptions. Consideration of several more or less equally plausible alternative models may help us to gauge the fragility of conclusions that we draw from some one of them. We can hope that assessment of error by bootstrap rather than by theoretical formula will become standard practice where experience does not already point to the successful applicability of theoretical error formulas.

These are practical guides to action as self-critic. In my estimation they are the important message in the careful and clearly stated critique given by Freedman and Navidi.

Comment

Gad Nathan

For many years the statistical and legal controversy about the necessity and the advisability of adjusting census counts on the basis of information available from evaluation surveys or from external sources has centered on the general principles involved, such as the definition of the concept of "statistical defensibility" (Spencer (1982) and the discussion thereof). While discussion of these general principles is important, and even necessary, Freedman and Navidi, together with Ericksen and Kadane (1984), are to be congratulated on getting down to the brass tacks of the problems involved in attempts at the real life application of adjustment methods as well as with the theoretical and empirical criticism of the methods proposed. Discussion of actual applications of adjustment methods is important since both proponents and opponents of the adjustment of census counts are generally in agreement that adjustment should be carried out if and only if there exists a method for carrying it out which meets certain conditions and standards of quality and accuracy, for example, by some definition of "statistical defensibility." Since obviously no formal existence or nonexistence theorems can be proved in this respect, the argument must hinge on the empirical results of the proposed methods of adjustment.

Even when an adjustment method can be demonstrated to be adequate, it will not generally be unique and different adjustment methods may be required or may be suitable for the different purposes for which census data are used. This raises the issue of who should do the adjusting, the producer of the data or their users. In any case, a fundamental decision to be made with respect to the adjustment procedure involves the definition of the unit of analysis and of prediction, and in particular, the geographical breakdown to be used. Although both Ericksen and Kadane (1984) and the present paper consider different alternative definitions of the geographical areas to be used, they both implicitly assume that the geographical breakdown must be such that reasonably adequate sample estimates of the undercount be available for each area considered.

In fact, the Lindley and Smith (1972) strategy used here has been generalized by Pfeffermann and Nathan (1981) to deal with the important case where observations on the dependent variable are not available for all units, whereas observations on the independent variables are. They also propose methods for estimating the variance $\sigma^2$, in this situation, which are similar to those discussed here. Using these results, alternative geographical breakdowns, not necessarily limited to those with sample data for all areas and possibly more suited to the required uses of census data, could be considered.

Most of the criticism of the New York proposals for the adjustment justifiably centers on the underlying assumptions of the regression model and on their justification or lack of it. Indeed, the correct identification of a working model is crucial in this situation where the pure design-based estimate for a single area has too large a sampling error to be of any practical use on its own and must "borrow strength" from other areas via the model-based approach. However, the search for a good model need not be limited to the aim of adjustment via a model-based estimator. A model which can be used not only to estimate more efficiently
REGRESSION MODELS

the undercount but also to explain it and to discover
the sources of underenumeration can become an im-
portant tool for census planners in their attempts to
reduce the undercount in future censuses, by attaining
a deeper understanding about the underlying mecha-
nism of underenumeration.

As pointed out in the paper by Freedman and Na-
vidi, the proponents of the New York adjustment
procedure failed to provide sufficient justification for
the model used. This failure was both with respect to
the inclusion of variables and the resulting potential
bias and with respect to the specification of the
variance and of the error structure. A long list of
additional potential variables is recommended for con-
sideration. This list includes “geographical location”
and interactions, so that the possibility of different
regression models for geographical regions, not only
with different constants but also with different regres-
sion coefficients, must be considered. If we add to this
the various possibilities for error structure (model
errors, sampling errors, and correlations between
them), the number of different models to be considered
and the number of their parameters becomes very
large indeed. The choice of the correct model among
these and the estimation of its parameters all on the
basis of 66 observations becomes a formidable prob-
lem. To this are added the problems due to the fact
that the observations are based on data from a com-
plex sample design, rather than on simple random
sampling, so that, for instance, the diagonality of the
sampling variance matrix, $K$, is indeed difficult to
justify.

However, in fact, the 66 estimates of undercounts
are each based on many observations (the Post Enu-
meration Program sample size in each area) and this
individual information for subunits might be utilized
for more efficient model search and identification. For
instance, some method of sample re-use or cross-
validation based on sample-splitting as proposed by
Pfeffermann and Nathan (1985) could be used. It is
shown there that efficient cross-validation can over-
come both the problem of overfitting and underesti-
mation of error due to the search among a large
number of alternatives and the problem of testing
goodness of fit on the basis of data from complex
samples.

The empirical results and simulation study of Sec-
tion 6 illustrate clearly the faults of the proposed
adjustment. However, it should be pointed out that
the fact that replacement of the crime rate variable
by an urbanization rate results in approximately the
same quality of fit (as measured under the model
assumptions) does not in itself invalidate either model
for purposes of adjustment. Similarly, the lack of
consistency in the choice of the best subset of three
explanatory variables in the simulation study does not
necessarily show inadequate adjustment. It is possible
that more than a single choice of a set of explanatory
variables can provide equally adequate estimates of
undercount, although, of course, the explanation pro-
vided by the models is thereby limited. In any case, as
pointed out, the estimates of standard errors used to
judge the quality of these models are definitely defi-
cient.

Finally, although the results of this paper show,
without doubt, that the adjustment procedure pro-
posed by New York is not “statistically defensible,”
this should, under no circumstances, be regarded as a
demonstration that an adequate adjustment procedure
cannot be found. The negative result should rather be
interpreted as implying that an adequate procedure
for adjustment of census counts has not yet been
found, either for a specific aim or for an official, all
purpose one. However, the methods proposed by
Ericksen and Kadane (1985) are certainly worthy of
further consideration and, above all, for further em-
pirical testing. In particular, suitable methods for
model choice and model identification for these cir-
cumstances should be developed and applied. The
results obtained should be continually scrutinized and
appraised by methods similar to those of the present
document.

ADDITIONAL REFERENCES

PFEFFERMANN, D. and NATHAN, G. (1981). Regression analysis of
data from a cluster sample. J. Amer. Statist. Assoc. 76
681–689.
identification based on data from complex sample surveys. Bull.
Int. Statist. Inst. 51 12.2.1–12.2.17.
Statistician 36 208–209.

Rejoinder

D. A. Freedman and W. C. Navidi

To begin with, we would like to thank Morrie
DeGroot for his editorial support and the discussants
for their careful work. We wish Jay Kadane weren’t
quite so angry with us, but then we are being very
negative about some of his work. He and Gene Erick-

sen are good statisticians who believe in what they do;