adjustment. The paper by Freedman and Navidi provides valuable early discussion on this important topic and contributes importantly to the continuing debate about census coverage error and the wisdom of census adjustment. Most statisticians should find their discussions informative, amusing, and provocative. I certainly did.

Comment

Albert Madansky

My comments on the Freedman–Navidi paper are of two sorts, one directed specifically to the content of the paper and the other a set of general remarks directed at the common theme of Freedman's recent papers (Freedman, 1981, 1985, Freedman et al., 1983, as well as this one), critiqueing the use of statistics in modeling.

1. COMMENTS ON THE FREEDMAN–NAVIDI PAPER

They describe the Post Enumeration Program (PEP) studies (Section 3) and point out that "about two dozen different series of PEP estimates were developed," each based on a different set of imputation rules for treating the missing data. They claim that the Bureau of the Census "was unwilling to use PEP to adjust the population counts" because 1) there was considerable variation across the series, 2) the probabilistic basis for the estimate was open to serious question, and 3) the standard errors of the estimate turned out to be quite large.

To impute these as the reasons for the "unwillingness" of the Bureau of the Census to use PEP to adjust the population counts lends a greater aura of ratiocinativity to that decision than actually was the case. In truth, the Bureau of the Census was unwilling *by any means* to adjust the population count, and never considered in a constructive way how one might use PEP to adjust the population counts. The Bureau of Census stance was more in the nature of "we don't want to adjust the raw census counts" and "even if we wanted to adjust, we don't know *how* to adjust using PEP data" than in the nature of the authors' imputed scenario, namely an implied willingness to adjust, recognition that methodology was available for effecting that adjustment, but, taking the view that "the PEP data are so problematical that we don't want to use them to adjust the raw census," and rationally deciding not to embark on an adjustment program. Indeed, Mitroff et al. (1982, 1983; see also Kadane, 1984) indicate that the principal motivating factor for the Bureau of the Census decision not to adjust was that the Bureau has historically been "nonpolitical and objective" and that use of any adjustment procedure would be in violation of that standard of Bureau behavior.

The positive contribution made by Ericksen and Kadane was to set forth an approach by which the PEP data could be used to adjust the census. Their paper merely suggested an approach toward adjustment; the work they did to implement their approach was in the nature of a constructive proof of an "existence theorem," used in an advocacy proceeding partially for the numbers it produced but primarily to make the point that indeed adjustment was feasible with the data at hand.

But let us get to the substance of the Freedman–Navidi paper. What should one make of the three "warts" in the PEP data? That the standard error of the PEP estimates turned out to be high is no reason not to use them if, in combination with the raw census data, one can produce demonstrably better estimates of the population counts than those achievable by using merely the raw census data. Let us see by a quick calculation whether this is in fact potentially the case.

The essence of the procedure for estimating the population count using the results of a postcensus sample (e.g., PEP) can be seen from a consideration of the following:

	Census	Sample
Respondents	n	n'
Nonrespondents	m	m'
Total	N	N'

Here N is the true census count, n is the observed census count, N' is the postcensus sample size, n' is the number in the postcensus sample who were also in the census, and $m' = N' - n'$ is the number in the postcensus sample who were not counted in the census. Now let $\theta = m/N$, the fraction undercount in the

Albert Madansky is Professor of Business Administration and Associate Dean for Ph.D. Studies, Graduate School of Business, University of Chicago, 1101 E. 58th Street, Chicago, IL 60637.

census. Since $n + N\theta = N$, $N = n/(1 - \theta)$ and, since $1 - \theta$ is estimated by n'/N', we can estimate the population count by nN'/n'.

Now since n' is bounded away from 0, we can use the normal approximation to the binomial distribution of n' to enable us to calculate the expected value and standard deviation of $1/n'$. For example, if $N' =$ 50,000 and $\theta = .002$ (selected by me not coincidentally but rather as the PEP 2/9 estimate of the undercount), the estimate of the expected value of $1/n'$ is .005 and the standard deviation of $1/n'$ is .0003532. Thus, the ratio of the standard deviation of the estimated population count to its expected value is .10. (If θ were as high as .07, this ratio would equal .016.) By not adjusting, one incurs the cost of saying that the population count is n when it is in fact N. By using the crude adjustment method just described one would have an unbiased estimate of the true population count and would incur the cost associated with the standard deviation of the adjusted estimate. And, as can be seen from these calculations, especially for highly undercounted groups, this standard deviation is relatively quite small.

That there are two dozen different imputation rules, the use of which produced considerable variation in the resulting estimates of the undercount, is again no reason to use the PEP data in the same manner as an ostrich would, burying his head in the sand to make the problem disappear. Rather, one should consider which of the imputation rules is most reasonable and go with the associated data. If one is to dither about the problems of using faulty data and wants to show that the results are not robust to small variations in the data, a better tack to have been taken by Freedman and Navidi would have been to compare use of the 2/9 data set with use of the PEP data set designated second-best by Ericksen, namely 3/9, than their exercise based on the 10/8 data set, selected, by their own admission, "cavalierly." Nowhere in the paper is there any citation of the justification given by Ericksen and Kadane for their choice of the 2/9 data set. Almost by innuendo the reader is left with the impression that its choice by Ericksen and Kadane was also made cavalierly. In point of fact, Ericksen gives a complete justification for use of the 2/9 data set over all others (cf. pp. 68–71 of Ericksen's affidavit in *Cuomo v. Baldrige, SDNY*), and in particular for setting aside data sets based on the August Current Population Survey (CPS) (such as 10/8).

It was no surprise to me that, as Freedman and Navidi demonstrated, if the 10/8 data set were used in the Ericksen–Kadane methodology instead of the 2/9 data set then different population estimates would have ensued ... especially given the different estimates of the undercount based on the two data sets (see Cowan and Bettin, 1982):

	10/8	2/9	3/9
Total	0.2%	1.4%	1.3%
Black	2.7	6.7	6.3
Hispanic	3.6	5.6	5.3
Other	−0.4	0.3	0.2

Note though that I included in this table the net undercount rates for Ericksen's second best data set, 3/9, and a quick comparison of these number with those for the selected data set, 2/9, would indicate that if Freedman and Navidi had "cavalierly" selected that data set they would not have been able to provide the reader of their paper with such a startlingly contrasting set of results.

With respect to the third concern about the data voiced by Freedman and Navidi, that the probabilistic basis for the estimate is open to question, again the concern is subject to analysis and not merely cause to drop the use of the data. In the case of the P sample, clearly if the conditional probability of not being in the P sample given not being in the census is 1 then there would have been no "recaptures" who were not previously "captures." Since there were "recaptures" who were not previously "captures," we know that we are not dealing with an extreme deviation from independence. The position implied by Freedman and Navidi, namely, "When in doubt, don't!" makes all degrees of dependence equally damaging. By contrast, a constructive approach would attack the technical issue of determining the magnitude at which a deviation from independence practically "matters." (For example, an investigation of the 2×2 table χ^2 test for test resistance, along the lines of Ylvisaker (1977), would be useful in determining whether the observed degree of dependence should be considered an important deterrent to the use of the P sample.)

Finally, I quote from Kadane (1984): "... the end product is an estimate, that all estimates require assumptions that in turn require justification, and that all estimates start on an even footing, including nonadjustment." I believe that Freedman and Navidi would have done their scientific cause more justice by analyzing as well the nonadjusted 1980 Census with the care, rigor, and vigor exhibited in this paper.

2. COMMENTS ON FREEDMAN'S GENERAL QUEST

There is a passage in the Jerusalem Talmud stating that religious scholars are the guardians and defenders of the city. The Aramaic for "guardians of the city" is "neturei karta," and it has been adopted as the name of a group of ultrareligious Jewish extremists, mainly in Jerusalem, whose view is that only God can reestablish a Jewish state in Israel, and that a Jewish state established by human beings is a violation of God's will and so should be combatted. They see their mission as that of "guardians of the city," defending it from encroachment by secularity. As I read the ever-growing collection of papers authored or coauthored by David Freedman on the use of statistical procedures in modeling, I cannot help but dub him the "neturei karta," the "guardian of the city" of statistics.

How can one object to what he is trying to do? His quest, after all, seems quite reasonable. He tilts with models that are used in public policy deliberations and decisions. And he only concerns himself with the issue of whether the assumptions underlying the model are credible. Someone has to be the "guardian of the city!" Freedman is without peer in both thoroughness and clarity of analysis.

The problem, though, with Freedman's quest is in many ways analogous to that of the neturei karta. If they are successful, then the State of Israel will cease to exist. And if Freedman successfully uncovers models based on invalid assumptions, the decision maker is left to make decisions using only his intuition, for decisions must be made, with or without statistical help. All Freedman has done is saved statisticians from "aiding and abetting" and/or being accessories to a decision which in any event will be made, even if based merely on intuition and judgment. Is that worse or better than the scenario in which the statistician at least shows the decision maker the direction in which a decision should go, given the available data, in a (possibly) fictitious world built upon a bed of (possibly erroneous) assumptions? My contention is that even such deductions are useful grist for the decision maker's mill. Indeed, even if the assumptions are valid but the model is incomplete, or is just plain wrong, insights can be obtained from working the model through to its implied conclusions. (One can even gain insight from implications of purely mathematical models with no statistical component.)

Yes, assumptions should be checked for validity, and procedures should be checked for robustness. And no, statisticians are not merely people who "draw a straight line from an unwarranted assumption to a foregone conclusion using a procedure optimal according to a criterion invented by the statistician." But perhaps a bit of the latter can be condoned in statistical practice, especially if the alternative is that of letting the policy decision maker "go it alone." The statistician, after all, has more than a science to offer. He has a developed skill to offer as well, namely an ability to get the "feel" of data even when the data do not conform to any textbook model or set of assumptions.

ADDITIONAL REFERENCES

COWAN, C. D. and BETTIN, P. J. (1982). *Estimates and Missing Data Problems in the Post Enumeration Program.* Statistical Methods Division, Bureau of the Census.

FREEDMAN, D. (1981). Some pitfalls in large econometric models: a case study. *J. Business* **54** 479–500.

FREEDMAN, D. (1985). Statistics and the scientific method. In *Cohort Analysis in Social Research* edited by W. M. Mason and S. E. Fienberg. New York, Springer-Verlag, pp. 343–366.

KADANE, J. B. (1984). Review of Mitroff, I. I., Mason, R. O. and Barabba, V. P. The 1980 census: policymaking amid turbulence. *J. Amer. Statist. Assoc.* **79** 467–469.

MITROFF, I. I., MASON, R. O. and BARABBA, V. P. (1982). Policy as argument. *Manag. Sci.* **28** 1391–1404.

MITROFF, I. I., MASON, R. O. and BARABBA, V. P. (1983). *The 1980 Census: Policymaking Amid Turbulence.* Lexington, MA, Lexington Books.

YLVISAKER, D. (1977). Test resistance. *J. Amer. Statist. Assoc.* **72** 551–556.

Comment

I. P. Fellegi

1. INTRODUCTION

I must state at the outset that I like the paper and would only have relatively unimportant technical "quibbles" to raise in *disagreement.* Instead, I will concentrate on some broader implications of the paper's findings. Another introductory comment is prompted by the paper's style, but applies to much of the written material on the topic of census adjustment. I would have preferred if the paper had more of a "sanitized" version of the authors' testimony, i.e., free of the debating style of courtrooms. The issues involved are both significant and complex, and it is all the more important that we should be able to debate our differences in a manner that makes it easier for our professional colleagues to understand our point of view, even if they disagree with it.

The paper clearly and, I believe conclusively, makes a case against a *specific* approach to adjustment. Yet its value goes well beyond its argument against a particular methodology. This is an important paper the careful reading of which imparts at the same time

I. P. Fellegi is Chief Statistician of Canada, Statistics Canada, R. H. Coats Building, 26-A, Ottawa, Ontario K1A 0T6, Canada.