

# The Relevance or Irrelevance of Weights for Confidentiality and Statistical Analyses

Stephen E. Fienberg\*

**Abstract.** Sample survey weights represent a standard part of the survey statistician’s repertoire, but they remain a mystery for many of those who work with survey data. They also pose a potential challenge for confidentiality protection. Much has been written about how to use weights in statistical analyses since the basic idea of weighted estimates—weighting units inversely proportional to their probability of selection—emerged from the classic paper by Horvitz and Thompson [12]. Very little has been written about the effect of releasing survey weights on the confidentiality of survey data. The two topics are inextricably intertwined. This paper discusses both, largely from the model-based perspective, and explains what is claimed and/or known about the issue of confidentiality protection. The paper also provides a prescription for ways to deal with weights both for analysis and for disclosure limitation.

**Keywords:** survey weights, confidentiality protection, weighted estimates

## 1 Introduction

There is a general consensus in the statistical literature that the use of sampling protects confidentiality, although not absolutely, cf. Fienberg and Slavkovic [10] and the papers cited therein. Roughly speaking, the probability that an individual record that is unique in the sample is also unique in the population from which the sample was drawn equals the sampling fraction,  $n/N$ , where  $N$  is the number of units in the population and  $n$  is the size of the sample, e.g., see Fienberg and Makov [9] and Skinner and Holmes [26]. Most of the literature that discusses the issue of sampling and confidentiality protection focuses on the use of simple random sampling. Yet when one looks at the actual practice associated with large-scale sample surveys it involves far more elaborate sampling schemes leading to unequal probabilities of selection for each population member and the use of survey weights.

While sample survey weights represent a standard part of the survey statistician’s repertoire, they remain a mystery for many of those who work with survey data. Much has been written about how to use weights in statistical analyses since the basic idea of weighted estimates—weighting units inversely proportional to their probability of selection—emerged from the classic paper by Horvitz and Thompson [12]. In survey practice, however, weights do not necessarily have the same meaning from one survey to the next, since they usually incorporate information beyond the probability of selection. Since survey weights may play a role in the analysis of survey data, and thus are expected as part of the release of survey data for secondary analysis, we need to ask the extent

---

\*Carnegie Mellon University, Pittsburgh, PA, <mailto:fienberg@stat.cmu.edu>

to which they raise issues regarding the confidentiality associated with released survey databases, and if so, how such issues should be addressed.

In this paper, we briefly review the concept of sample survey weights and their instantiation in modern large-scale sample surveys. Then we reconsider the seemingly “age-old” question of the role of survey weights in statistical analyses of data associated with sampling from finite populations. Most answers from a design-based perspective pay limited attention to the role of statistical models and how they are explored and used in statistical practice. We illustrate the problematic issue of how to use weights in complex statistical analyses through an example drawn from the National Long-Term Care Survey. Finally, we turn to issues of the extent to which information in survey weights might provide information that would allow an intruder to identify individual survey respondents with non-negligible probability.

In the ensuing sections, we attempt to discuss and answer the following series of questions:

- What are weights for and where do they come from?
- To weight or not to weight?
- Do weights pose issues for the analysis of released microdata?
- What hazards do weights pose for confidentiality protection?
- If there are confidentiality problems associated with survey weights, how can they be averted or mitigated?

We review what has been said about survey weights and confidentiality by others, and we present our own prescription for how to deal with weights in model-based analyses and the possible confidentiality problems they pose.

## 2 Some Background on Large Scale Sample Surveys

### 2.1 Survey Weights and Statistical Analysis

Sample surveys and their analyses lie at the core of official statistics activities in many countries. The vast majority of statistical tools used in this regard are based on the probabilities associated with the sampling design, i.e.,  $p(s)$  is the probability of selecting the sample  $s$  of  $n$  units from the population of  $N$  units. Let  $\pi_i$  be the probability of selection of the  $i$ th unit for  $i = 1, \dots, N$ . Then

$$\pi_i = \sum_s p(s).$$

For each unit in the population there is a measurement  $y_i$ , and the usual survey analysis treats the values  $(y_1, y_2, \dots, y_N)$  as fixed. The canonical survey analysis problem focuses

on estimating a population quantity such as the total

$$Y = \sum_{i=1}^N y_i$$

based on the observed sample. The classic Horvitz-Thomson estimator takes the form:

$$\hat{Y}^{HT} = \sum_{k=1}^N \frac{y_k}{\pi_k}. \quad (1)$$

This estimator is said to be design-unbiased, i.e., its expectation with respect to the design probabilities,  $\{p(s)\}$ , is the true population total. The weight associated with the  $k$ th realized sample unit is the inverse of the probability of selection,

$$w_k = 1/\pi_k, \quad (2)$$

and “informally” corresponds to the number of individuals in the population represented by the  $k$ th sample unit.

For simple random sampling the selection probabilities are equal, and design-based and model-based analyses tend to coincide, e.g., see Lohr [16]. When the sample design involves combinations of stratification and clustering, as is the case in most government surveys, then the selection probabilities are typically not equal, and design-based and model-based analyses tend to diverge; this is especially the case when the quantity of interest is not a simple quantity such as a population total or mean. In such circumstances, design-based survey statisticians often adopt an approach such as:

- Finding population quantities corresponding to the parameters of the model and then treating these as fixed but unknown, e.g., using the method of estimating equations, e.g., see Binder and Patak [2] and Kovačević and Binder [15].
- Estimating these population quantities using something like the Horvitz-Thomson weighted estimator. The Horvitz-Thompson estimator replicates sampled individuals by the weight in equation (2) to create a “mock” population.

The design-based estimation literature also argues that the use of weights provides robustness against model misspecification, e.g., see Hansen, et al. [11], Kalton [13], Pfefferman [19], and Pfefferman et al. [20]. But see Bertolet [1] for evidence to the contrary.

Statisticians who advocate making inference with respect to the model of interest and not with respect to the sampling design often note that:

- The sampling probabilities (or weights) are irrelevant to model-based analyses.
- Stratification adjustments go into the predictive structure of linear-like models.
- Clustering needs to be represented as an added component of variation.

Fienberg [6, 7] makes such arguments and others have developed models to clustering as a component of variation.

There are also proposals in the statistical literature to take weighting into account from a model-based perspective by replacing the likelihood function, which normally takes the form

$$L = \prod_{k=1}^n f(y_k|\theta)$$

where  $f$  is the probability density function and  $\theta$  is the parameter of interest, by a pseudo-likelihood function of the form

$$L = \prod_{k=1}^n f(y_k|\theta)^{w_k}. \quad (3)$$

Then we can compute the pseudo-likelihood estimate by maximizing

$$\log(L) = \sum_{k=1}^n w_k \log f(y_k|\theta)$$

with respect to  $\theta$ . Unfortunately, there is no general justification for this approach, and it makes little or no sense from a Bayesian perspective.

In typical official statistics settings, survey weights are also used to adjust for factors other than unequal probabilities of selection, and they are the product of at least three components:

$$w_k = \frac{1}{\pi_k} \times (\text{adjustment for non-response}) \times (\text{postratification adjustment}). \quad (4)$$

The second factor in equation (4) inflates the sample results to adjust for non-response, typically by invoking the assumption that the missing data are *missing at random*, at least within selected strata or post-strata, whereas the third component re-weights the population totals to add up to *control totals* coming from another source such as a census. Some survey organizations implement additional adjustments as well. Thus the typical weights not only reflect the sample design but also model-based adjustments to make the realized sample “match up” with information from other sources. Then for analysis purposes, the design-based perspective treats the weights as if they were fixed.

Using these more complex “sampling weights” in the pseudo-likelihood of equation (3) has even less justification than using the original base weights of equation (2).

## 2.2 To Weight or Not To Weight?

There is virtually no debate about the need to use some form of weights, at least the probability of selection, when one is interested in estimating population aggregates like totals and means. From a design-based perspective then, releasing the weights associated with sample individuals is crucial for accurate analysis. The issue of the relevance

of weights comes when we introduce statistical models: e.g., should a regression analysis be weighted vs. unweighted? For quite different discussions of the role of sample weights in regression modeling see DuMouchel and Duncan [4] and Kott [14].

Weights are almost always justified only from a frequentist, design-based perspective. In particular, they yield correct frequentist properties under finite population setup for aggregates. When combined with complex models via methods of estimating equations they appear to give a prescription for frequentist inference, and may provide protection from model misspecification, although only with respect to the distribution over all possible sample realizations.

From a model-based perspective, elements reflecting stratification and clustering belong in the model itself if they have any importance. Weights from equation (4) are insufficient for this purpose since they include non-sampling components due to adjustment for nonresponse and post-stratification. For a broader discussion on weighting from a Bayesian perspective see Sugden and Smith [29], Smith [27], Smith and Sugden [28], and Fienberg [8].

### 3 Example: Disability Among Elderly

Both the design-based estimating equation approach to defining population quantities and the pseudo-likelihood approach collapse for complex hierarchically-structured models with latent variables. We illustrate this point briefly by an example.

#### 3.1 Background on the National Long Term Care Survey

The National Long Term Care Survey (NLTCS) is a survey of the U.S. Medicare eligible population (aged 65+). The survey was originally carried out in 1982, and individuals were then reinterviewed in 1984, 1989, 1994, 1999, and 2004. Those who died were replaced with “newly-eligible” Medicare population members. From the 1989 wave onward, the primary sponsor of the survey has been the National Institute on Aging.

The NLTCS measures disability primarily based on 27 binary measures known as activities of daily living (ADL) and instrumental activities of daily living (IADL). Following Erosheva, Fienberg, and Joutard [5], we focus on a subset 16 binary ADLs and IADLS: eating, getting in/out of bed, getting around inside, dressing, bathing, using a toilet, doing heavy house work, doing light house work, doing laundry, cooking, grocery shopping, getting about outside, traveling, managing money, taking medicine, and telephoning.

Erosheva, Fienberg, and Joutard [5] describe an approach based on a Grade of Membership (GoM) model for analyzing disability using data on these 16 measures from the NLTCS for  $n = 21,574$  individuals. The resulting  $2^{16}$  contingency table contains 65,536 cells, only 3,152 of which are non-zero).

- 82% of the cell counts are less than 5.

- 4% of cell counts are greater than 20.
- 18% of the individuals exhibit no disabilities and 3% exhibit all 16.

This is a table of “unweighted” counts. The NLTCs has a relatively complex multistage design and uses a three-part weighting structure similar to that in equation (4). Individual weights have been part of the NLTCs data release, and for each individual we have the components of the weights, as well as their product.

It is worth noting that the weights associated with individuals change from one survey wave to another because of relative complex design rules, and there has in fact been considerable controversy associated with their calculation.<sup>1</sup> For the 2004 (and last) wave of the survey, the final combined weights can vary by as much as a factor of 10. Thus a table of “weighted” counts for the 16 disability measures would look quite different from an unweighted table.

### 3.2 Bayesian Grade of Membership (GoM) Model

Let  $\mathbf{x}$  be the matrix of observed responses  $x_{ij}$  for subjects  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, J$ , and let  $K$  be the number of mixture components (extreme profiles) in the GoM model. Every individual has a partial membership vector  $g_i = (g_{i1}, g_{i2}, \dots, g_{iK})$  of  $K$  nonnegative random variables that sum to 1. The membership scores define how close an individual is to each of the extreme profiles. The statistical task is to choose  $K$  and then estimate the the partial membership vectors.

For the NLTCs data,  $\lambda_{ikj} = P(x_{ij} = 1 | g_{ik} = 1)$ , is the probability of the  $i$ th individual being disabled on the activity  $j$  for a complete member of extreme profile  $k$ . Erosheva, Fienberg, and Joutard augment the observed data for each subject with realizations of latent classification variables  $z_i = (z_{i1}, \dots, z_{iJ})$ . We denote by  $\mathbf{z}$  the matrix of latent classifications  $z_{ij}$ , and let  $z_{ijk} = 1$ , if  $z_{ij} = k$ , and  $z_{ijk} = 0$  otherwise.

If the distribution of membership scores is Dirichlet with parameters  $\alpha$ , then the joint probability model for the parameters and augmented data is

$$\begin{aligned} p(\mathbf{x}, \mathbf{z}, \mathbf{g}, \boldsymbol{\lambda}, \alpha) &= p(\boldsymbol{\lambda}, \alpha) \cdot p(\mathbf{x}, \mathbf{z}, \mathbf{g} | \boldsymbol{\lambda}, \alpha) \\ &= p(\boldsymbol{\lambda}, \alpha) \prod_{i=1}^N [p(z_i | g_i) p(x_i | \boldsymbol{\lambda}, z_i) \cdot D(g_i | \alpha)], \end{aligned}$$

---

<sup>1</sup>See <http://www.icpsr.umich.edu/NACDA/news.html#nlctcs>

where

$$\begin{aligned} p(z_i|g_i) &= \prod_{j=1}^J \prod_{k=1}^K g_{ik}^{z_{ijk}}, \\ p(x_i|\boldsymbol{\lambda}, z_i) &= \prod_{j=1}^J \prod_{k=1}^K \left( \lambda_{kj}^{x_{ij}} (1 - \lambda_{kj})^{1-x_{ij}} \right)^{z_{ijk}}, \\ Dir(g_i|\alpha) &= \frac{\Gamma(\sum_k \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} g_{i1}^{\alpha_1-1} \dots g_{iK}^{\alpha_K-1}. \end{aligned}$$

Erosheva, Fienberg and Joutard use as a prior

$$p(\boldsymbol{\lambda}, \alpha) = p(\alpha) \prod_{k=1}^K \prod_{j=1}^J p(\lambda_{kj}), \quad (5)$$

where  $p(\lambda_{kj})$  is  $Beta(\eta_1 = \eta_2 = 1)$ . The joint distribution of the parameters and augmented data is then

$$p(\boldsymbol{\lambda})p(\alpha) \left( \prod_{i=1}^N D(g_i|\alpha) \right) \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^K \left( g_{ik} \lambda_{kj}^{x_{ij}} (1 - \lambda_{kj})^{1-x_{ij}} \right)^{z_{ijk}}. \quad (6)$$

To summarize, we have an elaborate Bayesian hierarchical mixed membership model for the cell counts in a  $2^{16}$  with informative latent structure. The membership scores are random and controlled via the hierarchical structure. The levels of hierarchy matter, as do choices of Dirichlet prior parameters. Doing a weighted analysis here using the complex survey weights makes little sense. Where would one place the weights and why? Clearly a blind application of equation (3) with weights (4) cannot work. Manton, Lamb, and Gu [17] do incorporate weights into their analyses of medicare costs associated with subgroups defined by a non-Bayesian version of the GoM model, but with virtually no discussion of the appropriateness.

We can now state the privacy protection questions for this type of data and analysis:

1. How do we protect the privacy of the individual responses represented by the  $2^{16}$  contingency table? Perhaps by standard methods focused on small counts? For example, see Fienberg and Slavkovic [10].
2. What about privacy protection for a weighted  $2^{16}$  contingency table?
3. For those who want to incorporate weights into their statistical analyses, can we release the weights with no alteration?
4. If the weights might be used by an intruder to help in the identification of individuals in the sample, how should they be altered to protect confidentiality?

## 4 Confidentiality and Weights

At last we are ready to address the issue of what is known about confidentiality protection and survey weights.

- **Question 1:** What information do weights provide to an intruder intent on identifying individuals in the sample?
- **Question 2:** Do they increase the probability of disclosures? If so, how?

We have identified four different sets of advice related to the protection of sampling weights in the literature and discuss them briefly below.

### 4.1 de Waal and Willenborg [3]

de Waal and Willenborg argue that sampling weights can provide indirect identifying information regarding membership in substrata defined by sets of post-stratification variables. The idea is that an intruder might have accurate information on post-stratification population counts and can then use the numbers of sample people with given weights to match individuals with post-strata. This really focuses only on the third component of the weights in equation (4). How likely is this problem to occur? Is the situation different in the Netherlands as compared to in the U.S.? The complexity of weight calculations in surveys such as the NLTCs suggests that this particular form of attack is unlikely to have much traction.

de Waal and Willenborg suggests two ways to protect the weights: (1) subsampling to reduce the probability of correct matches by an intruder attempting record linkage, and (2) adding noise to the weights so that the post-strata are not identifiable. We note that the second strategy messes up the original rationale of reweighting to match population controls!

### 4.2 Rubin [23] on Multiple Imputation for Confidentiality Protection

Rubin [23] proposed the use of multiple imputation (see [22]) to address the confidentiality problem. This amounts to first computing the relevant posterior distribution for the model of interest and generating multiple samples from it. These are synthetic samples and thus, Rubin argues, they automatically solve the confidentiality problem. Ragunathan, Reiter, and Rubin [21] provide details of the use of multiple imputation for this purpose.

How does this address the issue of survey weights? One might use weights to construct the posterior (although most Bayesians that I know wouldn't), but once one has the posterior, there is no requirement that the the sample draws from it should be weighted! In fact, this was one of the primary virtues of the multiple imputation approach as Rubin described it: No matter what the original sample design, the multiple-imputed data sets can take the form of simple random samples, and thus the



problem of survey weights vanishes.

If the claim of protection were correct, this would be a suitable prescription for dealing with the privacy protection problem for the  $2^{16}$  contingency table from the NLTCs!

### 4.3 Mitra and Reiter [18]

Mitra and Reiter [18] apply a variant of the multiple imputation approach to generate partially synthetic data and then investigate how to report sampling weights. They argue that, when synthesizing some sampling design variables, one must adjust weights to reflect the new synthesized values. They consider two approaches: (i) recalculating the weights (RCAL) to be consistent with the synthetic values, effectively making the synthetic sample “representative” of the population, and (ii) copying and pasting (CPP) the original weights of records whose original design variables match the synthetic ones. They note that, “[t]he RCAL method preserves some properties of the original sampling weights that the CPP method does not; for example, the sum of the RCAL weights equals the sum of the observed weights, whereas the sums are not necessarily equal for the CPP weights. Additionally, the CPP method cannot be applied unless exact matches are available.” They then proceed to use simulation studies for two simple designs to investigate the implications for data quality of using survey-weighted analyses based on these two methods. They also explore the implications for data swapping for confidentiality protection.

They conclude, “The simulations in this paper illustrate the importance of survey weights when altering design variables to limit disclosure risks. Releasing the original weights can lead to biased inferences or compromise identity of respondents. At least for partially synthetic data, recalculating the weights to be consistent with released values can improve design-based estimation. Unfortunately, this approach does not appear to improve inferences sufficiently when using data swapping of design variables. Further research is needed to investigate the viability of the recalculation approach for more complicated multi-stage sampling schemes.”

Throughout their discussion, Mitra and Reiter actually pay no attention to the issue of the privacy problems associated with the original survey weights, although their recalculation may provide some protection.

### 4.4 Singh et al. [24]: MASSC

Singh, Yu, and Dunteman [24] have developed an elaborate package for disclosure limitation of survey data which they label MASSC (for Micro-agglomeration, Substitution, Subsampling, and Calibration). It involves an elaborate process of aggregation, substitution and subsampling which is also described in the paper by Singh [25] presented at this workshop. They address the issue of disclosiveness of sampling weights indirectly through their multi-step procedure. Because of the subsampling in particular, they recalculate the sampling weights in a step they refer to as “Calibration” and claim that

this reduces the risk. For further details, see Singh [25]. It is difficult to evaluate this claim and to separate the treatment of sampling weights from the remaining alterations to the data performed by MASSC.

## 5 My Prescription

The different approaches described above for addressing the potential disclosiveness of survey weights are at best ad hoc, and never directly address why we should worry about weights in the first place, except perhaps for multiple imputation. Clearly, subsampling and adding noise to weights will have predictable protection at some level and thus may be reasonable strategies, but as we noted above, these alter the properties of weight that made them of value to design-based inference to begin with.

As we suggested at the outset of this paper, the privacy protection concern with weights is inextricably intertwined with the role weights play in statistical analyses. Thus, we believe that we need to get those carrying out surveys and preparing file for secondary data analysis to do several things:

- Get rid of population controls (component three of equation (4)) and thus remove the largest part of the confidentiality concerns that emanate from sampling weights.
- Stop insisting that model-based analyses incorporate weights (although retain the idea of weighting for aggregate quantities such as means and totals). For purposes of estimating aggregate quantities, it may well be advisable to report weights in altered form, e.g., shrunken in order to reduce the impact of large weights due to very small probabilities of selection.
- Think about new approaches to survey design that deal *de novo* with confidentiality concerns as well as with analytical goals, not just traditional sample efficiency goals.
- Share real design information so that it can be incorporated directly in statistical models for analysis rather than indirectly via sampling weights.
- Address perhaps the most important disclosure problems associated with cluster sampling at the design stage by eliminating clustering wherever possible.

In many ways this prescription represents a radical departure from current survey practice.

## 6 Summary

The disclosure dangers associated with sample survey weights are difficult to assess, and the current practice of most survey organizations, including the U.S. Bureau of the

Census and other U.S. statistical agencies, is to release survey weights with public use samples in unaltered form.

In this paper we have attempted to address the following series of questions: (1) What are survey sampling weights? (2) Why are statisticians interested in them? Or, “To weight or not to weight”? (3) Do weights pose issues for the analysis of released microdata? (4) What hazards do weights pose for confidentiality protection? (5) How can problems be averted or mitigated? We have summarized the limited literature on the topic of survey weights and disclosure limitation and offered our own prescription for addressing the weighting problem.

### Acknowledgments

The research reported here was supported in part by NSF grants EIA9876619 and IIS0131884 to the National Institute of Statistical Sciences, by NSF Grant SES-0532407 and NIH Grant No. R01 AG023141-01 to the Department of Statistics, Carnegie Mellon University, and by Army contract DAAD19-02-1-3-0389 to CyLab at Carnegie Mellon University. Based on a talk presented at the *Workshop on Data Confidentiality: The Next Five Years*, held at the National Center for Health Statistics, Hyattsville, MD, May 1–2, 2008.

### References

- [1] Bertolet, M. (2008). *To Weight or Not to Weight: Incorporating Sampling Effects into Model-Based Survey Analysis*. Ph.D. Dissertation. Department of Statistics, Carnegie Mellon University, Pittsburgh, PA.
- [2] Binder, D. A. and Patak, Z. (1994). Use of Estimating Functions for Estimation from Complex Surveys. *Journal of the American Statistical Association*, **89**, 1035–1043.
- [3] de Waal, T. and Willenborg, L. (1997). Statistical Disclosure Control and Sampling Weights. *Journal of Official Statistics*, **13**, 417–434.
- [4] DuMouchel, W. H. and Duncan, G. J. (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association*, **78**, 535–543.
- [5] Erosheva, E. A., Fienberg, S. E., and Joutard, C. (2007). Describing Disability Through Individual-Level Mixture Models for Multivariate Binary Data. *Annals of Applied Statistics*, **1**, 502–537.
- [6] Fienberg, S. E. (1980). The Measurement of Crime Victimization: Prospects for Panel Analysis of a Panel Survey. *The Statistician* **29**, 313–350.
- [7] Fienberg, S. E. (1989). Modeling Considerations: Discussion from a Modeling Perspective. In *Panel Surveys*, D. Kasprzyk, G. Duncan, G. Kalton, and M. P. Singh, Eds., 512–539. New York: Wiley.

- [8] Fienberg, S. E. (2008). To Weight or Not to Weight? A Bayesian Model-Based Update. Presented at *Sample Surveys and Bayesian Statistics 2008*, University of Southampton, August, 2008.
- [9] Fienberg, S. E. and Makov, U. E. (1998). Confidentiality, Uniqueness, and Disclosure Limitation for Categorical Data. *Journal of Official Statistics*, **14**, 385–397.
- [10] Fienberg, S. E. and Slavkovic, A. B. (2008). A Survey of Statistical Approaches to Preserving Confidentiality of Contingency Table Entries. In *Privacy Preserving Data Mining: Models and Algorithms*, C. Aggarwal and P. S. Yu, Eds., 289–310. New York: Springer-Verlag.
- [11] Hansen M. H., Madow, W. G. and Tepping, B. J. (1983). An Evaluation of Model-Dependent and Probability-Sampling Inferences in Sample Surveys (with discussion). *Journal of the American Statistical Association*, **78** (384), 776–807.
- [12] Horvitz, D. G. and Thompson, D. J. (1952). A Generalization of Sampling Without Replacement From a Finite Universe. *Journal of the American Statistical Association*, **47**, 663–685.
- [13] Kalton, G. (1989). Modeling Considerations: Discussion from a Survey Sampling Perspective. In *Panel Surveys*, D. Kasprzyk, G. Duncan, G. Kalton, and M. P. Singh, Eds., 575–587. New York: Wiley.
- [14] Kott, P. S. (1991). A Model-Based Look at Linear Regression with Survey Data. *American Statistician*, **45**, 107–112.
- [15] Kovačević, M. S. and Binder, D. A. (1997). Variance Estimation for Measures of Income Inequality and Polarization—The Estimating Equations Approach. *Journal of Official Statistics*, **13**, 41–58.
- [16] Lohr, S. (1999). *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.
- [17] Manton, K. G., Lamb, V. L., and Gu, X. (2007). Medicare Cost Effects of Recent U.S. Disability Trends in the Elderly: Future Implications. *Journal of Aging Health*, **19**, 359–381.
- [18] Mitra, R. and Reiter, J. P. (2006). Adjusting Survey Weights When Altering Identifying Design Variables Via Synthetic Data. In *Privacy in Statistical Databases*, J. D. Ferrer and L. Franconi, Eds., Lecture Notes in Computer Science, 177–188. Berlin: Springer-Verlag.
- [19] Pfeffermann, D. (1993). The Role of Sampling Weights When Modeling Survey Data. *International Statistical Review*, **61**, 317–337.
- [20] Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H. and Rasbash, J. (1998). Weighting for Unequal Selection Probabilities in Multilevel Models (with discussion). *Journal of the Royal Statistical Society, Series B*, **60**, 23–56.

- [21] Raghunathan, T.E., Reiter, J.P., and Rubin, D.B. (2003). Multiple Imputation for Statistical Disclosure Limitation. *Journal of Official Statistics*, **19**, 1–16.
- [22] Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- [23] Rubin, D. B. (1993). Discussion: Statistical Disclosure Limitation. *Journal of Official Statistics*, **9**, 462–468.
- [24] Singh, A. C., Yu, F., and Dunteman, G. H. (2003). MASSC: A New Data Mask for Limiting Statistical Information Loss and Disclosure. In *Proceedings of the Joint UNECE/EUROSTAT Work Session on Statistical Data Confidentiality*, Luxembourg, April 7–9, 2003. (Available at <http://www.unece.org>)
- [25] Singh, A. C. (2008). Maintaining Analytical Utility While Protecting Confidentiality of Survey and Nonresponse Data. Presented at *Workshop on Data Confidentiality: The Next Five Years*, National Center for Health Statistics, Hyattsville, MD, May 1–2, 2008.
- [26] Skinner, C. J. and Holmes, D. J. (1998). Estimating the Re-identification Risk Per Record in Microdata. *Journal of Official Statistics*, **14**, 361–372.
- [27] Smith, T. F. M. (1988). To Weight Or Not To Weight, That Is The Question (with discussion). In *Bayesian Statistics 3*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, Eds., 437–451. Oxford University Press.
- [28] Smith, T. F. M. and Sugden, R. A. (1988). Sampling and Assignment Mechanisms in Experiments, Surveys and Observational Studies. *Review of the International Statistical Institute*, **56** (2), 165–180.
- [29] Sugden, R. A. and Smith, T. F. M. (1984). Ignorable and Informative Designs in Survey Sampling Inference. *Biometrika*, **71**, 495–506.

