A COMPARISON OF TWO ARRESTEE DRUG USE ESTIMATION METHODS

Y. Michael Yang and Dean R. Gerstein

NORC at the University of Chicago, 1350 Connecticut Ave., N.W., Washington, DC 20036

KEY WORDS: Post-Sampling Stratification, Weighting Adjustment, Mean Square Error, Design Effect, Survey Cost

Sponsored by the National Institute of Justice (NIJ), the Arrestee Drug Abuse Monitoring program (ADAM) collects substance abuse and other data from arrestees in 35 sites, generally urban counties, across the United States. This paper compares two arrestee drug use estimation methods in terms of statistical efficiency, cost efficiency, and relevant administrative measures. The analysis is based on ADAM data collected from January 2000 through September 2001, with additional administrative data on cost and timeless of data delivery through mid-2003.

Section one reviews ADAM sample designs. Section two describes the old estimation method based on post-sampling stratification. Section three presents empirical evidence on the deficiencies of the old estimation method. Section four proposes a new estimation method based on a two-stage weighting scheme. Section five compares the statistical efficiency, cost efficiency, and timeliness of the two estimation methods. Finally, section six provides some remarks about possible future improvements in ADAM methodology.

1. ADAM Sample Designs

ADAM is designed to monitor drug use trends and provide data useful to local authorities for developing drug-control policies. NIJ initially did not seek to make national estimates based on ADAM data, so it did not require a probability sample of participating sites, which were selected instead through a nonrandom process. Within each site, however, sample selection of male arrestees is carried out based on probability principles[1]. Depending on the population structure of the site, ADAM features a range of sample designs. A *Single Jail Design* is used when a site has only one arrestee booking facility (jail). When a site has two to four facilities, ADAM uses a *Stratified Design* with proportional sample allocation across jails. When there are more than four jails, ADAM adopts a *Two-Stage Stratified Cluster Design* where jails are sampled in the first stage and arrestees are sampled from sampled jails in the second stage. Finally, the *Feeder Jail Design* is used in sites with a large central booking facility and many small satellite facilities with variable transportation from the satellite facilities to the central facility.

Within each sample facility, ADAM follows a systematic sampling protocol to select arrestees. Sampling and data collection typically take place over a 14-day period—usually two sequential weeks—per quarter. For any given day, the target population consists of all eligible arrestees who are booked during the 24-hour calendar day. ADAM interviewers typically work an 8-hour shift starting from 4:00PM to midnight. During any day, there is a continuous *flow* of arrestees into the jail. At any point in time during that day, there is an accumulated *stock* of arrestees who have been booked and are waiting for processing by a judge or other authority. ADAM stratifies the daily population of arrestees into stock arrestees and flow arrestees and selects a systematic sample of arrestees from each stratum (Rhodes, *et al*, 2002). Therefore, within each stratum (stock or flow) on a given day, sampling probabilities only depend on the sample size and the total number of eligible arrestees booked.

2. The Old Estimation Method

ADAM currently uses post-sampling stratification to adjust for potential bias due to unequal selection probabilities, nonresponse, and noncoverage[2]. Basically, the method is to define post-sampling strata such that arrestees within the same stratum have the same probabilities of being selected and interviewed by ADAM. The following factors are used to define the poststrata: (1) size of the jail, (2) stock or flow, (3) offense charge severity, and (4) daily case flow. The last variable measures the number of bookings on a daily basis, i.e., the daily population size. Table 1 summarizes the post-sampling stratification scheme under the old estimation method. Each large jail is

[1] ADAM collects data from adult male, adult female and juvenile arrestees. Female and juvenile data collections are based on convenient samples, so this analysis only applies to adult male estimation. Adult male arrestees comprise about 80% of all arrestees in ADAM sites.

[2] Existing ADAM documentation only discusses adjusting for unequal selection probabilities through post-sampling stratification weighting (Hunt and Rhodes, 2001). The weighting scheme actually adjusts for selection probabilities, nonresponse and noncoverage in one step.

stratified separately, whereas small jails are combined before arrestees are stratified by other variables.

Table 1: ADAM Post-Sampling Stratification Weighting Strata

Strata	Facility Size	Stock Flow	Charge Severity	Daily Case Flow
1	Large	Stock	Felony	High
2				Medium
3				Low
4			Misd.	High
5				Medium
6				Low
7			Other	High
8				Medium
9				Low
10		Flow	NA	High
11				Medium
12				Low
13	Small	Stock	Felony	NA
14			Misd.	
15			Other	
16		Flow	NA	

The weight for a respondent i in stratum h is computed as,

$$w_{ih} = \frac{N_h}{n_{rh}}, i \in h \qquad [1]$$

where N_h is the total number of eligible arrestees in stratum h, and n_{rh} is the number of respondents in stratum h[3].

Then, the estimator for the population mean $\overline{Y}$ is:

$$\hat{\overline{Y}} = \sum_{h=1}^{h} \left[\frac{N_h}{N} \right] \left(\sum_{i=1}^{n_{rh}} \frac{y_{ih}}{n_{rh}} \right) \qquad [2]$$

And the estimator for the sampling variance of $\hat{\overline{Y}}$ is:

$$\hat{V}(\hat{\overline{Y}}) = \sum_{h=1}^{h} \left[\left\{ \frac{N_h}{N} \right\}^2 \right] \left\langle \frac{\sum_{i=1}^{n_{rh}} \frac{\left(y_{ih} - \hat{\overline{y}}_h \right)^2}{n_{rh} - 1}}{n_{rh}} \right\rangle \qquad [3]$$

where N is the total number of arrestees per site; N_h is the total number of arrestees in stratum h; n_{rh} is the number of respondents in stratum h; y_{ih} is the observed value of Y for arrestee i in stratum h; and $\hat{\overline{y}}_h$ is the sample mean of Y in stratum h.

3. An Evaluation of the Old Method

For post-sampling stratification estimation to be effective, the following assumptions have to be true.

(a) Selection probabilities or response rates differ across strata;
(b) The distribution of survey variables is homogeneous within strata and heterogeneous across strata.

For subsequent presentation, we define the sampling probability in subgroup g as

$$p_g = \frac{n_g}{N_g} \qquad [4]$$

where n_g and N_g are the sample size and population size of subgroup g, respectively. We further define the response rate of subgroup g as

$$r_g = \frac{n_{rg}}{n_g} \qquad [5]$$

where n_{rg} is the size of the responding sample. With these definitions, we present some results below.

Table 2: ADAM Sampling Probabilities and Response Rates by Stratification Variables

Factor	Subgroup (g)	Sampling Probability (p_g)	Response Rate (r_g)
Facility Size	Large	0.20	0.55
	Small	0.38	0.52
Stock vs. Flow	Stock	0.21	0.49
	Flow	0.25	0.63
Charge Severity	NA[4]	0.25	0.63
	Felony	0.26	0.56
	Misd.	0.21	0.43
	Other	0.14	0.62
Daily Case Flow	NA[5]	0.38	0.52
	High	0.18	0.53
	Medium	0.20	0.57
	Low	0.24	0.58
Overall		0.23	0.54

[3] ADAM computes four weights. Our presentation here assumes that the total weight for interview data analysis is of interest. Generalization to the other three weights is straightforward.

[4] These are flow cases that are not subdivided by Charge Severity.
[5] These are cases from small facilities that are not subdivided by Daily Case Flow.

Table 2 shows the main effects of the poststratification variables on sampling probabilities and response rates where each subgroup is defined by a single variable. First, the relative magnitudes of sampling probabilities for the subgroups are consistent with expectations. Arrestees from small facilities have been sampled at a much higher rate than those from large facilities. Arrestees booked during the flow period have slightly higher probabilities of being sampled than those booked earlier during the day. Those charged with a felony are more likely to be sampled than those charged with a misdemeanor, who in turn are more likely to be in the sample than the residual group. The number of arrestees booked is inversely correlated with the sampling probability. Second, although all variables have an impact on p_g, the only drastic difference in sampling probabilities is between small and large facilities. Third, response rate does not seem to be affected by facility size and daily case flow, but it is clearly related to stock/flow and charge severity. In particular, flow cases and felony cases have a higher response rate than their counterparts.

Table 3 below presents similar information but each subgroup is a stratum under post-sampling stratification. Again, arrestees from small facilities have far higher selection probabilities. In addition, felony cases, flow cases and cases booked on low flow days are more likely to be selected into the sample. The stratum with the highest sampling probability is Small-Flow at .45, followed by Small-Stock-Felony at .37. With regard to response rates, again, felony and flow cases are more likely to respond to the survey.

Table 3: ADAM Sampling Probabilities and Response Rates by Weighting Strata

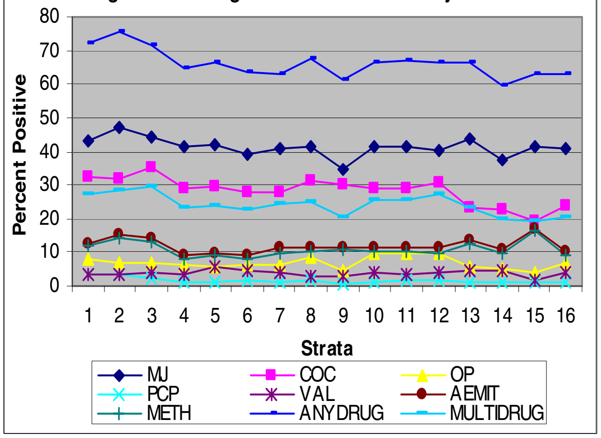
Strata (g)	Facility Size	Stock vs. Flow	Charge Severity	Daily Case Flow	Prob. (p_g)	Resp. Rate (r_g)
1	Large	Stock	Felony	High	0.24	0.52
2				Medium	0.23	0.57
3				Low	0.25	0.58
4			Misd.	High	0.18	0.41
5				Medium	0.17	0.43
6				Low	0.19	0.44
7			Other	High	0.10	0.63
8				Medium	0.13	0.63
9				Low	0.18	0.55
10		Flow	NA	High	0.18	0.62
11				Medium	0.23	0.67
12				Low	0.30	0.69
13	Small	Stock	Felony	NA	0.37	0.58
14			Misd.		0.36	0.44
15			Other		0.24	0.63
16		Flow	NA		0.45	0.53
Overall					0.23	0.54

The fact that sampling and response probabilities vary across strata does not ensure the effectiveness of poststratification adjustment. Another crucial condition is that there is significant variance across strata with respect to the survey variables, i.e., $\bar{Y}_h$ varies across strata. Assuming that data are missing completely at random within strata, we use the unweighted estimates per stratum $\hat{\bar{y}}_{uh}$ to approximate $\bar{Y}_h$.

$$\hat{\bar{y}}_{uh} = \sum_{i=1}^{n_{rh}} \frac{y_{ih}}{n_{rh}}, \; i \in h \qquad [6]$$

Figure 1 presents the unweighted estimates per strata for 9 ADAM drug test variables. The most remarkable feature shown in Figure 1 is the lack of variance in the unweighted estimates across strata. Secondly, estimates in the first three strata tend to be greater than the rest of the strata for most variables. It appears that arrestees in large facilities and arrestees with more serious charges tend to abuse drugs more frequently than their counterparts.

Figure 1: Unweighted Urine Estimates by Strata



Further analysis indicates that this pattern is largely driven by Charge Severity which is the only weighting variable that is correlated with the survey variables. Clearly, if $\bar{Y}_h$ does not show appreciable variation across strata, adjusting w_h to W_h will not have a significant impact on $\hat{\bar{Y}}$.

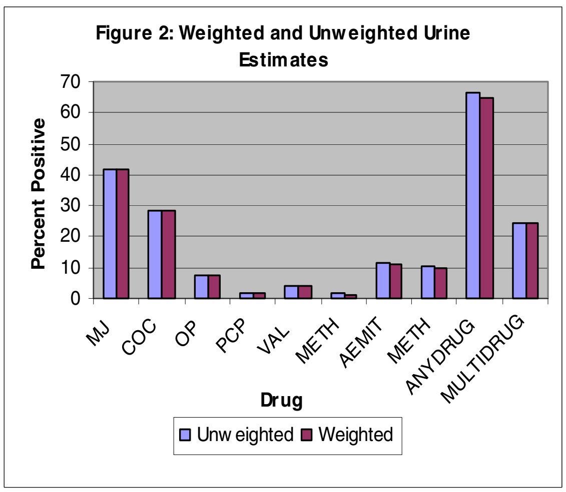
To assess the effectiveness of post-sampling stratification estimation in terms of reducing mean square error (MSE), we follow Kish's advice to compare the weighted and unweighted estimates and their relative MSEs (Kish, 1992). A total of 42 ADAM variables are involved in the comparisons. We first define the weighted, unweighted and bias estimates as follows.

$$\hat{\bar{y}}_w = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}, \qquad [7]$$

$$\hat{\bar{y}}_u = \sum_{i=1}^{n} \frac{y_i}{n}, \qquad [8]$$

$$\hat{B} = \hat{\bar{y}}_u - \hat{\bar{y}}_w, \qquad [9]$$

where n represents the total responding sample. We assume that the weighted estimates are unbiased relative to the unweighted estimates. Thus, $\hat{B}$ represents an estimate of the bias of the unweighted estimates[6]. Figure 2 compares the weighted and unweighted estimates of 10 drug test variables over all sites and quarters where ADAM data are available at the time of this evaluation. It shows that the differences between weighted and unweighted estimates are extremely small.



Similar analysis on other variables and at the site level reveals the same pattern. Across all sites and all quarters, there are 200 independent data files. For the 5 most important statistics derived from drug test data, we computed the following statistics from each file: unweighted estimate $\hat{\bar{y}}_u$, weighted estimate $\hat{\bar{y}}_w$, estimated bias $\hat{B}$, standard error of the unweighted estimate $\hat{\sigma}_{\hat{\bar{y}}_u}$, standard error of the weighted estimate $\hat{\sigma}_{\hat{\bar{y}}_w}$, bias ratio of the unweighted estimate B/S, mean square error of the unweighted estimate $MSE_{\hat{\bar{y}}_u}$, mean square error of the weighted estimate $MSE_{\hat{\bar{y}}_w}$, and the ratio of the two mean square errors $MSE_{\hat{\bar{y}}_w} / MSE_{\hat{\bar{y}}_u}$. In addition, we computed the design effects introduced by unequal weighting. Due to space limitation, we only provide the following summaries here. First, the weighted and unweighted estimates are remarkably close per site per quarter, giving little evidence of significant bias in the unweighted estimates even at the site level. Second, the poststratification estimator tends to have greater MSE than the unweighted estimator. Of the reported MSE ratios, two thirds of them are greater than 1. Obviously, bias is generally small relative to standard error with small samples at the site level. Finally, the average design effect due to unequal weighting is about 1.4 for both questionnaire weights and urine weights. For details about site level analysis, please see Yang and Gerstein, 2003.

In conclusion, the old estimation method has not been effective in terms of reducing bias or variance. Although some stratification variables influence selection and response probabilities, the population shows little variance across post-strata with respect to the survey variables. As Cochran (1977) points out, the gain in precision from a stratified sample over a random sample will be small unless the survey variables vary greatly across strata. The same general principle applies to poststratification. The current estimation method is not only statistically inefficient, it also causes delays in data dissemination because the collection and processing of population data (N_h in [1]) are expensive and time consuming. Besides, the population data collected, typically retrospectively, are often inaccurate due to deficiencies of the record keeping systems at the sites. In the next section, we describe an alternative estimation methodology that has superior statistical efficiency, cost efficiency, timeliness, and accuracy.

4. The New Estimation Method

The post-sampling stratification weight under the old method may be decomposed into two components: the base weight and the nonresponse weight. That is, equation [1] may be written as,

$$w_{ih} = \frac{N_h}{n_{rh}} = \left(\frac{N_h}{n_h}\right)\left(\frac{n_h}{n_{rh}}\right), i \in h \qquad [10]$$

Notice that the first term on the right hand side is the inverse of selection probability and the second term is the inverse of response rate of post stratum h. Sample size n_h cancels out because the weighting classes for

---

[6] As noted earlier, $\hat{\bar{y}}_w$ is not strictly unbiased if W_h is not accurate.

base weight and nonresponse weight are identical. However, our evaluation reveals that sampling probability and nonresponse probability are affected by different factors. The former is primarily determined by the size of the facility, with additional within-facility variation by date and between stock and flow. The latter is mainly affected by stock/flow and charge severity. To closely represent these two processes, we adjust for selection probabilities and nonresponse separately.

4.1 Stage One: Base Weight

If we stratify the sample by day and stock/flow within each day, we can reasonably assume that selection probabilities are the same within strata. We use the following notation.

- h sampling stratum defined by data collection date and *stock/flow*. The total number of strata is determined by the number of jails per site and the length of data collection period in each jail.
- N_h number of arrestees in stratum h
- n_h sample size of stratum h

Then, the base weight for a sample arrestee i of stratum h is:

$$w_{ih} = \frac{N_h}{n_h} \qquad [11]$$

Notice that the N_h in [11] is defined by date and stock/flow only within each facility. Charge severity, which contributes the most to cost and does not affect selection probability, drops out from the definition of population data.

4.2 Stage Two: Nonresponse Adjustment Weight

Charge severity, together with stock/flow, does affect nonresponse. We make the following assumptions regarding the mechanism of nonresponse, the population distribution of survey variables, and the pattern of missing data.

a) The probability of nonresponse is affected by stock/flow and charge severity (felony and other);
b) The means of survey variables are homogeneous within adjustment classes;
c) Data are missing at random within adjustment classes.

Let's define the following notation:

- c nonresponse adjustment class defined by *stock/flow* and *severity* (felony and other)
- n_c sample size of class c
- n_{rc} number of respondents in class c

Then, the nonresponse adjustment weight for respondent i in class c is:

$$w_{ic} = \frac{n_c}{n_{rc}} \qquad [12]$$

The final weight of i in sampling stratum h and nonresponse adjustment class c is

$$w_i = w_{ih} * w_{ic} * R_i \qquad [13]$$

where R_i is 1 for respondents and 0 for nonrespondents.

Since charge severity is only involved in nonresponse adjustment, the required data are already available from the ADAM samples. This change has significant cost implications for ADAM. Notice that expression [13] only addresses within-jail sampling and response probabilities. Under multistage designs where jails—clusters or PSUs—are sampled in the first stage, w_i should be multiplied by the first stage sampling weight appropriate for the design[7].

4.3 Estimator of $\overline{Y}$ and $V(\hat{\overline{Y}})$

To present general estimators that are applicable to all ADAM designs, we introduce the following notation.

- $h = 1,2,\ldots,H$ is the stratum number
- $i = 1,2,\ldots,n_h$ is the cluster—jail or PSU—number within stratum h
- $j = 1,2,\ldots,m_{hi}$ is the case number within cluster i of stratum h
- f_h is the first stage sampling rate for stratum h
- w_{hij} is the weight for observation j in cluster i of stratum h
- y_{hij} is the observed value of variable y for observation j in cluster i of stratum h

[7] We chose not to formally introduce the first stage sampling weight into the new weighting scheme due to the lack of population data about jails that are not included in the site sample.

The H strata should not be confused with the within-jail sampling or post-sampling strata defined earlier. These H strata are the first stage sampling strata under stratified two-stage cluster designs. For a multistage sample design, the variance estimation method depends only on the first stage of the sample design. So, the required input includes only first-stage cluster and first-stage stratum identification. For a design without stratification, we will set $H = 1$; for a design without clusters, we will let $m_{hi} = 1$ for all combinations of h and i. The survey weight w_{hij} incorporates selection probabilities of all sampling stages as well as adjustments for nonresponse. Here, we assume that w_{hij} is the product of the cluster weight and w_i as computed in expression [13].

The estimator of the population mean is

$$\hat{\bar{Y}} = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}} \qquad [14]$$

The estimator of the variance of $\hat{\bar{Y}}$ is

$$\hat{V}(\hat{\bar{Y}}) = \sum_{h=1}^{H} \frac{n_h(1-f_h)}{(n_h-1)} \sum_{i=1}^{n_h} (e_{hi.} - \bar{e}_{h..})^2 \qquad [15]$$

where

$$e_{hi.} = \frac{\left(\sum_{j=1}^{m_{hi}} w_{hij}(y_{hij} - \hat{\bar{Y}})\right)}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}} \qquad [16]$$

$$\bar{e}_{h..} = \frac{\sum_{i=1}^{n_h} e_{hi.}}{n_h} \qquad [17]$$

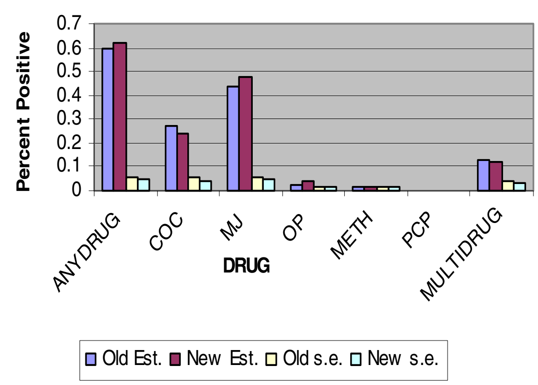
SAS PROC SURVEYMEANS may be used to carry out these estimations (SAS InstituteInc., 1999). For variance estimation, this method obtains a linear approximation for the estimator and then uses the variance estimate for this approximation to estimate the variance of the estimate itself (Woodruff, 1971).

5. Comparisons of the Two Estimation Methods

5.1 Statistical Properties

We conducted parallel tests with a sample of 10 ADAM sites to compare the properties of the post-sampling stratification estimator (OLD) and the alternative estimator (NEW). We compared the point estimates, standard errors, and overall design effects with respect to seven drug test variables. The old and new estimates are remarkably close; most of the differences are within two percentage points, and they rarely exceed five percentage points. The standard errors estimates are also very close, but the differences here are more systematic, as the new standard errors tend to be smaller than the old ones. Figure 3 compares the point estimates and standard errors between the two estimation methods for the Anchorage ADAM site. The other nine testing sites (not presented) show the same pattern.

Figure 3: Point and SE Estimates under the Two Estimators



The comparison of overall design effects depicts a similar picture, with the new estimator having smaller design effects than the old estimator. We then define two MSE ratios to compare the estimated total sampling errors under the two estimators. The first MSE ratio is the MSE of the old estimator divided by the MSE of the unweighted estimator, and the second MSE ratio is the MSE of the new estimator divided by the MSE of the unweighted estimator.[8] The old estimator has greater MSEs than the unweighted estimator for the vast majority of ADAM statistics, but

[8] The MSE of the weighted estimators are simply their respective variances, since we assume that the weighted estimators are unbiased. There are two unbiased estimators and they are very close. The size of the bias of the unweighted estimator depends on which unbiased estimator it is evaluated against.

the new estimator tend to have smaller MSEs than the unweighted estimator. This provides strong evidence that the new estimation method is superior in terms of bias reduction. Overall, the parallel tests show that the new estimation method produces similar point estimates but smaller standard errors, MSEs, and design effects.

5.2 Cost Efficiency

The new estimator is also more cost efficient because it does not require detailed population data as post-sampling stratification does. To estimate base weights all we need is the total number of eligible bookings by stock and flow per day, and the information needed for estimating the nonreponse adjustment weight is already included in the sample.

The cost of population data collection and processing has two components: site costs and central costs. Approximately, the former is the data collection costs and the latter data processing costs. ADAM sites take either of two approaches to providing population data: (1) submitting electronic files produced by the arresting agencies, or (2) submitting printed outputs or photocopies of booking logs, individual case sheets, or similar records. Although the difference in site costs between the two approaches would slightly favor the new method, this comprises a relatively small and negligible sum compared with the central costs.

ADAM central data processing costs differ dramatically for the two estimation methods. Table 4 compares the estimated annualized costs related to population data processing for 35 ADAM sites, based on FY2003 loaded hourly rates for each labor category, actual hours per site during the last half of 2002 to process the current population data, and actual hours per site during early 2003 to process the population data from the test sites.

Table 4. Comparison of Central Costs of Old Versus New Census Methods

Labor Category	OLD		NEW	
	Hours	Costs($)	Hours	Costs($)
Programming	1,280	143,104	0	0
Supervision	442	40,337	21	1,916
Clerical	759	32,282		0
Data Edit/Entry	3,660	107,201	70	2,050
Statistical	640	62,488	576	56,239
Total	6,781	385,412	667	60,205
Difference	-6,014	-325,207		

Once the new estimation method is fully implemented, the estimated cost savings will total approximately $325,000 per year; these savings derive in large part from reductions in data entry and programming costs. Data entry and associated support and supervision are negligible under the new method, and statistical costs are somewhat less. The large commitment of programmer time to fully process both forms of the current population data disappear under the new method, with tasks such as transforming the electronic files into the standard format, running and cleaning the results of the programs that recode arrest charges from the many input formats into ADAM charge codes and check for matching between population and questionnaires, obtaining new population files when those provided prove to be flawed, and so forth, are no longer necessary.

5.3 Timeliness

One of the major objectives of ADAM is to monitor trends in substance abuse and help inform local agency operations and policies. Therefore, timely data processing and dissemination are important goals.
Under the current design, it has taken an average of 49 calendar days for each site to deliver its population data to central processing after completion of quarterly data collection in 2002, with appreciable variation by site, from a minimum of 1 day to nearly 5 months. This average fell to about 40 days by the second quarter of 2003, and it does not seem possible to reduce this interval any further. Moreover, it has taken an average of 76 calendar days during 2002 from receipt of the initial population data to the point that both the population and questionnaire data are fully processed, cleaned, and ready to merge. This is the point at which we initiate the statistical estimation process, which itself is generally quite rapid (one to four days, depending on load factors), produce weighted data and post reports. This number fell to approximately 50 days by the second quarter of 2003, and this average also seems impossible to reduce further, in view of stubborn outliers and the continuing stream of changes in local personnel and procedures, with associated new problems to be solved, among the 35 ADAM sites.

Overall, then, completing the collection and processing of the old population data required an average of slightly more than 4 months from the end of the field period through the end of the estimation process. Based on eight quarters of experience, we have been able to reduce this by no more than another month.

Based on the parallel testing, we can deliver data appreciably faster under the new method. Under this estimation method, population data are to be collected on a daily basis during quarterly data collection. Therefore, the sites should be able to deliver the new population data to central processing at the same time as they deliver the questionnaire data, which now averages roughly 10 calendar days after completion of data collection. With the much reduced processing requirements (and reduced need for resubmission) of the new population data, we estimate that it will require an average of no more than 28 days to completely process the population, questionnaire, and urinalysis data from each site and begin to weight the data. We therefore estimate that the average time to produce final weighted data using the new census will be no more than 6 weeks, which is half of the minimum average time that can be achieved under the old methodology.

6. Concluding Remarks

With the accumulation of ADAM data, we have gained important insight into the structure of the arrestee population at each site and the properties of the post-sampling stratification estimator. The new estimation method represents the first fruit of integrating this information into a more efficient design that enhances the overall quality of ADAM data at reduced costs. The new method will be implemented in the fourth quarter of 2003.

A major limitation of the current ADAM methodology is associated with the population data and its role in estimation. ADAM produces site level estimates on a quarterly basis, but the so called population data are not true population data because (1) they are limited only to those arrestees who are booked into sampled facilities, and (2) they are limited to the two-week data collection period. Thus, the old estimation method does not account for the first stage sampling of jails under stratified cluster design, nor does it account for the entire quarterly population of arrestees within a site. This has at least three implications: (1) the current ADAM weights are not appropriate for estimating population totals beyond the sample facilities and the data collection period, (2) strong assumptions about between facility variances are needed to derive unbiased population estimates of means and proportions in cluster sites, and (3) the reported variances of population means or proportions for cluster sites are underestimated. Addressing this limitation requires expanding population data collection into all jails within a site throughout the quarter, which is prohibitively expensive, especially for those sites with very large numbers of jails. One less expensive alternative would be to collect population data from all jails at one point in time so the first stage sampling weights may be estimated. We have conducted limited but promising research to explore the possibility of using Uniform Crime Report (UCR) data to model and predict ADAM population data per facility.

Other aspects of the ADAM design that may be improved include the sample allocation among sites, sampling of female arrestees, nonresponse and noncoverage adjustments, and so on. In some ADAM sites, a probability sample of jails is not feasible given budgetary and operational constraints. Investigations on the between-jail variance structure with respect to important ADAM variables will be of great value.

REFERENCES

Cochran, W. G. (1977). *Sampling Techniques*. Third Edition. New York: John Wiley.

Hunt, D. and Rhodes, W. (2001). *Methodology Guide for ADAM*. Washington, D.C.: National Institute of Justice.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley.

Kish, L. (1992). “Weighting for Unequal P_i.” *Journal of Official Statistics*, Vol. 8. No. 2, 183-200.

Rhodes, W. et al. (2002). “Arrestee Drug Abuse Monitoring System: Sampling Design and Analysis,” Unpublished paper.

SAS Institute, Inc. (1999). SAS/STAT User's Guide, Version 8, Volumes 2, Cary, NC: SAS Institute, Inc.

Yang, M. and Gerstein, D. (2003). “Response to Reviews of ‘An Alternative Two-Stage ADAM Weighting Scheme,’” Report submitted to the National Institute of Justice.

Yang, M. and Gerstein, D. (2002). “An Alternative Two-Stage ADAM Weighting Scheme,” Report submitted to the National Institute of Justice.

Woodruff, R.S. (1971). “A Simple Method for Approximating the Variance of a Complicated Estimate,” *Journal of the American Statistical Association,* 66, 411-414.