

# Methodological Challenges in the Redesign of the 2015 Survey of Prison Inmates

Jennifer Unangst<sup>1</sup> and Marcus Berzofsky<sup>1</sup>,  
<sup>1</sup>RTI International, 3040 Cornwallis Rd. RTP, NC 27709

## Abstract

The Survey of Prison Inmates (SPI) is being redesigned for its intended implementation in 2015 by the Bureau of Justice Statistics (BJS). The goal of SPI is to provide estimates which describe the U.S. state and federal inmate populations across topics such as criminal history, mental and physical health status, and socioeconomic status at the national-level, as well as at the level of particular subdomains (e.g., gender) and jurisdictions (i.e., states with large prison populations such as California, Texas, and Florida). This paper discusses the methodologies and associated challenges encountered during the SPI redesign, including the determination of an optimal first-stage allocation scheme, utilization of both implicit and explicit stratification to select prison facilities across subdomains, and estimation of the required number of inmates to be sampled in order to achieve the desired precision. Using a redesigned instrument and historical experience from the 2004 SPI to approximate design effects, a simulation study was conducted to determine the proportion of outcomes that would meet the study's precision goals while minimizing respondent burden on inmates and facilities.

**Key Words:** Survey design

## 1. Introduction

### 1.1 Background

The Survey of Prison Inmates (SPI) is a periodic, omnibus survey of the national inmate population sponsored by the Bureau of Justice Statistics (BJS). The most recent iteration of SPI occurred in 2004 (United States Department of Justice, BJS 2004), and it is being redesigned for its next intended implementation in 2015. The survey covers a variety of topics, including inmate demographics, criminal justice history, mental and physical health status, as well as treatment services related to mental and physical health. Since each of these topics is of equal importance, the survey design for the 2015 SPI must be robust to achieve reasonable precision given a fixed cost for each of the estimate types.

The target population for SPI is all inmates residing in a prison. Inmates in the target population are housed within one of two jurisdiction types: (1) *state facility* (i.e., a prison operated by or for a state Department of Corrections) and (2) *federal facility* (i.e., a facility operated by the Federal Bureau of Prisons). Furthermore, the target population is housed in one of two facility types: (1) *confinement-based prisons* (i.e., a prison in which less than half of the inmate population is permitted to leave the facility unaccompanied) and (2) *community-based prisons* (i.e., a prison in which over half the inmates are permitted to leave unaccompanied). Within the target population, the 2015 SPI has

several inmate subpopulations for which inference is desired. These subpopulations include:

- Male inmate housed in state facilities
- Female inmates housed in state facilities
- Inmates housed in federal facilities
- Inmates housed in state jurisdictions with over 100,000 inmates

Table 1 presents the distribution of facilities and inmates in the target population based on the 2012 Census of State and Federal Adult Correctional Facilities. In total, the target population consists of 1,399,008 male inmates and 103,578 female inmates housed in 2,001 prisons.

**Table 1:** Number of Facilities and Inmates Held by Jurisdiction Type and Gender Housed

<i>Jurisdiction Type</i>	<i>Gender Housed</i>	<i>Number of Facilities</i>	<i>Number of Inmates Held</i>	
			<i>Male</i>	<i>Female</i>
State	Male only	1,362	1,173,144	0
	Female only	191	0	80,797
	Both	257	47,151	9,997
Federal	Male only	171	177,409	0
	Female only	19	0	12,731
	Both	1	1,304	53
Total		2,001	1,399,008	103,578

## 1.2 Study Objectives

There are three main objectives for the SPI: (1) to maximize precision of the estimates, (2) to minimize burden (i.e., the total time facilities and inmates will be asked to participate in the survey), and (3) to minimize cost. Since some of these objectives are competing (e.g., improving precision of the estimates usually increases cost and, most likely, burden), the overarching goal is to optimize across all three objectives.

In general, there are two ways to maximize precision: (1) increase the sample size, and (2) minimize design effects. Often, implementing either of these approaches to maximizing precision will increase cost and increase burden. For SPI, costs were fixed given BJS budget constraints. Therefore, our study objective was to develop a sampling plan that would maximize precision while ensuring burden was minimized within known budget constraints.

Given this objective, BJS required several additional constraints that needed to be achieved in the survey design. Namely,

- Precision goals were set at a coefficient of variation (CV) less than or equal to the CV achieved in the 2004 study or 0.10 for state male inmates or 0.15 for all other subpopulations of interest (i.e., female, state inmates, federal inmates, and jurisdictions with 100,000 or more inmates). The CV is defined as the ratio of the standard deviation to the prevalence estimate.
- The within-facility sample size (i.e., facility level burden) cannot exceed 75% of the total inmate population (including nonrespondents).
- The number of confinement facilities selected in any state cannot exceed 30% of the total number of confinement facilities in the state.

This paper discusses the methodological considerations taken to develop the best, if not optimal, design for the 2015 SPI.

## 2. Methods

In order to assess different options in a systematic manner, a three step process was used. First, a set of design assumptions which would be applied to all potential designs was developed. Second, the set of characteristics for which we wanted to vary the study design and the range for which they would be varied were determined. Third, we simulated each potential design and assessed the quality of the design in terms of the proportion of key estimates that had a CV below the desired amount.

### 2.1 Design Assumptions

Several assumptions were made across all design options considered for the 2015 design. These assumptions were either (1) necessary to meet one of BJS's criteria, or (2) from a statistical perspective, provided the smallest variance. The design features described in the following subsections were used in all design options under consideration for the 2015 SPI.

#### 2.1.1 Frame Creation

The frame used for our analysis was a modified version of the 2012 Census of State and Federal Adult Correctional Facilities. The modifications made were to account for projected future changes in the prison population. These changes included the following.

- Removing facilities that were expected to close prior to the start of data collection in 2015;
- Adding new facilities that were known to be operating as of May 2014 (after the Census was completed);
- Increasing or decreasing the inmate population in a facility based on facility indicated growth; and
- Altering the gender housed assignment (e.g., reclassifying some male only facilities to be both genders housed) based on changes indicated by the facility.

For each facility, the frame contained the following pieces of key information for the sample design.

1. The state in which the facility is located
2. The number of male inmates housed (male population count)
3. The number of female inmates housed (female population count)
4. The jurisdiction type (state or federal)
5. The facility type (confinement-based or community-based)
6. Whether the facility provided mental health or medical services

Using this modified Census, a separate frame was created for male inmates and female inmates. The male inmate frame consisted of all facilities that only house male inmates as well as the male portion (i.e., the male population count) of facilities that house both genders. The female inmate frame consisted of all facilities that only house female inmates as well as the female portion (i.e., the female population count) of all facilities that house both genders.

Furthermore, based on the modified frame, there were three state jurisdictions housing 100,000 inmates or more. These state jurisdictions include Texas, California, and Florida.

### *2.1.2 Selection Method*

A two-stage design was used, in which the facilities were selected in the first stage and inmates within the sampled facilities were selected in the second stage. The sample of facilities was selected using a sequential probability proportional to size measure (PPS) method. The composite size measure used to select the PPS sample was defined as the number of inmates housed in the facility adjusted by an oversampling factor as appropriate. In the second stage, within a stratum, a fixed number of inmates was selected per facility.

Based on our study objectives, this type of two-stage design was implemented across all potential designs for two reasons:

1. *Self-weighting*. Within a stratum, inmates will be self-weighting. This means that all selected inmates in a stratum will have the same design-based weight regardless of the size of the facility from which the inmate is selected. This minimizes the design effect caused by unequal weights.
2. *Equal workload*. By selecting the same number of inmates in facilities, the interviewer workload across facilities will be equal. This approach should reduce data collection costs which may be caused by needing to staff interviewing teams that are constantly changing size.

### *2.1.3 Stratification*

Explicit stratification (i.e., strata for which a fixed sample size is allocated) was used to allocate the total facility sample size across 10 explicit strata. Within each of the gender specific frames, the facility population was split into 5 strata:

- Federal prisons,
- Texas state prisons,
- California state prisons,
- Florida state prisons, and
- Prisons in the remaining 47 states.

Furthermore, implicit stratification (i.e., the sorting of the frame by particular facility characteristics to group them in a stratum-like fashion prior to sample selection) was also used to ensure sufficient representation of several important facility characteristics (e.g., whether a facility is a confinement or community type facility, whether a facility offers mental health or medical services, census region, and state jurisdiction).

### *2.1.4 Oversampled Populations*

Several minority inmate populations were oversampled to ensure reasonable representation and precision for key subpopulations. These include female inmates, inmates housed in facilities providing mental health or medical services, and non-drug offenders within federal facilities. For inmates housed in facilities providing mental health or medical services, oversampling was controlled by applying an oversampling factor to the number of inmates housed in the facility. The resulting product (i.e., the composite size measure) was used to select the PPS sample in each analytic stratum.

Oversampling was an important feature of all designs because (1) the population was of analytic interest to BJS (e.g., female inmates), (2) inmates were more likely to have an

outcome of interest to BJS (e.g., a mental health issue), or (3) the population was so small that not oversampling may, by chance, lead to its exclusion in the sample (e.g., non-drug offenders in federal facilities). Therefore, it was determined that all designs would have some level of oversampling of these populations.

### *2.1.5 Burden Restriction on Inmate Sample Size*

A limit on the second-stage inmate sample size was established to restrict the burden imposed upon smaller facilities. Based on prior BJS experience with facility-based inmate surveys, it was anticipated that if a facility was overburdened, the facility may decline to participate in the study. Therefore, no more than 75% of a facility's population will be selected for the 2015 SPI. While this restriction addresses burden concerns, it also further introduces the possibility of unequal weighting within a stratum. For example, while still self-weighting within a facility, inmates in the smaller facilities within a stratum will have different design-based weights than the larger facilities.

## **2.2 Design Characteristics under Consideration**

Given the common assumptions across all design options, there were several design features that were not fixed which would impact precision of the estimates. Because of the large number of analytic objectives that BJS is interested in for SPI (e.g., analyses by gender, among inmates with mental health issues, etc.), it was not possible for a single optimization to be conducted and account for all of the design criteria (i.e., precision and burden) for all analytic objectives. Therefore, some design characteristics were allowed to vary in order to assess the impact each has on the array of estimates BJS is interested in analyzing.

The design characteristics included (1) the number of facilities expected to participate, (2) the within-facility sample size, (3) the magnitude of the oversampling factor used for female inmates, (4) the magnitude of the oversampling factor used for inmates housed in facilities that provide mental health or medical services, and (5) the nominal within-facility (i.e., inmate) sample size. Furthermore, not all designs included the BJS criterion that no state has more than 30% of its confinement facilities selected. This was done to assess the impact that this additional restriction has on the jurisdiction level precision of these states.

As stated earlier, the basic budget available for SPI was known prior to this analysis. The largest driver in the cost of the survey is the number of facilities that will participate. This is because it costs more money to transport and house an interviewing team at a new facility than increasing the within-facility sample size at a facility. However, due to an increase in the design effect due to clustering (see Section 2.4.3), precision is better when more PSUs (primary sampling units) are selected with fewer SSUs (secondary sampling units) within selected PSUs. In order to assess the impact of this balance between first-stage and second-stage sample sizes, the number of facilities expected to participate was allowed to vary to levels higher than the budget could afford.

In total, 25 design options were considered for the 2015 SPI based on varying combinations of these characteristics. Table 2 summarizes the range of values considered across the 25 design options.

**Table 2:** Range and Number of Values for Design Options

<i>Design Characteristic</i>	<i>Range of Values Considered</i>	<i>Number of Values</i>
Number of facilities expected to participate	350 to 400	5
Oversampling factor for facilities housing female inmates	2.4 to 5	6
Oversampling factor for facilities offering mental health/medical services	3 to 4	3
Inmate sample size in state facilities	64	1
Inmate sample size in federal facilities	64 to 120	5
Additional burden restrictions	Restricting sampled confinement facilities in TX, CA, and FL to 30% of the confinement facilities on the frame	N/A

### 2.3 Simulation Study

A simulation study was performed to assess the impact of varying the undetermined design characteristics on the study's precision goals. In particular, the simulation provided insight on the degree of variability induced into the second-stage inmate sample size due to the restriction of selecting no more than 75% of a facility's inmate population. Without this restriction, the precision for the design could have been directly determined without a simulation. However, with this restriction, the number of facilities affected by the limit on the second-stage inmate sample size may vary for each realization of the sample design. Therefore, a simulation was necessary to determine the total expected number of inmates to participate in SPI and the variance around that expected number.

For each design option, 1,000 samples were simulated from the SPI facility frame which is based on the modified 2012 Census of State and Federal Adult Correctional Facilities. For each sampling stratum and each design option, the average number of participating inmates per facility ( $\bar{m}$ ) and its 95% confidence interval were determined across the 1,000 simulated samples. Similarly, an expected nominal inmate sample size ( $m_{h+}$ ) for each analytic stratum and its 95% confidence interval were also determined.

### 2.4 Estimating Precision

Formula (1) below (Kish 1995) defines the variance for a prevalence estimate obtained under a complex survey design. There are two design effects used in the estimation of standard errors based on this formula – the unequal weighting effect (UWE) and the design effect due to clustering (DEFF<sub>c</sub>). Additionally, the prevalence estimate and the nominal sample size are required to estimate the variance under simple random sample (SRS) assumptions. Each of these components of the variance formula is unknown prior to sampling. For each simulated design option, the components of the variance formula were estimated based on the 1,000 simulated samples and assumptions derived from the 2004 SPI. Details of the methods used to estimate each component are provided in the following subsections. These general processes were done for each analysis stratum.

$$\text{Var}(\hat{p})_{design} = UWE \times DEFF_C \times \text{Var}(\hat{p})_{srs} \quad (1)$$

#### 2.4.1 Estimation Method for Prevalence of Key Outcomes

Based on the 2004 SPI instrument, 68 of the study's key outcomes were identified. These outcomes were selected because they represented a variety of topics including inmate demographics, criminal justice history, and mental and physical health statuses. The parameter for each key outcome for the 2015 SPI ( $p$ ) was estimated using the parameter estimate from the 2004 SPI (i.e., for assessment of the precision, the analysis fixed the point estimates at 2004 levels). Parameter estimates were computed at the 2015 SPI analytic stratum level.

#### 2.4.2 Estimation Method for the Unequal Weighting Effect

Sampling weights were estimated for each of the expected participating inmates in each of the 1,000 simulated samples. The sampling weights assumed a constant rate of nonresponse both across facilities within a stratum and between strata. In other words, for this analysis, no UWE due to nonresponse was incorporated. A UWE was then calculated for the simulated sample, and a final arithmetic average UWE was obtained across all 1,000 simulated samples.

#### 2.4.3 Estimation Method for the Design Effect due to Clustering

Let  $ICC$  denote the intra class correlation coefficient, and let  $\bar{m}$  denote the average inmate sample size. Then, the design effect due to clustering can be written as a function of these variables (Kish 1995).

$$DEFF_C = 1 + ICC(\bar{m} - 1) \quad (2)$$

The average inmate sample size,  $\bar{m}$ , was estimated by the mean inmate sample size across the 1,000 simulated samples. Because  $\bar{m}$  varied across samples, an upper and lower bound value was estimated based on its variance across the 1,000 simulated samples. The  $ICC$  for the 2015 SPI was assumed to be the observed  $ICC$  value from the 2004 SPI.

#### 2.4.4 Estimation Method for the Variance under SRS Assumptions

For prevalence estimate  $\hat{p}$  and nominal inmate sample size within stratum  $h$ ,  $m_{h+}$ , the variance under SRS assumptions can be written using the following traditional formula.

$$\text{Var}(\hat{p})_{srs} = \frac{\hat{p}(1-\hat{p})}{m_{h+}} \quad (3)$$

The nominal inmate sample size,  $m_{h+}$ , for each stratum  $h$ , was computed by calculating an estimate for the nominal inmate sample size for each simulated sample and then averaging the simulated sample sizes across all 1,000 simulations.

#### 2.4.5 Estimation Method for the 2015 CV

After estimating each component of the variance formula, it was possible to estimate a standard error for each of the 68 key outcome estimates. An estimated CV for the 2015 SPI was then obtained for each estimate. Due to the variability in the within-facility sample size ( $\bar{m}$ ), a lower bound, mean, and upper bound CV was calculated using the

upper and lower bound average within facility sample sizes. This process was completed for each analysis stratum.

### 3. Results

The 25 simulated design options were assessed in three ways: (1) by comparing the precision estimates derived for each design to the study's precision goals, (2) by comparing the characteristics of the facility sample to an expected distribution based on the frame, and (3) by comparing the specifications of the design option to the specifications of an "ideal" design. Each of these assessment methods is described in more detail in the following subsections.

#### 3.1 Assessment of Achieving Precision Goals

As described in Section 2.4, a lower bound, mean, and upper bound CV estimate were calculated for each parameter estimate in each analytic stratum. Estimated CV values less than or equal to the desired CV were flagged. Since the CV estimates were based on several assumptions, we additionally reviewed the CV estimates under a relaxed set of precision goals, where estimates within a specified threshold of the desired CV were also flagged. For state male inmates, CV estimates which were within 1% of the desired CV level were flagged, while for all other analysis strata, CV estimates which were within 10% of the desired CV level were flagged. A more relaxed criterion was established for the non-state male analysis strata due to the understanding that these strata would have smaller sample sizes and consequently have worse precision in all scenarios. This relaxed threshold identified estimates with a CV within the substantive equivalent of the desired CV, but just slightly above. Given the uncertainty of some of the assumptions made, these estimates were considered acceptably precise under the relaxed threshold.

Table 3 presents the percentage of estimates by estimate type (i.e., demographics, criminal history, physical health, and mental health) which were less than or equal to the stratum's desired CV level for one of the simulated designs. The number of key outcome estimates belonging to the estimate type is provided in parentheses below each type.

**Table 3:** Percentage of Mean Estimates Below Desired CV Threshold

<i>Stratum</i>	<i>Demographics</i> (22)	<i>Criminal History</i> (12)	<i>Physical Health</i> (10)	<i>Mental Health</i> (24)
State Male	72.7	41.7	100.0	83.3
State Female	18.2	33.3	20.0	4.2
Federal	86.4	58.3	80.0	87.5
Texas	81.8	100.0	80.0	79.2
California	13.6	25.0	10.0	41.2
Florida	63.6	75.0	80.0	45.8

Table 4 presents the percentage of estimates by estimate type (i.e., demographics, criminal history, physical health, and mental health) which were less than or equal to the stratum's *relaxed* desired CV level for one of the simulated designs. The number of key outcome estimates belonging to the estimate type is provided in parentheses below each type. Table 4 summarizes the same simulated design option presented in Table 3.



**Table 4:** Percentage of Mean Estimates Below *Relaxed* Desired CV Threshold

<i>Stratum</i>	<i>Demographics</i> (22)	<i>Criminal History</i> (12)	<i>Physical Health</i> (10)	<i>Mental Health</i> (24)
State Male	86.4	66.7	100.0	91.7
State Female	77.3	83.3	80.0	29.2
Federal	90.9	66.7	80.0	91.7
Texas	81.8	100.0	80.0	79.2
California	40.9	33.3	40.0	58.3
Florida	68.2	83.3	80.0	50.0

The percentage of estimates meeting precision goals ranged depending on the specifications of each design option and thus illustrated the pros and cons of each design. The percentages presented in Tables 3 and 4 are based on a design option with a state female oversampling factor of 3.5. As shown in Table 4, this particular design option resulted in a range of 29.2% to 83.3% of mental health and criminal history outcomes, respectively, of the state female estimates that achieved the relaxed CV threshold. Increasing the state female oversampling factor to 4.0 in an otherwise comparable design improved precision in the state female stratum and resulted in a range of 77.3% to 100% of the state female estimates meeting the relaxed CV threshold. However, there was also a corresponding negative impact on the precision observed in the state male stratum. More specifically, increasing the state female oversampling factor reduced state male estimates achieving the relaxed CV threshold down to 16.7% from 66.7% (as seen in Table 4) for criminal history outcomes.

The relatively low percentages of mean estimates meeting the CV threshold in strata like California and Florida are a result of imposing a burden restriction on the number of confinement facilities selected in each of these strata – namely, no more than 30% of the confinement facilities on the frame would be selected in those states. However, it was anticipated that increasing the number of sampled facilities in these states might cause the states to decline participation in the study. Thus, from a practical perspective, the precision could not be significantly improved in these states. In other words, for states with large prison populations, the need to minimize burden outweighed the goal to maximize precision due to the practical consideration that some participation by these states was better than none.

The pros and cons of each design were reviewed with BJS to determine which strata precision goals should be prioritized, with additional design options incorporated based on BJS feedback.

### 3.2 Assessment of Sampling Properties

The properties of each design option were assessed by comparing distributions of the expected number of selected facilities to the average number of facilities selected across the 1,000 simulated samples by key sampling characteristics. These characteristics included the state in which the facility is located, gender of the inmates housed in the facility, size of the facility, census region in which the facility is located, facility type (i.e., confinement or community), and whether the facility provides mental health and medical services.

Based on this assessment it became evident, for example, that facilities with fewer than 500 inmates were selected more frequently than expected in nearly all of the design options. This was likely due to the explicit stratification of the facility sample by gender, since the facilities on the frame which house female inmates were typically smaller than the facilities which house males. Table 5 provides an example of the expected versus observed sampling distributions for facility size. Typical simulated designs illustrated a similar average distribution of selected facilities. The other key sampling characteristics were close to the expected distributions in most cases.

**Table 5:** Distribution of Facility Size by the Number of Expected and Average Observed Selected Facilities

<i>Sampled Facility Size</i>	<i>Expected Sample Count</i>	<i>Average Observed Sample Count</i>
Less than 50 inmates	4.5	6.2
51-100 inmates	4.9	8.5
101-250 inmates	13.3	14.7
251-500 inmates	27.1	30.6
501-750 inmates	33.5	33.6
751-1,000 inmates	30.4	33.6
1,001-1,500 inmates	90.6	86.4
1,501-2,000 inmates	55.3	53.6
More than 2,001 inmates	90.5	82.9

### 3.3 Comparison to the “Ideal” Design

Rearranging Formula (1) allows the nominal inmate sample size in stratum  $h$ ,  $m_{h+}$ , to be written as a function of the overall DEFF (i.e., the product of the  $DEFF_c$  and the UWE), the prevalence, and the CV. By fixing several design characteristics used in the formula, it is then possible to empirically determine the “ideal” design necessary to achieve the desired CV.

$$m_{h+} = \frac{DEFF \times p(1-p)}{(pCV)^2} \quad (4)$$

The following assumptions were made in order to estimate the nominal inmate sample size required to achieve the desired CV for each analysis stratum’s estimates.

- The CV equaled the minimum desired CV.
- The prevalence was the estimate obtained in the 2004 SPI.
- The design effects were fixed based on the average inmate sample size selected.

The “ideal” facility sample size was then determined for each analysis stratum’s estimates by dividing the estimated “ideal” nominal inmate sample size by the average number of inmates sampled per facility ( $\bar{m}$ ). The facility sample sizes necessary to meet the precision goals for various percentages (e.g., 85%) of the estimates in each analysis stratum were reviewed and compared to the design options. The design which met the precision goals for 100% of the estimates was not considered due to the presence of outlier estimate types which would require an unrealistic number of facilities to meet the precision goals.

This exercise illustrated that achieving precision goals in all six analysis strata was unlikely due to burden and cost constraints. Table 6 shows the number of facilities

required to achieve the 70<sup>th</sup>, 75<sup>th</sup>, 85<sup>th</sup>, 90<sup>th</sup>, and 95<sup>th</sup> percentiles of estimates meeting the desired CV for each stratum. The facility allocation for one of the simulated design options is also included in Table 6 for comparison. The total sample size required in order to obtain the 70<sup>th</sup> percentile of estimates achieving the desired CV in every stratum is 361, while the design option only allocates 350 total facilities. Based on Table 6, it would then be expected that this particular design option would demonstrate lower precision in the state female, California, and Florida strata due to smaller sample sizes in the strata relative to the “ideal” design.

**Table 6:** Number of Participating Facilities Required to Achieve the Desired CV

	<i>Ideal Facility Sample Size by Percentile of Estimates Achieving the Desired CV</i>					<i>Design Option</i>
	<i>70th</i>	<i>75th</i>	<i>85<sup>th</sup></i>	<i>90th</i>	<i>95th</i>	
State Male	222	225	229	231	234	225
State Female	97	98	141	201	348	74
Federal	42	43	59	78	108	51
Texas	27	27	40	50	67	29
California	25	25	32	36	52	19
Florida	23	25	32	46	81	18
Total	361	366	429	510	690	350

#### 4. Discussion

While the primary goal of the study was to identify a design which optimized precision and burden given a fixed cost, the competing nature of these factors prevented the identification of a design which clearly achieved all three aims. Consequently, two design options were recommended for the SPI 2015 design – (1) a cost considerate recommendation and (2) an optimal design recommendation which met as many of the burden and precision goals as best as possible while simultaneously minimized cost as best as possible.

The SPI 2015 budget anticipated 350 participating facilities. Therefore, the cost considerate recommendation was the design which had 350 participating facilities and performed best in the assessment of precision (347 out of 408 total CV estimates across all analysis strata achieved the relaxed desired CV). The optimal design recommendation was the option which provided the desired precision for as many estimates as possible and had 367 participating facilities (365 out of 408 total CV estimates across all analysis strata achieved the relaxed desired CV).

The results of the study supported the notion that formal precision goals must be balanced by practical considerations such as cost and burden. The design which leads to the best set of precision estimates may not be feasible in reality due to burden or cost constraints. However, it is possible to achieve a design that meets a majority of BJS’s objectives under their known cost constraints.

### **Acknowledgements**

The authors would like to thank BJS for sponsoring this research and comments provided by Lauren Glaze. However, we would like to note that the views expressed in this paper are those of the authors only and do not reflect the views or position of BJS or the Department of Justice.

### **References**

Kish, Leslie (1995). *Survey Sampling*. New York, NY: John Wiley & Sons, Inc.

United States Department of Justice, Bureau of Justice Statistics. Survey of Inmates in State and Federal Correctional Facilities, 2004 (ICPSR 04572). Retrieved from <http://doi.org/10.3886/ICPSR04572.v1>