December, 2006

Optimizing the Expected Overlap of Survey Samples via the Northwest Corner Rule

Lenka Mach
Philip T. Reiss, *Columbia University*
Ioana Schiopu-Kratina

Optimizing the Expected Overlap of Survey Samples via the Northwest Corner Rule

Lenka MACH, Philip T. REISS, and Ioana ŞCHIOPU-KRATINA

In survey sampling there is often a need to coordinate the selection of pairs of samples drawn from two overlapping populations so as to maximize or minimize their expected overlap, subject to constraints on the marginal probabilities determined by the respective designs. For instance, maximizing the expected overlap between repeated samples can stabilize the resulting estimates of change and reduce the costs of first contacts; minimizing the expected overlap can avoid overburdening respondents with multiple surveys. We focus on the important special case in which both samples are selected by simple random sampling without replacement (SRSWOR) conducted independently within each stratum. Optimizing the expected sample overlap can be formulated as a linear programming problem known as a transportation problem (TP). We show that by appropriately grouping and ordering the possible samples in each survey, one can reduce the initial TP to a much smaller TP amenable to solution by an algorithm known as the Northwest Corner Rule (NWCR). The proposed NWCR method proceeds in two easily implemented steps: first selecting the numbers of births (new units) and deaths (deleted units) by a random selection from a hypergeometric distribution, and then selecting the births and deaths by SRSWOR. We formally prove properties of the NWCR solutions, including a minimal variance property of the minimal overlap solution. In a simulation study, the NWCR method compares favorably with a popular method based on assignment of permanent random numbers to each sampling unit.

KEY WORDS: Integration of surveys; Monge property; Sample coordination; Simple random sampling; Transportation problem.

1. INTRODUCTION

1.1 General Formulation of the Problem and the NWCR Solution

When conducting multiple sample surveys on overlapping populations, it is often important to control the number of units that the selected samples have in common. Such coordination of samples can facilitate data collection and reduce response burden, which reduces costs and ensures good quality of data at the source. The size of the overlap of samples affects the precision of estimates of change between two occasions. Coordination of samples has come to encompass a panoply of techniques applied in diverse settings (see McKenzie and Gross 2000; Ohlsson 2000; Royce 2000); here we aim more specifically to *integrate* two surveys (Mitra and Pathak 1984), or find a joint probability distribution whose marginal probabilities equal the original design probabilities.

Given a joint probability of selection for two surveys, we say that two samples are *coordinated* if their joint probability of selection is not equal to the product of the respective probabilities of selection in each survey (Cotton and Hesse 1992). Samples are *positively* (*negatively*) coordinated if the expected overlap in this joint distribution is greater (smaller) than it would be if the samples were selected independently.

We consider two stratified designs, with simple random sampling without replacement (SRSWOR) conducted independently in each stratum. Such designs are often used for establishment surveys. We seek to *maximize* or *minimize* the expected sample overlap under a joint (integrating) probability. Because the expected overlap is a linear function in the unknown joint probabilities and the constraints are linear [see (2) in Sec. 2.2], this can be viewed as a linear programming (LP) problem, with the expected overlap playing the role of the objective function. This LP problem is a particular instance of a transportation problem (TP) (Raj 1956; Causey, Cox, and Ernst 1985).

One technical obstacle in solving the LP problem is the large number of variables involved: as many as there are possible pairs of samples. To surmount this obstacle, we first reduce the problem to that of optimizing an objective function defined at the level of groups of samples. Next, because the coefficients of this objective function have the appropriate *Monge property* (see Definitions 1 and 2 in Sec. 3.2), we can obtain an optimal solution using the Northwest Corner Rule (NWCR), an algorithm well known in optimization theory. Interestingly, the minimization and maximization problems give rise to quite different objective functions and hence are not dual to each other.

The NWCR method applies to large sample sizes and is easy to implement. On the other hand, the only population updates that it can handle are *deaths* (units removed from the population) and *births* (units added to the population).

1.2 Motivation for Controlling Sample Overlap

Sample coordination can facilitate two key stages of a survey: data collection and estimation. Applications include the following:

1. *Reducing interviewing costs.* Raj (1956) considered the problem of carrying out two different surveys on a geographically dispersed population. To minimize the interviewers' cost of transportation, a maximal number of locales had to be visited once to gather information for both surveys. Survey takers may also wish to maximize the overlap of primary sampling units (PSUs), geographic areas assigned to interviewers, to reduce the cost of hiring interviewers. Maximizing the overlap of samples of ultimate units reduces the cost of initiation interviews (Ernst 1999).

Lenka Mach is Senior Methodologist, SSMD, Statistics Canada, Ottawa, Ontario K1A 0T6, Canada (E-mail: *lenka.mach@statcan.ca*). Philip T. Reiss is Associate Research Scientist, New York University Child Study Center, New York, NY 10016 (E-mail: *reissp01@med.nyu.edu*). (Work was done while the author was a Graduate Student, Department of Biostatistics, Columbia University, New York, NY 10032.) Ioana Şchiopu-Kratina is Senior Methodologist, HSMD, Statistics Canada, Ottawa, Ontario K1A 0T6 (E-mail: *ioana.schiopu-kratina@statcan.ca*). The authors thank the referees for useful comments and invaluable suggestions for improving the original manuscript. They also thank P. Lavallée and J. N. K. Rao for valuable comments on an earlier version of the manuscript. I. Şchiopu-Kratina acknowledges the support of the Statistics Canada Sabbatical Research Programme.

Journal of the American Statistical Association
December 2006, Vol. 101, No. 476, Theory and Methods
DOI 10.1198/016214506000000320

2. *Reducing response burden.* Being selected for multiple surveys may represent a burden, leading to nonresponse and diminished data quality; we would then wish to minimize the expected number of individuals asked to respond to several surveys.
3. *Frame updates.* Survey takers often need to update the frame of establishment surveys for births and deaths, or when a large number of units change strata. Retaining the maximum number of units from the previous sample will reduce the cost of first contacts and increase the precision of the estimates of change.

1.3 Some Previous Work on Overlap

There is a vast literature on coordination of samples, starting with the pioneering work of Keyfitz (1951). We now briefly review some of this work and discuss some of its limitations. More complete discussions have been given by Ernst (1999) and Ohlsson (1995).

Following Ernst (1999), we classify overlap methods as designed for sequential or simultaneous selection. Sequential methods are adequate when a survey is taken on two different occasions, whereas simultaneous methods coordinate samples selected at the same time for several surveys. Procedures for sequential selection can be adapted for simultaneous selection, but procedures designed specifically for simultaneous selection usually yield a better overlap. For two successive redesigns, sequential methods lacking the independence property in point 7 of Ernst (1999) are not recommended.

Keyfitz's sequential method with samples of size 1 in each stratum attains the maximum expected overlap. The first extension to larger sample sizes is due to Fellegi (1966), who suggested two procedures that are not optimal. The optimal procedure of Brewer, Early, and Joyce (1972) applies to surveys with large sample sizes but does not guarantee a fixed sample size. Ernst and Paben (2002) proposed an optimal procedure that has many desirable properties but seems somewhat complex to implement.

Raj (1956) introduced LP methods for sample coordination. For small sample sizes (e.g., when selecting PSUs), LP methods can be used to integrate surveys and optimize the expected overlap (e.g., Causey et al. 1985). For the selection of ultimate units, LP procedures are difficult to apply because of the large number of variables. Aragon and Pathak (1990) and Ernst and Ikeda (1994) have attempted to simplify this problem. The latter authors' method substantially reduces the size of the problem but is not optimal. In general, LP procedures for large sample sizes preserve only first-order inclusion probabilities.

Due to their ease of implementation, methods that use *permanent random numbers* (PRNs) (Ohlsson 1995; Rivière 2001) have become quite popular. Although some PRN methods are not optimal (see Sec. 6), or do not guarantee a fixed sample size, they are very flexible and can accommodate changes in the composition of the strata from one occasion to the next.

In summary, and as pointed out by a referee, methods for maximizing or minimizing the expected overlap of two samples tend to suffer from at least one of the following five limitations:

1. The procedure does not generally yield the optimal overlap.
2. The procedure is either impossible or impractical to use for surveys with large sample sizes.
3. Either the procedure is applicable only to simultaneous selection or, if it is also applicable to sequential selection, it does not yield the optimal overlap for simultaneous selection.
4. The sample size in at least one of the designs is variable.
5. The method applies to a restrictive class of surveys.

The NWCR method is one of the few methods (see also Ernst 1998) that have none of the first four limitations, but it does suffer from the fifth. It essentially applies to two surveys with SRSWOR selection in each stratum (the most commonly used design for establishment surveys), but cannot be used as such when many units have changed strata or for more than two surveys.

1.4 Organization of This Article

The remainder of the article is organized as follows. Section 2 presents the mathematical setup. Section 3 introduces the NWCR and some relevant definitions, and Sections 4 and 5 apply these ideas to maximizing and minimizing the overlap of two surveys, with illustrative examples. Section 6 presents a simulation study comparing the NWCR method and a PRN method known as *sequential SRSWOR*. Section 7 summarizes the advantages, the limitations, and some possible extensions of the NWCR method. An Appendix provides all proofs.

2. A TRANSPORTATION PROBLEM FOR TWO SURVEYS

2.1 Problem Setup

We consider two stratified SRSWOR designs on two overlapping finite populations. This covers the situation of two distinct surveys as well as two different selections for the same survey when the population has been updated for deaths and births. We assume that the definition of the strata is the same in both surveys, so we may confine our attention to the population of a single stratum, which is subject to updates. We have N units in the stratum initially and N' units after the updates. If D denotes the number of deaths, B denotes the number of births, and C denotes the number of units that belong to both populations, then $N' = N - D + B = C + B$. For the first (second) survey, n (n') units are selected by SRSWOR. We let S (S') denote the set of all possible samples in the first (second) survey. The overlap of the samples $s \in S$ and $s' \in S'$, denoted by $o(s, s')$, is the number of units that s and s' have in common. We seek a joint distribution for all pairs of samples (s, s') that will maximize or minimize their expected overlap, given the marginal distributions in each of the two surveys.

2.2 Initial Transportation Problem Representation

To formalize this integration-of-surveys problem, let $P = (p(s))_{s\in S}$ and $Q = (q(s'))_{s'\in S'}$ define the probability distributions on the set of samples for the two surveys. If $s_1, \dots, s_K$ is an ordering of the samples in S, and $s'_1, \dots, s'_L$ is an ordering of the samples in S', then finding the optimum expected overlap amounts to finding the maximum or minimum, over all

$\mathbf{X} = (x_{ij})_{1\le i\le K, 1\le j\le L}$, of

$$E_{\mathbf{X}}(o) = \sum_{i=1}^{K}\sum_{j=1}^{L} o(s_i, s'_j)x_{ij}, \qquad (1)$$

subject to

$$x_{ij} \ge 0, \qquad \sum_{j=1}^{L} x_{ij} = p(s_i), \qquad 1 \le i \le K,$$

$$\sum_{i=1}^{K} x_{ij} = q(s'_j), \qquad 1 \le j \le L, \qquad (2)$$

$$\sum_{i=1}^{K} p(s_i) = \sum_{j=1}^{L} q(s'_j) = 1.$$

The last constraint is automatically satisfied because P and Q are probability distributions. Under our assumption of two SRSWOR designs, we have $p(s) \equiv \binom{N}{n}^{-1}$, $q(s') \equiv \binom{N'}{n'}^{-1}$, $K = \binom{N}{n}$, and $L = \binom{N'}{n'}$. This constrained optimization is a TP that in general has too many variables to be readily solved.

2.3 Reduced Transportation Problem Representation

The TP representation for optimizing the expected overlap is not new (see, e.g., Raj 1956; Mitra and Pathak 1984; Causey et al. 1985). Our contribution arises from grouping and ordering the rows and columns of $(x_{ij})_{1\le i\le K, 1\le j\le L}$—and of the associated *matrix of sample overlaps* $(o(s_i, s'_j))_{1\le i\le K, 1\le j\le L}$—in a way that replaces the initial TP with a much more tractable one. We first describe the grouping; we discuss the ordering and objective function, which depend on whether we are maximizing or minimizing the expected overlap, in Sections 4 and 5.

Let C be the set of units common to the two overlapping populations of the stratum. The grouping of rows (i.e., of samples s) is by number of units in $\mathsf{C} \cap s$, which we denote by $c = c(s)$. For the setup of Section 2.1, c has the hypergeometric distribution given by

$$p(c) = p(c; N, n, C) = \binom{C}{c}\binom{D}{n-c}\Big/\binom{N}{n}, \qquad (3)$$

with $\max\{n - D, 0\} \le c \le \min\{C, n\}$. Similarly, if we group the columns (samples s') by $c' = c'(s')$, the number of units in $\mathsf{C} \cap \mathrm{s}'$, then c' has the hypergeometric distribution

$$q(c') = q(c'; N', n', C) = \binom{C}{c'}\binom{B}{n'-c'}\Big/\binom{N'}{n'}, \qquad (4)$$

with $\max\{n' - B, 0\} \le c' \le \min\{C, n'\}$.

Having thus grouped the rows and columns into what we call *super-rows* and *super-columns*, which form a matrix of *blocks*, we can reduce the optimal overlap problem to a smaller TP, that of maximizing or minimizing, over all probability distributions π, the objective function

$$\sum_{c}\sum_{c'} \rho(c, c')\pi(c, c'), \qquad (5)$$

subject to

$$\sum_{c} \pi(c, c') = q(c'), \qquad \sum_{c'} \pi(c, c') = p(c). \qquad (6)$$

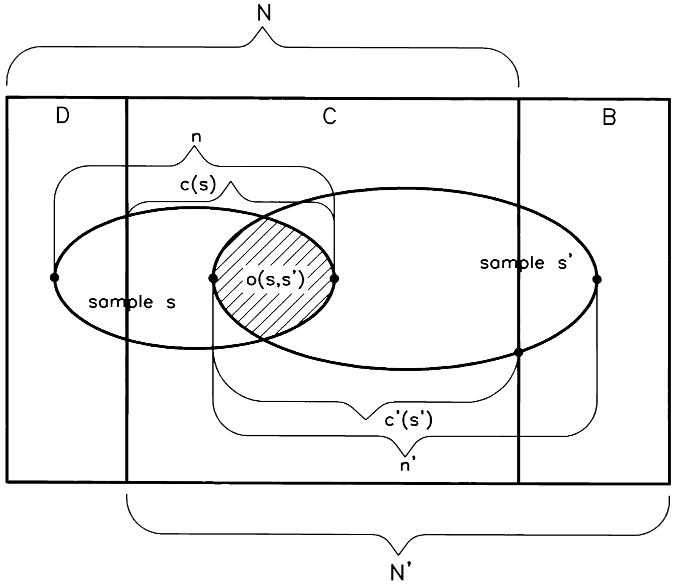

Figure 1. Key Quantities in the Optimal Overlap Problem.

Here $\pi(c, c')$ denotes the probability that the samples s and s' contain c and c' units from C, and $\rho(c, c')$ is a *block-optimum* function (defined later). In this way, we replace a transportation problem with $\binom{N}{n}\binom{N'}{n'}$ variables by one involving $(\min\{C, n\} - \max\{n - D, 0\} + 1)(\min\{C, n'\} - \max\{n' - B, 0\} + 1)$ variables. In the optimal solution, (5) will equal (1), but to prevent confusion, we sometimes refer to (1) as the *sample overlap* and to (5) as the *block overlap*. Figure 1 pictorially summarizes the definitions associated with the two designs.

3. THE NWCR METHOD FOR OPTIMIZING THE EXPECTED OVERLAP

Before defining the block-optimum function $\rho(c, c')$ we describe the NWCR algorithm in Section 3.1, define the Monge properties in Section 3.2, and argue that if, with a specific ordering of its rows and columns, the matrix of block optima $(\rho(c, c'))_{(c,c')}$ satisfies one of these properties, then the NWCR yields a solution to the reduced transportation problem of Section 2.3.

3.1 The NWCR Algorithm

Given orderings $c_1, \ldots, c_K$ of the super-rows and $c'_1, \ldots, c'_L$ of the super-columns, where now $K = \min\{C, n\} - \max\{n - D, 0\} + 1$ and $L = \min\{C, n'\} - \max\{n' - B, 0\} + 1$, we define $\pi_{ij} = \pi(c_i, c'_j)$, $p_i = p(c_i)$, and $q_j = q(c'_j)$ for $i = 1, \ldots, K$, $j = 1, \ldots, L$. The NWCR gives a greedy solution to (6) by allocating the maximum possible probability (mass) to each π_{ij}, given the marginals p_i and q_j and the mass assigned to previous blocks. Hoffman (1985) formulated the NWCR algorithm as follows: Set $\pi_{i0} = 0$ and $\pi_{0j} = 0$, and, if π_{rs} has been defined for all pairs (r, s), $r \le i$, $s \le j$, $(r, s) \ne (i, j)$, then

$$\pi_{ij} = \min\left\{p_i - \sum_{s=0}^{j-1}\pi_{is},\, q_j - \sum_{r=0}^{i-1}\pi_{rj}\right\}, \qquad i, j \ge 1.$$

The foregoing expression equals the as yet unassigned mass of that row or column, whichever is less. The following is a more

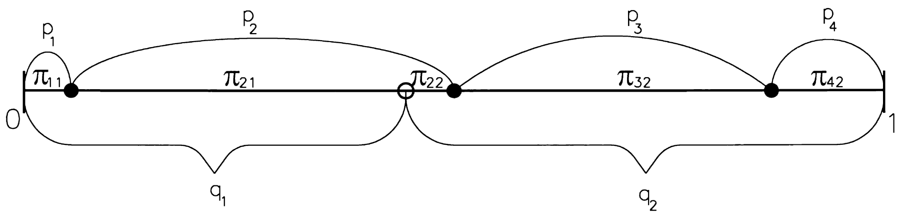

Figure 2. Pictorial Representation of an NWCR Solution With K = 4 and L = 2.

constructive algorithm for NWCR (see Arthanari and Dodge 1981, pp. 248–250, for some examples):

1. Define $\pi_{i0} = 0$ and $\pi_{0j} = 0$.
2. Set $i = j = 1$.
3. a. If $p_i - \sum_{s=0}^{j-1} \pi_{is} < q_j - \sum_{r=0}^{i-1} \pi_{rj}$ then $\pi_{ij} = p_i - \sum_{s=0}^{j-1} \pi_{is}$ and $\pi_{is} = 0$ for $s > j$; increment i by 1.
   b. If $p_i - \sum_{s=0}^{j-1} \pi_{is} > q_j - \sum_{r=0}^{i-1} \pi_{rj}$ then $\pi_{ij} = q_j - \sum_{r=0}^{i-1} \pi_{rj}$ and $\pi_{rj} = 0$ for $r > i$; increment j by 1.
   c. If $p_i - \sum_{s=0}^{j-1} \pi_{is} = q_j - \sum_{r=0}^{i-1} \pi_{rj}$ then set π_{ij} to their common value, $\pi_{is} = 0$ for $s > j$, and $\pi_{rj} = 0$ for $r > i$; increment i and j by 1.
4. If $i = K + 1$ or $j = L + 1$, stop; otherwise, return to step 3.

A straightforward noniterative implementation of the NWCR algorithm is illustrated in Figure 2. If the unit interval is partitioned into subintervals of length $p_1, \dots, p_K$ and again into subintervals of length $q_1, \dots, q_L$, then the joint probability π_{ij} is obtained as the length of the intersection of the subintervals of length p_i and q_j.

3.2 Monge Matrices and the NWCR

Our development requires the so-called Monge properties (Ross 1983).

Definition 1. The matrix $(\rho_{ij})_{1 \le i \le K, 1 \le j \le L}$ is *supermodular* (SM) if it satisfies

$$\rho_{ij} + \rho_{rs} \ge \rho_{is} + \rho_{rj} \quad \text{for all } 1 \le i < r \le K, 1 \le j < s \le L. \tag{7}$$

Definition 2. The matrix $(\rho_{ij})_{1 \le i \le K, 1 \le j \le L}$ is *submodular* (sM) if it satisfies

$$\rho_{ij} + \rho_{rs} \le \rho_{is} + \rho_{rj} \quad \text{for all } 1 \le i < r \le K, 1 \le j < s \le L. \tag{8}$$

Notation. Viewing $\mathbf{\Pi} = (\pi_{ij})_{i,j \ge 1}$ and $(\rho_{ij})_{i,j \ge 1}$ as vectors in $\mathbb{R}^{KL}$, we can write the objective function as an inner product, $\langle \mathbf{\Pi}, (\rho_{ij})_{i,j \ge 1} \rangle = \sum_{i,j} \rho_{ij} \pi_{ij}$. Alternatively, viewing $\mathbf{\Pi}$ as a joint probability distribution and given corresponding values $(\rho_{ij})_{i,j \ge 1}$ of a function ρ, we can define the expectation by $E_{\mathbf{\Pi}}(\rho) = \sum_{i,j} \rho_{ij} \pi_{ij}$ and the corresponding variance by $V_{\mathbf{\Pi}}(\rho) = E_{\mathbf{\Pi}}(\rho^2) - [E_{\mathbf{\Pi}}(\rho)]^2$. We shall use inner product or expectation notation as called for by the context. For $1 \le i < I \le K$ and $1 \le j < J \le L$, we define $\mathbf{\Delta}_{IJ}^{ij}$ as the $K \times L$ matrix with 1 as its (i, j) and (I, J) entries, -1 as its (i, J) and (I, j) entries, and 0 elsewhere. Clearly, $\langle \mathbf{\Delta}_{IJ}^{ij}, (\rho_{ij})_{i,j \ge 1} \rangle \ge 0$ (≤ 0) if $(\rho_{ij})_{i,j \ge 1}$ is SM (sM). We call a constant multiple of such a matrix an *elementary difference matrix*.

It is well known that if the matrix $(\rho_{ij})_{i,j \ge 1}$ is SM (sM), then the NWCR solution $\mathbf{\Pi}_0$ maximizes (minimizes) the objective function (5) subject to (6). We present this result as the Corollary to the following Proposition; see the Appendix for proofs.

Proposition. If $\mathbf{\Pi}_0$ is given by NWCR and $\mathbf{\Pi}$ is any solution to (6), then we have $\mathbf{\Pi}_0 = \mathbf{\Pi} + \sum_{\gamma=1}^{G} \alpha_\gamma \mathbf{\Delta}_\gamma$ for some positive numbers $\alpha_1, \dots, \alpha_G$, where $\mathbf{\Delta}_\gamma = \mathbf{\Delta}_{IJ}^{ij}$ for some $1 \le i < I \le K$ and $1 \le j < J \le L$ as above.

Corollary. If the matrix $(\rho_{ij})_{i,j \ge 1}$ is SM (sM), then $E_{\mathbf{\Pi}_0}(\rho)$ is maximal (minimal); that is, $E_{\mathbf{\Pi}_0}(\rho) \ge E_{\mathbf{\Pi}}(\rho)$ $(\le E_{\mathbf{\Pi}}(\rho))$ for any other solution $\mathbf{\Pi}$ of (6).

4. MAXIMIZING THE EXPECTED OVERLAP

In Section 2 we proposed to reduce the original TP to a more tractable TP by grouping the old and new samples, s and s', according to the number of common units and introducing a block-optimum function $\rho(c, c')$. By the Corollary, the key to maximization through the NWCR is to ensure that the matrix $(\rho(c, c'))_{(c,c')}$ of block optima is SM. In this section we show how this is accomplished.

4.1 Joint Distribution for Maximal Overlap

To solve the reduced TP for maximization, take ρ in (5) to be

$$\rho^{(M)}(c, c') = \min\{c, c'\}. \tag{9}$$

This represents the largest possible overlap of samples s and s' containing c and c' elements of C. The joint distribution is derived in two steps. The first step is to form a matrix whose rows correspond to the possible values of c and whose columns correspond to the possible values of c', *both in descending order*, and whose entries are the corresponding values of $\rho^{(M)}(c, c')$. Given the hypergeometric probabilities (3) and (4) for super-rows and super-columns, we use NWCR to obtain the joint probabilities $\pi(c, c')$. In the second step, within each block (c, c'), we divide the mass $\pi(c, c')$ equally among all those pairs of samples (s, s') such that $c(s) = c$, $c'(s') = c'$, and $o(s, s') = \rho^{(M)}(c, c')$, that is, the pairs of samples with maximal overlap within the block (c, c').

Example 1. Let $N = 6$, $n = 3$, $D = 3$, and $B = 2$. Then $N' = 5$, from which we select $n' = 4$ units by SRSWOR. Table 1 gives the matrix of block optima $\rho^{(M)}(c, c')$, and Table 2 shows the joint probabilities $\pi(c, c')$ given by the NWCR solution.

We now illustrate the second step for block $(c, c') = (2, 2)$ of Table 1. The samples s and s' in this block are listed in the margins of Table 3, whose entries give the overlap of each pair of samples, where u, d, and b denote units from the common pool, deaths, and births, respectively. Because $\pi(2,2) = .10$ (Table 2) and there are nine pairs of samples with maximal overlap 2 (Table 3), each of these pairs receives probability .0111, whereas the rest of the pairs in the block receive zero probability (Table 4).

The procedure can be summarized as follows:

1. Form a table whose marginals are the super-row and super-column probabilities arranged in descending order of c and c' as above, and derive $\mathbf{\Pi}_0$, the joint distribution of the blocks given by the NWCR.
2. Within each block, divide the mass equally among those individual pairs of samples (s, s') that attain the largest possible overlap for that block.

Theorem 1, proved in the Appendix, gives the key properties of this procedure.

Theorem 1. (a) The joint density $\mathbf{X}_0$ defined on the set of pairs of samples (s, s') by the foregoing two-step procedure has SRSWOR marginals.

(b) $\mathbf{X}_0$ has the maximum expected sample overlap within the set of joint densities with SRSWOR marginals.

(c) All joint densities $\mathbf{X}$ satisfying (a) and (b) have the same variance $V_{\mathbf{X}}(o) = E_{\mathbf{X}}(o^2) - [E_{\mathbf{X}}(o)]^2$.

Remark. If $(s, s') \in (c, c')$, then $o(s, s') \le \rho^{(M)}(c, c')$. Hence, given any solution $\mathbf{X}$ to (2) and the resulting solution $\mathbf{\Pi} = \mathbf{\Pi}(\mathbf{X})$ to (6), we have $E_{\mathbf{X}}(o) = E_{\mathbf{\Pi}(\mathbf{X})}[E_{\mathbf{X}}(o|\rho^{(M)})] \le E_{\mathbf{\Pi}(\mathbf{X})}(\rho^{(M)})$. Equality holds if $\mathbf{X}$ distributes mass optimally within blocks as in step 2 of our procedure, since then $E_{\mathbf{X}}(o|\rho^{(M)})] = \rho^{(M)}$. Thus proving (b) reduces to establishing that $E_{\mathbf{\Pi}}(\rho^{(M)}) \le E_{\mathbf{\Pi}_0}(\rho^{(M)})$ for any $\mathbf{\Pi}$ satisfying (6). Optimal distribution within blocks also implies that $V_{\mathbf{X}}(o) = E_{\mathbf{\Pi}(\mathbf{X})}[V_{\mathbf{X}}(o|\rho^{(M)})] + V_{\mathbf{\Pi}(\mathbf{X})}[E_{\mathbf{X}}(o|\rho^{(M)})] = V_{\mathbf{\Pi}(\mathbf{X})}(\rho^{(M)})$, so proving (c) reduces to showing that $V_{\mathbf{\Pi}}(\rho^{(M)}) = V_{\mathbf{\Pi}_0}(\rho^{(M)})$ whenever $E_{\mathbf{\Pi}}(\rho^{(M)}) = E_{\mathbf{\Pi}_0}(\rho^{(M)})$.

4.2 Selection Algorithms

Theorem 1 motivates the following two algorithms for selecting samples with maximal expected overlap.

Selection Algorithm 1 (Simultaneous selection for maximization). From the NWCR distribution, randomly select a block (c, c'). If $\rho^{(M)}(c, c') = c'$, randomly select c' units from C. To complete the selection of s' with $c'(s') = c'$, randomly select $n' - c'$ births. To complete s with $c(s) = c$, randomly select $c - c'$ units from the remaining $C - c'$ common units and then $n - c$ deaths. The $\rho^{(M)}(c, c') = c$ case is similar.

Selection Algorithm 2 (Sequential selection for maximization). Here s has already been drawn and belongs to super-row $c = c(s)$. We must select s' so that the expected sample overlap is maximized and the second design is preserved. By Theorem 1, it suffices for the conditional selection to correspond to a joint distribution as described in Section 4.1. All rows within a super-row have equal probabilities, so it suffices to consider the conditional distribution of super-columns given the selected super-row c. First, randomly select super-column c' from this conditional distribution. If $c < c'$, then s' consists of the c units from $s \cap \mathsf{C}$, $c' - c$ randomly selected units from the remaining $C - c$ units in C, and $n' - c'$ births. If $c' < c$, randomly select c' units from the c units in $s \cap \mathsf{C}$ and then complete s' by randomly selecting $n' - c'$ births.

Example 2. In Example 1, assume that $s = u_1u_2d_1$ was selected first. For super-row $c = 2$, Table 2 gives $\pi(2,3) = .35$ and $\pi(2,2) = .1$. We randomly pick $r \in [0, .45]$, say $r = .15$. Since $.15 < .35$, the new sample s' is in super-column $c' = 3$. To select s', we retain u_1 and u_2, add u_3 from C, and randomly select one of the two births, say b_2, to obtain $s' = u_1u_2u_3b_2$.

Special Case 1. When $D = 0$, there is only one super-row, with $c(s) = n$. We choose the number of births from the appropriate hypergeometric distribution and select these units as before. We applied this method to select a sample for Statistics Canada's Local Government Surveys. When $B = 0$, there is only one super-column. We select or deselect sequentially to

Table 1. Block Maxima in Example 1

	c'	
c	3	2
3	3	2
2	2	2
1	1	1
0	0	0

Table 2. NWCR Block Probabilities in Example 1

	$q(c')$	
$p(c)$	.40	.60
.05	.05	0
.45	.35	.10
.45	0	.45
.05	0	.05

Table 3. Overlaps Within Block (2, 2) in Example 1

	$u_1u_2b_1b_2$	$u_3u_2b_1b_2$	$u_1u_3b_1b_2$
$u_1u_2d_1$	2	1	1
$u_3u_2d_1$	1	2	1
$u_1u_3d_1$	1	1	2
$u_1u_2d_2$	2	1	1
$u_3u_2d_2$	1	2	1
$u_1u_3d_2$	1	1	2
$u_1u_2d_3$	2	1	1
$u_3u_2d_3$	1	2	1
$u_1u_3d_3$	1	1	2

Table 4. Probabilities Within Block (2, 2) in Example 1

	$u_1u_2b_1b_2$	$u_3u_2b_1b_2$	$u_1u_3b_1b_2$
$u_1u_2d_1$	.0111	0	0
$u_3u_2d_1$	0	.0111	0
$u_1u_3d_1$	0	0	.0111
$u_1u_2d_2$	.0111	0	0
$u_3u_2d_2$	0	.0111	0
$u_1u_3d_2$	0	0	.0111
$u_1u_2d_3$	.0111	0	0
$u_3u_2d_3$	0	.0111	0
$u_1u_3d_3$	0	0	.0111

attain the sample size n'. When $B = D = 0$ (i.e., only the sample size has changed), the maximum overlap $\min\{n, n'\}$ is attained.

5. MINIMIZING THE EXPECTED OVERLAP

By the Corollary to the Proposition, the expected block overlap can be *minimized* by setting up a matrix $(\rho(c, c'))_{(c,c')}$ of block optima that is sM. To that end, we define the block-optimum function to be used in (5) by

$$\rho^{(m)}(c, c') = \max\{0, c + c' - C\}; \tag{10}$$

this represents the smallest possible overlap attainable by a pair of samples within the block (c, c'). As with maximization, a solution to (2) that minimizes the expected sample overlap can be obtained by a two-step procedure:

1. Form super-rows and super-columns as in Section 4.1. Arrange the super-rows in *increasing* order of c and the super-columns in *decreasing* order of c'. (Note the contrast with the maximization procedure; this is needed to obtain an sM rather than SM matrix.) Derive the joint distribution $\mathbf{\Pi}_0$ of the blocks by the NWCR.
2. Within each block, divide the mass equally among those pairs of samples (s, s') that have the smallest possible overlap for that block.

The following result, proved in the Appendix, enables us to devise an algorithm for simultaneous selection for minimization, in analogy to that for maximization.

Theorem 2. (a) The joint density $\mathbf{X}_0$ defined on the set of pairs of samples (s, s') by the foregoing two-step procedure has SRSWOR marginals.

(b) $\mathbf{X}_0$ has the minimum expected sample overlap within the set of joint densities with SRSWOR marginals.

(c) $\mathbf{X}_0$ has the minimum variance within the set of joint densities with SRSWOR marginals that attain the minimum expected sample overlap.

Example 3. By an argument like that in the Remark after Theorem 1, establishing (c) reduces to proving that $V_{\mathbf{\Pi}_0}(\rho^{(m)}) \leq V_{\mathbf{\Pi}}(\rho^{(m)})$ whenever $E_{\mathbf{\Pi}}(\rho^{(m)}) = E_{\mathbf{\Pi}_0}(\rho^{(m)})$. We now present a case in which this inequality is strict. Table 5 displays the block optima (10) for Example 1. Although the rows are now in increasing order of c, the block probabilities are still given by Table 2.

With the notation before the Proposition, consider $\mathbf{\Pi} = \mathbf{\Pi}_0 - .35\mathbf{\Delta}_{32}^{21}$. We have $\langle \mathbf{\Delta}_{32}^{21}, \rho \rangle = 0$ from Table 5, whence $\mathbf{\Pi}$ is a solution that minimizes the expected block overlap. Its variance (.89) exceeds that of $\mathbf{\Pi}_0$ (.19).

Special Case 2. For two identical populations, there is only one block ($c = n$, $c' = n'$). To minimize the expected overlap

Table 5. Block Minima in Example 3

	c'	
c	3	2
0	0	0
1	1	0
2	2	1
3	3	2

when $n + n' \leq N$, we randomly select n units for the first survey and n' different units for the second.

6. COMPARISON WITH A PRN METHOD

In this section we compare the NWCR method with sequential SRSWOR, a PRN method also used to integrate and coordinate stratified SRSWOR designs. Sequential SRSWOR was developed by Fan, Muller, and Rezucha (1962) as a fast computer technique for selecting an SRSWOR sample. Ohlsson (1995) described the method and its use for sample coordination. Briefly, all units on the frame are independently and permanently assigned random numbers from the uniform distribution on [0, 1]. Units corresponding to the first n ordered PRNs encountered when one starts from a point $a \in [0, 1]$ and moves in a specified direction (right or left) are included in the sample. Ohlsson (1992) proved that this technique yields the SRSWOR design. For positive (negative) coordination, one can select s and s' by moving in the same direction (opposite directions) from the same starting point.

Some properties of the PRN methods, such as the expected overlap, apparently have not been studied. We undertook a simulation study to compare the empirical expectations and variances of the overlap $o(s, s')$ for sequential SRSWOR with the values given by the NWCR method. All simulations were based on 10,000 repetitions except those for the small-stratum examples, which were based on 100,000 repetitions.

6.1 Results for Positive Coordination

For positive coordination, the expected sample overlap $E_{\mathbf{X}}(o)$ and variance $V_{\mathbf{X}}(o)$ in strata with small, medium, and large numbers of units are given in Table 6. The NWCR method has larger expected overlap in all scenarios, except Scenario 2, for which the expected overlap and the variance are equal for both methods. Because the frequency distribution of the block overlap is shifted toward higher values for NWCR, higher variance results in four of the six scenarios. A somewhat higher variance may well be a price worth paying for attaining our primary objective of maximal expected overlap.

The differences for $E_{\mathbf{X}}(o)$ in Table 6 may seem too small to be practically significant. However, most establishment surveys have hundreds of strata, and over all strata, the expected overlap could be much larger for the NWCR method.

Table 6. Expectations for Positive Coordination

Scenario	*Sample sizes*	*Method*	$E_{\mathbf{X}}(o)$	$V_{\mathbf{X}}(o)$
Small stratum: $N = 6$, $C = 3$, $N' = 5$				
1	$n = 3$, $n' = 4$	NWCR	1.50	.45
		PRN	1.49	.43
Medium stratum: $N = 40$, $C = 37$, $N' = 41$				
2	$n = 22$, $n' = 19$	NWCR	17.15	.92
		PRN	17.15	.92
3	$n = 20$, $n' = 20$	NWCR	18.05	.92
		PRN	17.78	.65
4	$n = 5$, $n' = 5$	NWCR	4.51	.40
		PRN	4.31	.43
Large stratum: $N = 476$, $C = 466$, $N' = 514$				
5	$n = 50$, $n' = 50$	NWCR	45.33	3.92
		PRN	45.31	3.66
6	$n = 10$, $n' = 10$	NWCR	9.07	.83
		PRN	9.00	.77

Table 7. Empirical Block Probabilities for Sequential SRSWOR in Scenario 1

	$q(c')$	
$p(c)$	.40	.60
.05	.04	.01
.45	.20	.25
.45	.15	.30
.05	.01	.04

Within blocks, both methods divide the probability equally among the pairs of samples (s, s') with $o(s, s') = \rho^{(M)}(c, c')$. (An analogous statement holds for negative coordination.) Thus insight into how the methods differ can best be gleaned by comparing the block probabilities $\pi(c, c')$ for the two methods. These are shown for Scenario 1 in Tables 2 (for NWCR) and 7 (for sequential SRSWOR). We observe that all block probabilities are positive for sequential SRSWOR, whereas some NWCR block probabilities are zero.

Taking $\mathbf{\Pi}$ to be the block probability distribution for sequential SRSWOR, we can deduce that the NWCR method will have higher expected overlap if inequality (7) is strict for some of the $\mathbf{\Delta}_\gamma$'s referred to in the Proposition (see the proof in the App.). Empirically, NWCR realizes the largest gains in expected overlap when $K \approx L$ and $n = n'$. In Scenario 2 $\rho_{ij}^{(M)} + \rho_{rs}^{(M)} = \rho_{is}^{(M)} + \rho_{rj}^{(M)}$ for all i, j, r, and s, and so both methods yield the same expectation and variance.

6.2 Results for Negative Coordination

Table 8 compares the two methods for three negative coordination scenarios. In each case $n + n' > C$, because when $n + n' \leq C$, $\rho^{(m)} \equiv 0$, whence $E_{\mathbf{X}}(o) = 0$ and $V_{\mathbf{X}}(o) = 0$ for both methods. For all scenarios, both the expected overlap and the variance are smaller for the NWCR method than for the sequential SRSWOR method.

As for maximization, the NWCR method yields improvements due to instances of strict inequality in (8). We observed that the range of different $\rho^{(m)}$ values with $\pi(c, c') > 0$ is narrower for NWCR than for sequential SRSWOR; this helps explain the NWCR method's lower variance.

7. DISCUSSION AND EXTENSIONS

The NWCR solution is easily implemented and gives optimal expected overlap while preserving the designs of the surveys considered. The NWCR for maximization applies to both sequential and simultaneous selection, solves well-defined mathematical problems, and has verifiable properties, such as minimum variance in the case of negative coordination (Thm. 2). In this section we discuss possible extensions to accommodate migration among strata, as well as the case of more than two surveys.

Table 8. Expectations for Negative Coordination

Scenario	Sample sizes	Method	$E_{\mathbf{X}}(o)$	$V_{\mathbf{X}}(o)$
Small stratum: $N = 6$, $C = 3$, $N' = 5$				
7	$n = 3$, $n' = 4$	NWCR	.90	.19
		PRN	.91	.52
Medium stratum: $N = 40$, $C = 37$, $N' = 41$				
8	$n = 22$, $n' = 20$	NWCR	1.40	.24
		PRN	1.45	1.25
9	$n = 20$, $n' = 20$	NWCR	0	0
		PRN	.27	.32

7.1 Change in the Composition of Strata

The NWCR method is readily adapted to accommodate changes in the composition of strata when each stratum in the new survey can be paired with a stratum in the old survey (e.g., when unit classification codes must be updated but the classification system defining the strata remains unchanged). The goal in such applications is generally *positive* coordination. Within each such pair of strata, we treat units from the old stratum that moved outside the corresponding new stratum as deaths and units that migrated to the new stratum from outside the corresponding old stratum as births. This approach produces an SRSWOR design on the second occasion but no longer optimizes the expected overlap.

A second, more general approach for positive coordination is not based on the NWCR algorithm, but retains the basic idea of reducing the TP through a grouping procedure, and does maximize the expected overlap. For each new stratum S', define $S'_i = S_i \cap S'$, $i = 1, \ldots, l$, where $S_1, \ldots, S_l$ are the first-design strata for which this intersection is nonempty. In practice, l is generally quite small. A super-row is defined by $\tilde{c} = (c(i))_{i=1,\ldots,l}$, where $c(i) = \text{card}\{s \cap S'_i\}$ for each s in the super-row. The super-row probabilities can be calculated from the original probabilities of the stratified design. The super-columns are similarly defined by $\tilde{c}' = (c'(i))_{i=1,\ldots,l}$, where $c'(i) = \text{card}\{s' \cap S'_i\}$, and have a hypergeometric distribution. Defining the block-optimum function as $\rho(\tilde{c}, \tilde{c}') = \sum_{i=1}^{l} \min\{c(i), c'(i)\}$, we solve a TP to maximize the expected block overlap, and distribute the mass within blocks as in step 2 of Section 4.1. For $l = 1$, this approach reduces to that of Section 4.

7.2 More Than Two Surveys

With $d > 2$ surveys, the first step toward obtaining a reduced TP is to divide the units among $2^d - 1$ classes according to which survey populations they belong to (e.g., units belonging only to the first and third populations constitute one class). Samples for each survey can then be grouped according to number of units from each such class (generalizing the super-rows and super-columns of the $d = 2$ case), and d-dimensional blocks can be formed from these groups. There are different ways to define the overlap. For $j = 2, \ldots, d$, we assign costs or utilities for a unit's membership in j of the samples; these quantities are then used to define a block overlap function. We can then solve a reduced TP and distribute the mass within blocks as for $d = 2$. It may be possible to maximize the expectation of at least some overlap functions through a greedy algorithm. Optimality proofs would depend on higher-dimensional analogs of Monge matrices (Bein, Brucker, Park, and Pathak 1995). The $d > 2$ case is the subject of ongoing work.

APPENDIX: PROOFS

Proof of the Proposition

The proof is similar to that of result 5.6.2 of Arthanari and Dodge (1981). Order the nonzero entries of $\mathbf{\Pi}_0$ as in Section 3.1. We suppose that the first k nonzero entries of $\mathbf{\Pi}_0$ equal the corresponding entries of $\mathbf{\Pi}$, and describe a procedure to obtain $\mathbf{\Pi}^* = \mathbf{\Pi} + \sum_{\gamma=1}^{G^*} \alpha_\gamma \mathbf{\Delta}_\gamma$, another solution of (6), the first $k+1$ nonzero entries of which equal the corresponding entries of $\mathbf{\Pi}_0$. This will prove the Proposition, because by applying this procedure repeatedly, we can obtain $\mathbf{\Pi}_0$. Assume that (i,j) corresponds to the $(k+1)$st nonzero entry of $\mathbf{\Pi}_0$ and that $\pi_{ij}^0 \neq \pi_{ij}$. Because of the constraints, $\mathbf{\Pi}$ also coincides with $\mathbf{\Pi}_0$ on all previously assigned zero entries. Consequently, on row i, $\pi_{ij'} = \pi_{ij'}^0$ for $j' < j$, and on column j, $\pi_{i'j} = \pi_{i'j}^0$ for $i' < i$ and thus $\pi_{ij} < \pi_{ij}^0$. For $\mathbf{\Pi}$ to have the correct marginals, there must then exist nonempty sets $R = \{r > i : \pi_{rj} > \pi_{rj}^0\}$ and $S = \{s > j : \pi_{is} > \pi_{is}^0\}$. Pick $r \in R$, $s \in S$, let $\alpha_1 = \min\{\pi_{rj} - \pi_{rj}^0, \pi_{is} - \pi_{is}^0\}$, and define $\mathbf{\Pi}_1 = \mathbf{\Pi} + \alpha_1 \mathbf{\Delta}_{rs}^{ij}$. $\mathbf{\Pi}_1$ is a solution to (6) such that (using obvious notation) either $R_1 = \{r > i : \pi_{rj}^1 > \pi_{rj}^0\}$ or $S_1 = \{s > j : \pi_{is}^1 > \pi_{is}^0\}$ has one fewer element than R or S. Form $\mathbf{\Pi}_2$ from $\mathbf{\Pi}_1$, R_1, and S_1 in the same way that $\mathbf{\Pi}_1$ was formed from $\mathbf{\Pi}$, R, and S. Continue in this way, forming $\mathbf{\Pi}_3, \mathbf{\Pi}_4, \ldots, \mathbf{\Pi}_{G^*}$ and stopping when either R_{G^*+1} or S_{G^*+1} is empty and $\pi_{ij}^{G^*} = \pi_{ij}^0$. Now all of $\mathbf{\Pi}_{G^*}$'s nonzero entries equal the corresponding entries of $\mathbf{\Pi}_0$, so we can take $\mathbf{\Pi}^* = \mathbf{\Pi}_{G^*} = \mathbf{\Pi}_0$.

Proof of the Corollary to the Proposition

Because the matrix of block optima is SM (sM), we have $\langle \mathbf{\Delta}_\gamma, \rho \rangle \geq 0$ (≤ 0), $\gamma = 1, \ldots, G$. The maximality (minimality) of $E_{\mathbf{\Pi}_0}(\rho)$ follows directly from the Proposition since $E_{\mathbf{\Pi}_0}(\rho) = \langle \mathbf{\Pi}_0, \rho \rangle = \langle \mathbf{\Pi}, \rho \rangle + \sum_{\gamma=1}^{G} \alpha_\gamma \langle \mathbf{\Delta}_\gamma, \rho \rangle$.

Proof of Theorem 1

(a) Because the super-rows (-columns) have the correct marginals by construction, it suffices to show that all $\binom{C}{c}\binom{D}{n-c}$ rows s within super-row c [all $\binom{C}{c'}\binom{B}{n'-c'}$ columns s' within super-column c'] receive equal probabilities in the NWCR method. This follows from the symmetry of the distribution of mass within blocks, where each row (column) receives equal mass.

(b) By the Remark after Theorem 1, it suffices to prove that $E_{\mathbf{\Pi}}(\rho^{(M)}) \leq E_{\mathbf{\Pi}_0}(\rho^{(M)})$ for any $\mathbf{\Pi}$ satisfying (6). By the Corollary, we need only show that the matrix of block optima given by (9) is SM, that is, $\min\{c_i, c_j'\} + \min\{c_r, c_s'\} \geq \min\{c_i, c_s'\} + \min\{c_r, c_j'\}$ whenever $1 \leq r \leq i$, $1 \leq s \leq j$. But this inequality is equivalent to $h(c_r) \geq h(c_i)$, where $h(k) = \min\{k, c_s'\} - \min\{k, c_j'\}$. Since $c_s' \geq c_j'$, h is nondecreasing. Hence, because $c_r \geq c_i$, we have $h(c_r) \geq h(c_i)$, as required.

(c) We drop the superscript from $\rho^{(M)}(c, c') = \min\{c, c'\}$ for this part of the proof. By the Remark after Theorem 1, it suffices to show that $E_{\mathbf{\Pi}_0}(\rho) = E_{\mathbf{\Pi}}(\rho)$ implies $E_{\mathbf{\Pi}_0}(\rho^2) = E_{\mathbf{\Pi}}(\rho^2)$. The Proposition enables us to write $\mathbf{\Pi}_0 = \mathbf{\Pi} + \sum_{\gamma=1}^{G} \alpha_\gamma \mathbf{\Delta}_\gamma$, whence $E_{\mathbf{\Pi}_0}(\rho^2) = E_{\mathbf{\Pi}}(\rho^2)$ reduces to

$$\langle \mathbf{\Delta}_\gamma, \rho^2 \rangle = 0 \tag{A.1}$$

for each $\gamma = 1, \ldots, G$. Since the matrix of block optima is SM, $\langle \mathbf{\Delta}_\gamma, \rho \rangle \geq 0$ for each γ. But $E_{\mathbf{\Pi}_0}(\rho) = E_{\mathbf{\Pi}}(\rho)$ implies that $\langle \mathbf{\Delta}_\gamma, \rho \rangle = 0$ for each γ. Writing $\mathbf{\Delta}_\gamma = \mathbf{\Delta}_{IJ}^{ij}$, the equality of expectations is equivalent to $\rho_{ij} + \rho_{IJ} = \rho_{iJ} + \rho_{Ij}$, where $\rho_{ab} = \min\{c_a, c_b'\}$, $c_i \geq c_I$, and $c_j' \geq c_J'$. Without loss of generality, $c_I \geq c_J'$. Then $\rho_{IJ} = \rho_{iJ}$, and hence $\rho_{ij} = \rho_{Ij}$, so that $\rho_{ij}^2 + \rho_{IJ}^2 = \rho_{iJ}^2 + \rho_{Ij}^2$, which is equivalent to (A.1).

Proof of Theorem 2

(a) This follows as in the proof of Theorem 1.

(b) Writing $\rho(c, c') = \rho^{(m)}(c, c') = \max\{c + c' - C, 0\}$ and reasoning as in the proof of Theorem 1(b), it suffices to show that the matrix of block optima is sM, that is, that $\rho(c_1, c_1') + \rho(c_2, c_2') \leq \rho(c_1, c_2') + \rho(c_2, c_1')$ whenever $c_1 < c_2$ and $c_1' > c_2'$. This inequality is equivalent to $g(a+h) - g(a) \leq g(b+h) - g(b)$, where $g(t) = \max\{t, C\}$, $a = c_1 + c_2'$, $b = c_1 + c_1'$, and $h = c_2 - c_1$. The latter inequality follows from the convexity of g.

(c) Let $\rho_{ab} = \max\{c_a + c_b' - C, 0\}$ denote a typical element of the matrix of block minima. Arguing as in the proof of Theorem 1(c), we reduce the problem to that of assuming $\langle \mathbf{\Delta}_{IJ}^{ij}, \rho \rangle = 0$, with $c_i < c_I$ and $c_j' > c_J'$, and proving that $\langle \mathbf{\Delta}_{IJ}^{ij}, \rho^2 \rangle \leq 0$, that is,

$$\rho_{ij}^2 - \rho_{iJ}^2 + \rho_{IJ}^2 - \rho_{Ij}^2 \leq 0. \tag{A.2}$$

Because $\rho_{ij} - \rho_{iJ} = \rho_{Ij} - \rho_{IJ}$, $\rho_{ij}^2 - \rho_{iJ}^2 + \rho_{IJ}^2 - \rho_{Ij}^2 = (\rho_{ij} - \rho_{iJ}) \times [(\rho_{ij} + \rho_{iJ}) - (\rho_{Ij} + \rho_{IJ})]$. The first factor of the latter expression is nonnegative, whereas the second is the sum of $\max\{c_i + c_j', C\} - \max\{c_I + c_j', C\}$ and $\max\{c_i + c_J', C\} - \max\{c_I + c_J', C\}$, neither of which is positive since $c_i < c_I$. Thus (A.2) holds.

[Received September 2004. Revised December 2005.]

REFERENCES

Aragon, J., and Pathak, P. K. (1990), "An Algorithm for Optimal Integration of Two Surveys," *Sankhyā*, Ser. B, 52, 198–203.

Arthanari, T. S., and Dodge, Y. (1981), *Mathematical Programming in Statistics*, New York: Wiley.

Bein, W. W., Brucker, P., Park, J. K., and Pathak, P. K. (1995), "A Monge Property for the *d*-Dimensional Transportation Problem," *Discrete Applied Mathematics*, 58, 97–109.

Brewer, K. R. W., Early, L. J., and Joyce, S. F. (1972), "Selecting Several Samples From a Single Population," *Australian Journal of Statistics*, 14, 231–239.

Causey, B. D., Cox, L. H., and Ernst, L. (1985), "Applications of Transportation Theory to Statistical Problems," *Journal of the American Statistical Association*, 80, 903–909.

Cotton, F., and Hesse, C. (1992), "Méthodes d'échantillonnage pour l'enquête annuelle d'entreprises," *Actes des Journées de Méthodologie Statistique*, 13 et 14 mars 1991, Insee Méthodes n° 29-30-31, 21–37.

Ernst, L. R. (1998), "Maximizing and Minimizing Overlap When Selecting a Large Number of Units per Stratum Simultaneously for Two Designs," *Journal of Official Statistics*, 14, 297–314.

——— (1999), "The Maximization and Minimization of Sample Overlap Problems: A Half Century of Results," in *Proceedings of the International Statistical Institute*, pp. 168–182.

Ernst, L. R., and Ikeda, M. M. (1994), "A Reduced-Size Transportation Algorithm for Maximizing the Overlap Between Surveys," *Survey Methodology*, 21, 147–157.

Ernst, L. R., and Paben, S. P. (2002), "Maximizing and Minimizing Overlap When Selecting Any Number of Units per Stratum Simultaneously for Two Designs With Different Stratifications," *Journal of Official Statistics*, 18, 185–202.

Fan, C. T., Muller, M. E., and Rezucha, I. (1962), "Development of Sampling Plans by Using Sequential (Item-by-Item) Techniques and Digital Computers," *Journal of the American Statistical Association*, 57, 387–402.

Fellegi, I. P. (1966), "Changing the Probabilities of Selection When Two Units Are Selected With PPS Without Replacement," in *Proceedings of the Social Statistics Section*, American Statistical Association, pp. 434–442.

Hoffman, A. J. (1985), "On Greedy Algorithms That Succeed," in *Surveys in Combinatorics*, ed. I. Anderson, Cambridge, U.K.: Cambridge University Press, pp. 97–112.

Keyfitz, N. (1951), "Sampling With Probabilities Proportionate to Size: Adjustment for Changes in Probabilities," *Journal of the American Statistical Association*, 46, 105–109.

McKenzie, B., and Gross, B. (2000), "Synchronized Sampling," in *ICES II, The Second International Conference on Establishment Surveys*, American Statistical Association, pp. 237–243.

Mitra, S. K., and Pathak, P. K. (1984), "Algorithm for Optimal Integration of Three Surveys," *Scandinavian Journal of Statistics*, 11, 257–316.

Ohlsson, E. (1992), "SAMU, a System for Coordination of Samples From the Business Register at Statistics Sweden: A Methodological Description," Research and Development Report 1992:18, Statistics Sweden.

——— (1995), "Coordination of Samples Using Permanent Random Numbers," in *Business Survey Methods*, eds. B. G. Cox, D. A. Binder, D. N. Chinnappa, A. Christianson, M. J. Colledge, and P. S. Kott, New York: Wiley, pp. 153–169.

——— (2000), "Coordination of PPS Samples Over Time," in *ICES II, The Second International Conference on Establishment Surveys*, American Statistical Association, pp. 255–264.

Raj, D. (1956), "On the Method of Overlapping Maps in Sample Surveys," *Sankhyā*, 17, 89–98.

Rivière, P. (2001), "Coordinating Samples Using the Microstrata Methodology," in *Proceedings of Statistics Canada Symposium Session 8*, CDROM 11-522-XIE; available at *http://www.statcan.ca/english/freepub/11-522-XIE/11-522XIE2001001.htm*.

Ross, S. M. (1983), *Introduction to Stochastic Dynamic Programming*, San Diego: Academic Press.

Royce, D. (2000), "Issues in Coordinated Sampling at Statistics Canada," in *ICES II*, *The Second International Conference on Establishment Surveys*, American Statistical Association, pp. 245–254.