Part B

Optimal Design For the 2001 Redesign of the Monthly Population Survey

Introduction

107. The MPS adopts a multi-stage design in which the sample is clustered within first stage units. A high level of clustering will reduce costs in travel between first stage units which account for a high proportion of overall costs. However this will result in higher variances as fewer first stage units will be selected (for a fixed total sample size). On the other hand a low level of clustering will cost more in travel but will produce lower variances on estimates.

108. The objective of the optimisation process is to determine the level of clustering that achieves the best trade-off between costs and variance by minimising total cost for a fixed level of accuracy. Key components of the optimisation process are the cost and variance models which provide the link between sample sizes at each stage of selection and resulting costs and variances, respectively. As survey accuracy deteriorates over the life of a design, the prime objective of the sample redesign is to return to the level of accuracy achieved at the beginning of the current (1996) design period.

109. In the following sections we outline a couple of the previous sample optimisation methods and propose an alternative method for the 2001 redesign.

1986 and 1991 Optimisation Methods

110. In the 1986 and 1991 redesigns the optimal cluster size for each area type was determined by minimising area type cost subject to the variance for the given area type. The optimisation problem was formulated as follows:

Minimise $\left(c_{0i} + c_{1i}m_i + c_{2i}m_i q_i\right)$

subject to $\left(v_{0i} + \frac{v_{1i}}{m_i} + \frac{v_{2i}}{m_i q_i}\right) = K_i$

where the subscript *i* indicates a geographical area type.

111. The area type level variance model in this approach assumes that the number of first stage units in an area type (m_i) are distributed across states in accordance with the same relativities as the state skips. In this way area type variance models can be thought of as being implicitly built up from variance models at state by area type level, which were never produced in redesigns prior to 2001, due to the limitations in computing capabilities.

112. The optimal cluster size, which is the most important design parameter to be output from the optimisation process, is given by:

$$q_i = \sqrt{\frac{c_{1i}}{c_{2i}} \frac{v_{2i}^{(N)}}{v_{1i}^{(N)}}} \qquad \boxed{1}$$

113. The cluster size is optimal in the sense that it does not depend upon the state skips which are set in advance, independently of this optimisation process, and are based on the Carroll allocation with some adjustments. The total sample size ($m_i q_i$) is actually determined by applying the state skips to state by area type population sizes and then summing the resulting sample size across states within an area type. This means that the final design sample size is not optimally decided.

1996 Redesign Optimisation

114. The 1996 optimisation method involved cost and variance models, again at the area type level, but differed from the 1986/1991 method in that a sample size constraint was introduced. The 1996 method was of the form:

Minimise National Cost = $\sum_i \left(c_{0i} + c_{1i} m_i + c_{2i} m_i q_i \right)$

subject to $\sum_i \left(v^{(N)}{}_{0i} + \frac{v^{(N)}{}_{1i}}{m_i} + \frac{v^{(N)}{}_{2i}}{m_i q_i} \right) = K^{(N)}$

and $n_i = \psi n_{0i}$ $\boxed{2}$

and optimising for q_i, m_i and ψ.

115. The sample size constraint, involving benchmark area type sample sizes n_{0i}, was introduced as a way of controlling the sample allocation to meet design objectives. The main advantages of the 1996 optimisation method over the previous method were that firstly equal probability sampling within state was included in the optimisation (see below) and secondly that cost and variance across all area types were optimised and constrained, not just at individual area type level as in 1986/1991.

116. The value n_{0i} in the sample size constraint equation $\boxed{2}$ was based on the achieved area type sample size at the beginning of the 1991 design period. Clark and Steel (2000) show how this constraint equation is derived from three implicit constraints. The first of these reflected the perceived priorities of the accuracies of national, state and territory estimates. The second constrained the cluster sizes to depend on area type and not on state, while the third ensured equal probability selection for all dwellings within a state. Mathematically these three constraints can be expressed as follows:

$$n_s = \beta \, z_s$$

$$q_{is} = q_i$$

$$\frac{n_{is}}{N_{is}} = \theta_s$$

where z_s = the measure of importance for the accuracy of estimates of state s.

117. Clark and Steel (2000) show that these three constraints are satisfied if and only if

$$n_0^i = \alpha \, N_i^* \qquad \boxed{3}$$

where N_i^* can be thought of as a "state weighted population count" for area type *i* and is given by:

$$N_i^* = \sum_s \frac{z_s / z}{N_s / N} N_{is}$$

118. As $z_s \propto n_s$ it can be readily shown that constraint equation $\boxed{3}$ is equivalent to constraint $\boxed{2}$ used in the optimisation.

119. The area type sample size values (n_{0i}) used in the constraint equation were obtained by dividing the latest state by area type population counts by a preliminary state skip and then aggregating the result across state. The preliminary state skip was produced by adjusting the previous design state skip for any changes in coverage rates and changes in population size since the previous redesign. The optimisation was then carried out giving the following solutions for the optimal design parameter values:

$$q_i^{opt} = \frac{\sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}}}{\sqrt{\sum_i \left(c_{2i} n_{0i}\right)}} \sqrt{\frac{c_{1i}}{v_{1i}^{(N)}}}\; n_{0i} \qquad \boxed{4}$$

$$m_i^{opt} = \left(\frac{\sum_i \sqrt{v_{1i}^{(N)} c_{1i}} + \sqrt{\sum_i \left(c_{2i} n_{0i}\right)} \sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}}}{\left(K - \sum_i v_{0i}^{(N)} \right)} \right) \sqrt{\frac{v_{1i}^{(N)}}{c_{1i}}} \qquad \boxed{5}$$

$$n_i^{opt} = \frac{\sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}}}{\left(K - \sum_i v_{0i}^{(N)} \right)} \left(\frac{\sum_i \sqrt{v_{1i}^{(N)} c_{1i}}}{\sqrt{\sum_i \left(c_{2i} n_{0i}\right)}} + \sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}} \right) n_{0i} \qquad \boxed{6}$$

$$\psi = \frac{\sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}}}{\left(K - \sum_i v_{0i}^{(N)} \right)} \left(\frac{\sum_i \sqrt{v_{1i}^{(N)} c_{1i}}}{\sqrt{\sum_i \left(c_{2i} n_{0i}\right)}} + \sqrt{\sum_i \frac{v_{2i}^{(N)}}{n_{0i}}} \right) \qquad \boxed{7}$$

120. As mentioned earlier, the state skips are not "optimal" as they are determined by expedient considerations regarding the relative importance of the accuracy of state estimates, together with a desire to maintain, where possible, previous state sample sizes. State skips generally remain close to those determined by the Carroll allocation method which controls the sample allocation to states in accordance with the perceived relative priorities of state estimates. This results in an allocation that is essentially proportional to the square root of the state's population size. By simply updating the state skips according to changes in population the optimal nature of the state skips should be preserved. However this is approximate in nature as the Carroll allocation method is based on a single stage design and did not take account of differences in variance structure between the states and territories. We propose an alternative method which gets around this problem.

A Possible Alternative Optimisation Method for 2001

121. The main drawback of the 1996 and previous redesign optimisation methods is that relative priorities for the accuracies of state and national estimates were determined indirectly using state sample sizes rather than by desired state and national relative variances. What we require from an optimisation method is to minimise national costs while ensuring that the following constraints are met:

- the national relative variance returns to the same value as at the beginning of the 1996 design period,
- state relative variances are in the same ratios to one another as they were for the 1996 design period, and
- equal probability selection within states is achieved resulting in constant state skips.

122. The last two constraints involve relative variances and design parameters at the state level. Area type is still a key determinant of costs, so the cost models still need to be the area type level. Likewise state by area type variance models provide a convenient way of forming variance models at the state level. Therefore an alternative sample allocation is suggested by making use of state by area type cost and variance models to specify these constraints more precisely. This approach at the present time is premised on conceptual grounds only and has not been empirically evaluated to determine whether it compares favourably with previous optimisation methods in terms of efficiency or robustness.

123. Variance models have never been available at the state by area type level in previous redesigns but an issue worth considering before adopting this method is the robustness of the 2001 state by area type variance models. While the model R^2 values are not quite as high as they are for area type models they are still generally quite acceptable with most R^2 values being 90% or higher and a considerable number being 97% or more. (Refer to Tables 1 and 2 of Attachment A for details) For this reason it is proposed to optimise cluster sizes at the state by area type level rather than at area type level alone as in previous redesigns.

124. Nevertheless, further validation on state by area type variance models is intended to be carried out by calculating the standard error on the optimal q_{is} values (conditioning on cost model parameters) to gauge the impact of variance model errors on the key design parameter. If some state by area type variance model parameters do have sufficiently high errors, the affected state by area types can be collapsed to generate more robust models. Alternatively the affected state by area type model parameters can be determined by a calibration approach that ensures additivity to the area type level.

125. The proposed optimisation process will still be iterative in nature in that initial values for the relative importance of state estimates will be reassessed in light of cost and considerations of expediency. Any changes to the values for state or national accuracy levels can be made to the variance constraint values and the optimisation process run again. The trade-off or gain in costs resulting from the change in relativities of state accuracy levels can then be examined to determine whether further changes are required. The process thus becomes an iterative one.

126. Cost models at state by area type level are premised on the assumption that costs are linear with the number of blocks and dwellings which implies that the cost model parameters at the state by area type level are essentially the same as those at area type level.

127. The proposed optimisation problem can thus be formulated as follows:

$$\left.\begin{array}{lll}
\textit{Minimise} & \sum_s \sum_i \left(C_{0is} + C_{1is} m_{is} + C_{2is} m_{is} q_{is}\right) & \\
\textit{subject to} & \sum_s \sum_i \left(V_{0is}^{(N)} + \dfrac{V_{1is}^{(N)}}{m_{is}} + \dfrac{V_{2is}^{(N)}}{m_{is} q_{is}} \right) = K_N^{96} & (a) \\
& \sum_i \left(V_{0is}^{(S)} + \dfrac{V_{1is}^{(S)}}{m_{is}} + \dfrac{V_{2is}^{(S)}}{m_{is} q_{is}} \right) = \phi\ K_s^{96} & (b) \\
& \qquad \forall \text{ states } s=1,\ldots,8 & \\
& m_{is} = \dfrac{\theta_s N_{is}}{q_{is}} & (c)
\end{array}\right\} (A)$$

where the superscript (N) denotes the model parameter relative to the square of national totals and the superscript (S) denotes the model parameter relative to the square of state totals. Thus the left hand side of constraint (a) represents the national relative variance while the left hand side of constraint (b) represents the state relative variance.

128. The relative variance parameters above are actually a hybridisation between employment and unemployment variance models and are different at state and national levels. Further details regarding the hybridisation method can be found in Attachment B. If we were optimising for employment or unemployment alone the national relative variance constraint (a) would be a linear combination of the state relative variance constraints (b), with $\phi = 1$. However in the hybridised case, constraints (a) and (b) should be independent.

129. The national and state level variance constraint values are obtained from the predicted values of the 2001 national and state variance models using the 1996 redesign parameter values for cluster size (q_i) and number of selected clusters (m_{is}). Thus

$$K_s^{96} = \sum_i \left(V_{0is}^{(S)} + \frac{V_{1is}^{(S)}}{m_{is}^{96}} + \frac{V_{2is}^{(S)}}{m_{is}^{96} q_i^{96}} \right)$$

$$K_N^{96} = \sum_s \sum_i \left(V_{0is}^{(N)} + \frac{V_{1is}^{(N)}}{m_{is}^{96}} + \frac{V_{2is}^{(N)}}{m_{is}^{96} q_i^{96}} \right)$$

130. There are several issues that need to be addressed concerning how best to ensure that the constraint values in problem (A) best reflect MPS circumstances during the 2001-2006 design period. These are:

i. the variance model terms and constraint values K_N^{96} and K_S^{96} appearing in constraints (a) and (b) are relative variances. Employment and unemployment total values used in the denominators of the variance models should probably reflect those values expected to be realised during most of the 2001 design period. The denominators for the relative variance constraint values are likely to be similar to those from the beginning of the 1996 design period. However adjustments will need to be made to any states that have changed population size appreciably since then and will therefore have more or less impact on national estimates.

ii. An assessment needs to be made as to whether the variance structure implicit in the variance models derived from 1996 census data, has changed since then. To determine whether this is the case empirical variance models derived from the MPS would have to be determined, but at this stage resources are not sufficient to make this possible.

iii. A possible modification is to produce post-stratified variance models which more closely reflect the estimation method used in the MPS. However again there are not sufficient resources to enable us to do this.

iv. The variance models assume that sample selections have no sample loss. In reality around 15% of selected dwellings are either unoccupied or do not have in-scope persons in them, although this figure varies somewhat between states and area types. Allowance needs to be made for this in the optimisation so that achieved variances reflect the likely amount of sample loss.

Solution to Optimisation Problem (A)

131. Optimisation problem (A) can be re-parameterised in terms of $\frac{1}{m_{is}}, \frac{1}{\theta_s}$ *and* ϕ by substituting constraint (c) for $m_{is} q_{is}$ into the remainder of the optimisation problem (A).

132. This optimisation problem is not quite as straight forward to solve as in previous methods, however with modern computing capabilities, numerical non-linear optimisation methods are far more feasible than they have been in the past. However the optimisation problem given by (A) can at least be partly solved analytically using the Lagrangian method. By simultaneously solving those equations in the Lagrangian condition that relate to first order partial derivatives with respect to θ_s and m_{is} we obtain the following expressions:

$$\theta_s = \frac{\sqrt{\sum_i \left(\mu \frac{V_{2is}^{(N)}}{N_{is}} + \lambda_s \frac{V_{2is}^{(S)}}{N_{is}} \right)}}{\sqrt{\sum_i \left(N_{is} C_{2is} \right)}} \qquad \boxed{8}$$

$$m_{is} = \frac{\sqrt{\mu \; V_{1is}^{(N)} + \lambda_s V_{1is}^{(S)}}}{\sqrt{C_{1is}}} \qquad \boxed{9}$$

Substituting expressions $\boxed{8}$ and $\boxed{9}$ into the constant state sampling fraction constraint (c) gives an expression for the optimal cluster size, viz.:

$$q_{is} = \frac{\sqrt{\sum_i \left(\mu \frac{V_{2is}^{(N)}}{N_{is}} + \lambda_s \frac{V_{2is}^{(S)}}{N_{is}} \right)}}{\sqrt{\sum_i \left(N_{is} C_{2is} \right)}} \frac{\sqrt{C_{1is}} \; N_{is}}{\sqrt{\mu \; V_{1is}^{(N)} + \lambda_s V_{1is}^{(S)}}} \qquad \boxed{10}$$

133. Comparing the expression for the optimal cluster size under the 1986/91 optimisation method (equation $\boxed{1}$) with that under the proposed method (equation $\boxed{10}$) we observe that firstly state and national relative variance parameters are weighted by the Lagrangian multipliers λ_s and μ, respectively. Secondly the second stage variance and cost model parameters are adjusted by the state by area type population size, due to the constant within state sampling fraction constraint.

134. As mentioned earlier the optimisation process is an iterative one in which adjustments may be made to the relative priorities of state relative variances. Note that unlike previous optimisation methods in which the optimal cluster size remain unchanged by this process, under the proposed method all optimal design parameters will be adjusted to some degree at each iteration via updating of the λ_s's and μ.

135. To determine λ_s and μ, and for that matter ϕ, we simultaneously solve the remaining equations from the Lagrangian condition. These are constraint equations (a) and (b) (after re-parameterisation in terms of λ_s and μ by substitution of equations $\boxed{8}$ and $\boxed{9}$) and a third equation which states that the λ_s's are orthogonal to the state relative variance constraint values K_s^{96}. As there are actually eight equations under constraint (b), this gives us ten equations in ten variables, which may be solved simultaneously using the Newton-Raphson method in several variables (See Johnson and Riess (1977)).

Action items for Further Work

136. In order to properly evaluate this method against previous optimisation methods we propose the following set of action items, subject to available resources:

a. Evaluate the fitness of state by area type variance models for use in this method. One possible measure is the calculation of an estimate for the standard error on the optimal cluster size, under this approach, conditioning on the cost model parameters.
b. Deal with state by area type models that are not sufficiently accurate by collapsing with similar state by area types.
c. Apply the method verifying that any solution obtained by the use of the Newton-Raphson method is a global solution to the optimisation problem. Newton-Raphson may fail to converge for particular initial values, making it necessary to trial numerous initial values.
d. Compare the efficiency of this method with those of the 1996 and 1986 optimisation methods.
e. Assess the stability or robustness of the sample allocation method through a sensitivity analysis of the cost and variance parameters as well as the constraint values. The presence of uncharacteristically large or small cluster sizes or state skips that are substantially different to traditional values is likely to suggest instability in the method.

Attachment A

Table 1: State by Area Type Variance Models - Employment Stratum Based Selections

State	Area Type	V_0	V_1	V_2	Sqrt(V_2/V_1)	R^2
1	1	-2862231	5.16E+008	3.77E+009	2.70	87.4%
	2	-2652168	1.73E+009	5.38E+010	5.57	98.2%
	3	-19047171	2.30E+010	4.53E+011	4.44	99.6%
	4	-9411727	1.11E+010	1.33E+011	3.46	98.4%
	6	-489543	1.13E+008	1.52E+009	3.67	97.7%
	7	-27256450	1.54E+010	2.08E+011	3.68	97.7%
	8	-2376389	5.39E+008	7.49E+009	3.73	98.9%
	9	-2996998	2.45E+008	4.10E+009	4.10	99.0%
	10	-2827654	1.00E+009	1.51E+010	3.88	98.8%
	11	-473667	5.39E+008	1.03E+010	4.38	98.9%
	12	-34926	2.22E+006	1.85E+007	2.89	88.1%
2	1	-1805867	3.35E+008	3.99E+009	3.46	91.8%
	2	-588677	1.12E+008	3.88E+009	5.89	98.1%
	3	-20892278	2.41E+010	5.66E+011	4.85	99.1%
	4	-6379500	4.13E+009	5.39E+010	3.61	98.5%
	6	-478424	5.22E+007	1.05E+009	4.49	97.3%
	7	-6522462	2.57E+009	4.22E+010	4.05	99.0%
	8	-2968290	2.12E+008	3.23E+009	3.91	98.2%
	9	-963573	9.99E+007	1.94E+009	4.40	98.5%
	10	327970	3.29E+007	2.66E+009	8.99	97.7%
	11	-2123389	5.00E+008	7.31E+009	3.82	98.6%
3	2	46416	7.83E+006	3.87E+008	7.04	97.3%
	3	-3517785	2.02E+009	4.93E+010	4.94	99.2%
	4	-8667419	5.56E+009	6.75E+010	3.48	99.0%
	6	-2294748	1.48E+008	1.19E+009	2.83	94.1%
	7	-12286346	5.73E+009	8.83E+010	3.93	98.3%
	8	-3546594	1.02E+009	1.49E+010	3.82	97.3%
	9	-1688277	2.53E+008	3.99E+009	3.97	99.2%
	10	-1656450	3.82E+008	4.88E+009	3.58	98.0%
	11	-3011258	4.92E+008	9.28E+009	4.34	98.0%
	12	-791371	5.08E+007	3.60E+008	2.66	97.0%
4	3	-3499003	3.25E+009	5.67E+010	4.18	97.7%
	4	-1042571	1.23E+009	1.90E+010	3.93	97.6%
	6	-175471	5.35E+006	1.17E+008	4.67	93.2%

	7	-403456	1.03E+008	1.48E+009	3.79	93.5%
	8	-1056180	6.74E+007	1.12E+009	4.09	97.9%
	9	-19802	2.42E+007	4.84E+008	4.48	93.1%
	10	98779	5.24E+006	2.32E+008	6.65	96.5%
	11	-370754	5.07E+007	6.23E+008	3.51	98.3%
	12	27595	2.23E+006	9.09E+006	2.02	94.6%
5	3	-1837828	9.20E+008	2.29E+010	4.99	97.2%
	4	-3582539	3.83E+009	8.63E+010	4.75	98.7%
	6	-178778	2.55E+007	3.82E+008	3.87	96.4%
	7	-1257738	3.02E+008	3.69E+009	3.50	97.7%
	8	-781883	2.26E+007	3.41E+008	3.88	95.2%
	9	-294908	4.58E+006	8.91E+007	4.41	94.4%
	10	-48278	2.64E+007	5.65E+008	4.63	98.1%
	11	-850908	8.28E+007	7.50E+008	3.01	98.1%
	12	660298	4.37E+007	1.48E+008	1.84	91.2%
6	7	-206770	1.52E+008	8.38E+008	2.34	91.6%
	8	-149155	3.50E+007	4.04E+008	3.40	89.5%
	9	-206216	2.23E+007	4.58E+008	4.53	94.7%
	10	12645	2.04E+005	1.54E+007	8.70	86.1%
	11	35534	5.42E+006	7.22E+007	3.65	89.9%
	15	-335363	2.71E+008	4.50E+009	4.08	89.1%
7	7	-250469	8.09E+006	7.64E+007	3.07	89.9%
	8	12426	1.99E+006	3.26E+007	4.05	85.6%
	9	-80999	7.27E+005	1.50E+007	4.54	85.4%
	12	-49485	-7.20E+004	3.63E+007	ERR	78.4%
	13	-1027979	2.00E+007	4.35E+007	1.48	85.8%
	16	-297012	3.12E+007	8.69E+008	5.28	85.0%
8	3	-336179	1.99E+008	5.71E+009	5.35	91.4%
	4	-291623	7.51E+007	6.17E+008	2.87	86.1%
	14	-12713	4.50E+004	2.87E+006	7.99	81.5%

Table 2: State by Area Type Variance Models - Unemployment Stratum Based Selections

State	Area Type	V_0	V_1	V_2	Sqrt(V_2/V_1)	R^2
1	1	-95505	4.97E+006	7.39E+008	12.19	97.8%
	2	161319	6.44E+007	5.42E+009	9.17	99.3%
	3	-373582	1.48E+008	4.82E+010	18.02	99.5%
	4	-113454	3.00E+008	1.82E+010	7.79	99.1%
	6	10370	1.02E+005	1.22E+008	34.53	97.2%
	7	-267388	6.35E+008	3.10E+010	6.98	99.4%
	8	63539	2.05E+007	1.50E+009	8.56	98.8%
	9	-376991	1.23E+007	7.28E+008	7.69	98.1%
	10	108686	3.69E+007	2.37E+009	8.00	99.5%
	11	20593	1.40E+007	1.69E+009	10.98	98.6%
	12	34131	1.98E+005	2.41E+006	3.49	87.3%
2	1	-262508	1.10E+007	8.91E+008	9.02	95.4%
	2	89061	6.37E+006	6.46E+008	10.07	98.5%
	3	-778667	9.12E+008	7.59E+010	9.12	99.5%
	4	-219973	9.33E+007	9.40E+009	10.03	99.1%
	6	9027	1.71E+006	1.06E+008	7.88	97.6%
	7	-193791	1.29E+008	7.04E+009	7.39	99.3%
	8	-78519	3.99E+006	4.97E+008	11.16	98.4%
	9	-2213	2.35E+006	2.47E+008	10.25	97.3%
	10	8422	4.61E+006	3.27E+008	8.42	97.9%
	11	10997	6.98E+006	1.05E+009	12.24	98.5%
3	2	40129	1.93E+005	5.54E+007	16.95	96.1%
	3	-137473	9.78E+007	6.08E+009	7.88	99.4%
	4	-210292	1.85E+008	9.54E+009	7.18	98.8%
	6	-78224	4.10E+006	1.74E+008	6.52	96.1%
	7	-368617	2.58E+008	1.43E+010	7.45	99.2%
	8	-82169	2.82E+007	2.57E+009	9.54	98.6%
	9	63575	8.82E+006	6.51E+008	8.59	98.8%
	10	-140889	1.07E+007	7.20E+008	8.20	98.9%
	11	-123807	1.50E+007	1.26E+009	9.15	98.8%
	12	114	1.02E+006	3.32E+007	5.70	94.9%
4	3	-58439	9.99E+007	8.00E+009	8.95	98.4%
	4	-120898	1.20E+007	3.77E+009	17.70	97.0%
	6	28453	-1.65E+005	1.41E+007	ERR	90.0%
	7	-73614	5.60E+006	2.95E+008	7.26	97.4%
	8	-50900	1.32E+006	1.68E+008	11.30	98.1%
	9	-13872	5.14E+005	6.95E+007	11.63	97.0%

	10	17358	-3.91E+003	2.75E+007	ERR	95.7%
	11	-33245	2.79E+006	7.95E+007	5.34	98.1%
	12	13110	5.47E+004	9.49E+005	4.17	84.5%
5	3	-27406	1.57E+007	2.36E+009	12.26	99.2%
	4	-224745	3.59E+007	1.20E+010	18.27	99.2%
	6	1985	-2.02E+004	4.13E+007	ERR	94.7%
	7	-85440	2.23E+006	6.57E+008	17.17	98.3%
	8	1048	4.36E+005	4.00E+007	9.57	95.2%
	9	-28904	3.44E+005	1.15E+007	5.78	91.5%
	10	4692	1.18E+006	5.96E+007	7.12	96.6%
	11	-41810	2.88E+006	1.01E+008	5.93	96.0%
	12	15166	5.31E+005	2.07E+007	6.25	90.6%
6	7	38790	3.60E+006	2.51E+008	8.35	92.4%
	8	10715	2.19E+006	6.89E+007	5.61	94.6%
	9	69	1.08E+006	8.80E+007	9.05	94.8%
	10	1096	2.99E+004	2.01E+006	8.20	83.9%
	11	3424	7.94E+005	1.12E+007	3.76	91.9%
	15	-95591	2.20E+007	6.34E+008	5.37	96.6%
7	7	-7562	3.38E+005	1.26E+007	6.11	92.4%
	8	8815	1.29E+005	6.33E+006	7.01	82.6%
	9	-842	1.93E+004	1.46E+006	8.70	88.3%
	12	-4360	7.78E+004	1.51E+006	4.41	89.9%
	13	-64405	4.95E+006	2.83E+006	0.76	94.5%
	16	11585	2.55E+006	8.98E+007	5.94	83.5%
8	3	-6259	8.28E+006	6.87E+008	9.11	90.6%
	4	-5387	3.57E+006	9.91E+007	5.27	86.7%
	14	-2094	5.80E+003	3.19E+005	7.42	83.4%

Attachment B

137. The hybridised national relative variance is a weighting of employment and unemployment relative variances using the 0.9 and 0.1 values determined for the 1996 redesign, and is given by:

$$\sum_{s}\sum_{i}\left(V_{0is}^{(N)}+\frac{V_{1is}^{(N)}}{m_{is}}+\frac{V_{2is}^{(N)}}{m_{is}q_{is}}\right) =$$

$$0.9\,\frac{\sum_{s}\sum_{i}\left(V_{0is}^{E}+\frac{V_{1is}^{E}}{m_{is}}+\frac{V_{2is}^{E}}{m_{is}q_{is}}\right)}{E_{N}^{2}} + 0.1\,\frac{\sum_{s}\sum_{i}\left(V_{0is}^{U}+\frac{V_{1is}^{U}}{m_{is}}+\frac{V_{2is}^{U}}{m_{is}q_{is}}\right)}{U_{N}^{2}}$$

Similarly the hybridised state relative variance is given by:

$$\sum_{i}\left(V_{0is}^{(S)}+\frac{V_{1is}^{(S)}}{m_{is}}+\frac{V_{2is}^{(S)}}{m_{is}q_{is}}\right)=$$

$$0.9\,\frac{\sum_{i}\left(V_{0is}^{E}+\frac{V_{1is}^{E}}{m_{is}}+\frac{V_{2is}^{E}}{m_{is}q_{is}}\right)}{E_{S}^{2}} + 0.1\,\frac{\sum_{i}\left(V_{0is}^{U}+\frac{V_{1is}^{U}}{m_{is}}+\frac{V_{2is}^{U}}{m_{is}q_{is}}\right)}{U_{S}^{2}}$$

where

V_{ris}^{E} = the r'th parameter of the variance model for employment, r = 0,1,2.

V_{ris}^{U} = the r'th parameter of the variance model for unemployment, r = 0,1,2.

E_N = national employment total

E_S = employment total for state S

U_N = national unemployment total

U_S = unemployment total for state S

$$V_{ris}^{(S)} = 0.9\frac{V_{ris}^{E}}{E_s^2} + 0.1\frac{V_{ris}^{U}}{U_s^2}, \quad \text{i} = 0,1,2$$

$$V_{ris}^{(N)} = 0.9\frac{V_{ris}^{E}}{E_N^2} + 0.1\frac{V_{ris}^{U}}{U_N^2}, \quad \text{i} = 0,1,2$$

References

Clark, R.G and Steel, D.G. (2000) Optimum Allocation of sample to strata and stages with simple additional constraints. *University of Wollongong, School of Mathematics and Applied Statistics,* Preprint 3/2000.

Johnson, L.W. and Riess, R.D. (1977) *Numerical Analysis*, Reading, Massachusetts: Addison-Wesley.