KM 40 (1992)
pg.81-95

COLLECTION OF DATA TO DETERMINE TOTALS OVER SUBPOPULATIONS OF UNKNOWN SIZE

Dirk Sikkel *)

ABSTRACT

When in a random sample the size of the target population is unknown, this may cause a considerable extra variance when totals within such a population have to be estimated. Sometimes it is relatively cheap to obtain extra information about the target population. The conditions are analyzed under which it makes sense to obtain such information and it is determined what the optimum allocation of resources is under a linear cost function. Results are derived for a simple random sample and for a stratified random sample in case there are errors in the registration of the strata. An example is given of expenses of firms based on a polluted sampling frame.

Key words: sampling, stratification, optimum allocation

---

*) Research International Nederland, PO Box 1257, 3000 BG Rotterdam

1. The problem

In market research often situations are encountered where the total of a variable must be measured within a target population of unknown size. Obvious examples are buying intentions of bus travellers for a new type of discount ticket or the expenditures on copying machines in a branch of industry where the number of firms is unknown. Of course, in a sample survey it is possible to determine during the interview whether a drawn element belongs to the target population or not. It is a well known fact that the need to establish the size of the target population by the sampling procedure leads to a substantial increase in variance. Let y be the target variable and x the variable that indicates whether an element belongs to the target population. $X_k = 1$ if element k of the population belongs to the target population and $X_k = 0$ otherwise. The universe consists of N elements. Now the fraction of elements in the target population is:

$$P = \sum_{k=1}^{N} X_k / N \; ; \tag{1}$$

the total X = NP; the total of y in the target population is

$$Y_t = \sum_{k=1}^{N} X_k Y_k \, . \tag{2}$$

We estimate Y_t using a simple random sample with replacement of size n. The sample values of x and y are denoted by $\underline{x}_i$ and $\underline{y}_i$, i = 1, 2, ..., n, respectively. The usual, unbiased, estimator for Y_t is

$$\hat{\underline{Y}}_{t0} = \frac{N}{n} \sum_{i=1}^{n} \underline{x}_i \underline{y}_i \, , \tag{3}$$

which has variance

$$S_{t0}^2 = \frac{N^2}{n} P(S_t^2 + Q\bar{Y}_t^2). \tag{4}$$

with $\bar{Y}_t = Y_t/X$, $S_t^2 = \sum X_k(Y_k - \bar{Y}_t)^2/X$ and $Q=1-P$, see e.g. Cochran (1977, ch 2). The term $Q\bar{Y}_t^2$ represents the effect of not knowing X, the size of the target population (if P were known, a sample with nP observations in the target population would have yielded the unbiased estimator $\hat{\underline{Y}} = NP\Sigma_1^{nP}\underline{y}_i/nP$ with variance $S_0^2 = N^2PS_t^2/n$).

Formula (4) represents the variance when the values of x and y are measured together in one sample. Often, however it can be relatively cheap to obtain information only about the value of X. This can be the case in a mixed-mode survey when the value of x can be established by a cheap and simple telephone interview and y has to be measured in an expensive face to face interview. The question then arises how resources should be allocated to the cheap survey mode in which only x is measured and the more expensive survey mode in which both x and y are measured. A similar question, but then concerning response effects in different survey modes is treated in Groves and Lepkovski (1985) and Groves (1989).

In this paper we will first consider the immediate generalization of (4) in the case where also cheap limited interviews are being held only to measure x and solve the allocation problem in case of a simple random sample with a fixed budget. Next we consider a somewhat more complicated problem when we have a stratified population.

2. The simple random case

Let there be an extra sample of size m in which only x is measured. The values of x in the sample are denoted by $\underline{x}_{n+1}, \ldots, \underline{x}_{n+m}$. We then have the following estimator for the sum of the target variable y in the target population:

$$\hat{\underline{Y}}_t = \frac{N}{n+m} \frac{\sum_{i=1}^{n+m} \underline{x}_i}{\sum_{i=1}^{n} \underline{x}_i} \sum_{i=1}^{n} \underline{x}_i \underline{y}_i \tag{5}$$

$\hat{\underline{Y}}_t$ is an unbiased estimator. This can be seen by writing (5) as

$$\hat{\underline{Y}}_t = \frac{N}{n+m} \sum_{i=1}^{n} \underline{x}_i \underline{y}_i \left[1 + \sum_{i=n+1}^{n+m} \underline{x}_i \; / \sum_{i=1}^{n} \underline{x}_i \right] ; \tag{6}$$

now we have $E\Sigma_1^n \underline{x}_i \underline{y}_i = nP\overline{Y}_t$, $E\Sigma_1^n \underline{x}_i \underline{y}_i / \Sigma_1^n \underline{x}_i = \overline{Y}_t$ (cf. Cochran, 1977, pp. 35-36). Furthermore,

$$E\left(\sum_{i=n+1}^{n+m} \underline{x}_i \right) * \left(\sum_{i=1}^{n} \underline{x}_i \underline{y}_i \; / \sum_{i=1}^{n} \underline{x}_i \right) = nP\overline{Y}_t \tag{7}$$

as the first and second factor between brackets of (7) are independent and the separate expectations can be multiplied.

We will derive the leading term of the variance of $\hat{\underline{Y}}_t$ by the delta method (see e.g. Bishop e.a., 1975, p. 486). To this end we write (6) as:

$$\hat{\underline{Y}}_t = \frac{N}{n+m} \varphi(\underline{a}, \underline{b}, \underline{c}) \tag{8}$$

with

$$\varphi(a,b,c) = (1 + \frac{a}{b})c \; , \tag{9}$$

and where

$$\underline{a} = \sum_{i=n+1}^{n+m} \underline{x}_i \; , \tag{10}$$

$$\underline{b} = \sum_{i=1}^{n} \underline{x}_i \ , \tag{11}$$

and

$$\underline{c} = \sum_{i=1}^{n} \underline{x}_i \underline{y}_i \ . \tag{12}$$

Because $\underline{a}$ is independent of $\underline{b}$ and $\underline{c}$, the leading term of the approximation of the variance of $\hat{\underline{Y}}_t$ is

$$\sigma^2(\hat{\underline{Y}}_t) \approx \frac{N^2}{(n+m)^2}\Big(\varphi_a^2(\bar{a},\bar{b},\bar{c})\sigma^2(\underline{a}) + \varphi_b^2(\bar{a},\bar{b},\bar{c})\sigma^2(\underline{b}) + \varphi_c^2(\bar{a},\bar{b},\bar{c})\sigma^2(\underline{c}) + +2\varphi_b(\bar{a},\bar{b},\bar{c})\varphi_c(\bar{a},\bar{b},\bar{c})\text{cov}(\underline{b},\underline{c})\Big) \ , \tag{13}$$

where $\sigma^2(.)$ denotes variance, $\bar{a}$=mP, $\bar{b}$=nP, $\bar{c}=nP\bar{Y}_t$, $\sigma^2(\underline{a})$=mPQ, $\sigma^2(\underline{b})$=nPQ, $\sigma^2(\underline{c})=nPS_t{}^2+nPQ\bar{Y}_t{}^2$ and $\text{cov}(\underline{b},\underline{c})=nPQ\bar{Y}_t$. Substitution into (13) yields

$$\sigma^2(\hat{\underline{Y}}_t) \approx N^2 P\left(\frac{S_t^2}{n} + \frac{Q\bar{Y}_t^2}{n+m}\right) \ . \tag{14}$$

Compared to (4), the interpretation of (14) is clear. The effect of not knowing the size of the target population is diminished by the factor n/(n+m). This, however, does not necessarily mean that it is sensible to allocate resources to this sample of size m to measure x only. This depends on the costs of the different interview procedures. Here we consider the (very plausible) linear cost function. The objective is to minimize (14) under the condition that

$$c_1 n + c_2 m = C \ , \tag{15}$$

where c_1 is the cost of one full interview and c_2 is the cost of an interview in which only x is measured. By writing r=m/n and using (15), the variance (14) is proportional to

$$f(r) = \frac{S_t^2(c_1+c_2 r)}{C} + \frac{QY_t^2(c_1+c_2 r)}{C(1+r)} \tag{16}$$

By setting the first derivative of f to zero, we find

$$(1+r)^2 = Q\,\frac{c_1 - c_2}{c_2 V^2}, \tag{17}$$

where V is the coefficient of variation of y in the target population. Since r can only be positive for the positive root of this equation we have for r

$$r = -1 + \frac{1}{V}\sqrt{Q\left(\frac{c_1}{c_2} - 1\right)} \tag{18}$$

As the second derivative of f equals $2(c_1 - c_2)Q\bar{Y}_t^2/(1+r)^3$, (18) gives the value of r which minimizes the variance. Moreover, (18) makes clear that it is not always useful to allocate resources to a sample in which only x is measured. This is useful only in case r>0, or

$$V^2 < Q\left(\frac{c_1}{c_2} - 1\right) \tag{19}$$

In terms of costs it is clear that the larger the rate c_1/c_2, the more likely that it is useful to conduct the limited interviews. Note that, in order to apply (18) and (19), both V and Q have to be known. In case of repeated surveys these quantities may be estimated from previous measurements. Otherwise, a pilot study is necessary to obtain information about V and Q. Because (19) gives a yes/no criterion (to have limited interviews or not), even rough estimates of V and Q may give useful information for the design of the survey.

3. The stratified case

In this section we will use the above derived results in a slightly more complicated situation, which is encountered in practice by many research institutes. For a survey of expenditures of firms the sampling frame is the register of the Chamber of Commerce. Addresses can be bought from different strata. One can be confident that the population is completely registered. Unfortunately, however, firms are slow to communicate changes in size to the Chamber of Commerce; hence the allocation to the strata is not completely correct. The errors usually concern the sizes of the registered firms. In our context, the strata are the size categories in which the firms are registered. The target populations are the true size categories. We want to make inferences about the expenditures within the true size categories. The true sizes can be established only of those firms which are interviewed.

Let Y_{ij} be the the total of the target variable within the true stratum i which is registered in stratum j. The overall total of the target variable in stratum i is equal to $Y_i=\Sigma_j Y_{ij}$. This total is estimated by $\hat{\underline{Y}}_i=\Sigma_j \hat{\underline{Y}}_{ij}$. According to (14), the variance of $\hat{\underline{Y}}_{ij}$ is equal to

$$\sigma^2(\hat{\underline{Y}}_{ij}) \approx N_j^2 P_{ij} \left(\frac{S_{ij}^2}{n_j} + \frac{Q_{ij}\bar{Y}_{ij}^2}{n_j+m_j} \right) . \qquad (20)$$

Here, P_{ij} is the probability that a randomly drawn firm from stratum j in reality belongs to stratum i; $Q_{ij}=1-P_{ij}$; n_j and m_j are the sizes of the samples with unrestricted and restricted interviews, respectively. S_{ij}^2 is the variance of Y within the part of stratum i which is registered in j. The variance of $\hat{\underline{Y}}_i$ is equal to $\Sigma_j \sigma^2(\hat{\underline{Y}}_{ij})$; the size of the registered stratum j is equal to N_j. We now want to choose n_j and m_j in such a way that a reasonable loss function is minimized. Different loss functions are conceivable, e.g. the sum of standard errors of $\hat{\underline{Y}}_i$ or the maximum variance of each individual estimator $\hat{\underline{Y}}_i$; the best choice of loss function depends on the purpose of the survey involved, but has a subjective element because there is no single criterion to be optimized as the precision of every single $\hat{\underline{Y}}_i$ is of importance.

The loss function we consider here is a weighted sum of the variances of the estimators $\hat{\underline{Y}}_{ij}/N_j$. The importance of a stratum is considered to be proportional to its size, but is not necessary to estimate the total within a large stratum with the same absolute precision as the total within a small stratum (it is implicitly assumed that the stratum sizes are proportional to the totals). We denote the size of the true stratum i by N_{ti}. We may not know exactly the sizes of the true strata, but when the number of errors is small or the errors are randomly spread, the sizes of the registered strata (N_j) may be taken as to approximate the true sizes. We may consider $W_i = N_{ti}/N$ (N is the total population size) to be a measure of importance of registered stratum i. Hence

$$\ell(\mathbf{n},\mathbf{m}) = \sum_i \sum_j W_i P_{ij} \left(\frac{S^2_{ij}}{n_j} + \frac{Q_{ij}\bar{Y}^2_{ij}}{n_j + m_j} \right) . \qquad (21)$$

is a plausible loss function. The vectors $(n_1, \ldots, n_K)'$ and $(m_1, \ldots, m_K)'$ are denoted by $\mathbf{n}$ and $\mathbf{m}$, respectively. Again, it will be more convenient to work with r_j instead of m_j, so we write $m_j = r_j n_j$, $j = 1, 2, \ldots, K$ and $\mathbf{r}$ for the vector $(r_1, \ldots, r_K)$. The budget restriction then becomes

$$\sum_j (c_{1j} n_j + c_{2j} r_j n_j) = C . \qquad (22)$$

By adding restriction (22) to the loss function ℓ, we obtain a new loss function $L(\mathbf{n},\mathbf{r}) = \ell + \lambda(\Sigma(c_{1j}n_j + c_{2j}r_j n_j) - C)$, where λ is a Lagrange multiplier. Now let $\alpha_j = \Sigma_i W_i P_{ij} S^2_{ij}$ and $\beta_j = \Sigma_i W_i P_{ij} Q_{ij} \bar{Y}^2_{ij}$. Then the loss function ℓ can be written as $\Sigma\{\alpha_j/n_j + \beta_j/(n_j(1+r_j))\}$. We will derive all theory in terms of α_j and β_j, so the particular choice of the loss function (21) is not a very critical assumption. The partial derivatives of L are (in terms of α_j and β_j)

$$\frac{\partial L}{\partial n_j} = \frac{-1}{n_j^2} \{\alpha_j + \beta_j/(1+r_j)\} + \lambda(c_{1j} + c_{2j} r_j) \qquad (23)$$

and

$$\frac{\partial L}{\partial r_j} = \frac{-1}{n_j(1+r_j)^2}\beta_j + \lambda c_{2j} n_j \tag{24}$$

By setting the partial derivatives to zero and solving for **r**, **n** and λ we find

$$r_j = -1 + \sqrt{\frac{\beta_j}{\alpha_j}\left(\frac{c_{1j}}{c_{2j}} - 1\right)} \tag{25}$$

and

$$n_j = \sqrt{\frac{\alpha_j}{(c_{1j}-c_{2j})\lambda}}\ , \tag{26}$$

where λ is the normalizing constant which satisfies

$$\sqrt{\lambda} = \frac{1}{C}\sum_j \sqrt{\alpha_j(c_{1j}-c_{2j})} + \frac{1}{C}\sum_j \sqrt{\beta_j c_{2j}} \tag{27}$$

As is shown in the appendix, inspection of the matrix of second order derivatives shows that (25), (26) and (27) do minimize L. This solution can be used when all r_j are positive or zero. Negative values of r_j have no practical interpretation. From (25), however, it follows that r_j depends only on the given α_j, β_j and the proportion of the costs c_{1j}/c_{2j} and *not* on C and λ. This suggests the following procedure (which is justified more precisely in the appendix) in case some of the r_j are negative. Define for those j for which r_j is negative, c_{2j}^* to be the cost per unit in the sample which provides for the extra population estimates such that for c_{2j}^* instead of c_{2j} the optimum value of r_j would be zero. For such j we have

$$c_{2j}^* = \frac{\beta_j c_{1j}}{\alpha_j+\beta_j} \tag{28}$$

Insertion of the c_{2j}^* at the appropriate places into (22) will yield the optimum values of r_j and n_j.

4. An example

In 1987 a sample from the data set of the Chamber of Commerce was bought by Research International Nederland. During the interview it was determined whether a firm belonged to the stratum in which it was registered or belonged to another stratum. For the category "industry/construction" this produced the results that are given in table 1. In table 2 the registered stratum sizes are given and the sample means and standard deviations of a typical variable that describes expenditures with respect to office equipment Note that the coefficients of variation are rather high, which is characteristic for financial data. The category "100-199 employees" has the smallest coefficient of variation, "200-499 employees" has the largest.

The data are analyzed as if the sample values represent the true scores. We assume that means and standard deviations are constant within the true strata over the registered strata. We also assume that the proportion of costs c_1/c_2 is constant over the strata. In figure 1, r is analyzed as a function of c_1/c_2. The intersection with the line r=0 corresponds with the proportion c_1/c_2 where it becomes profitable to conduct extra interviews to determine the size of the target population. For the category "100-199 employees" this intersection corresponds with $c_1/c_2=2$, i.e. when an interview in which only the membership of the target population is determined, costs less than one half of the costs of a complete interview, then such limited interviews are profitable. For the category "1-9 employees" it is a very different matter. Limited interviews are profitable only when complete interviews are 6 times more expensive. The difference between these zero-points is partly due to the coefficient of variation and partly to the fact that the smallest stratum is relatively well registered. The categories which are analyzed in figure 1 are the extreme cases. The zero-points of the other categories are between these extremes.

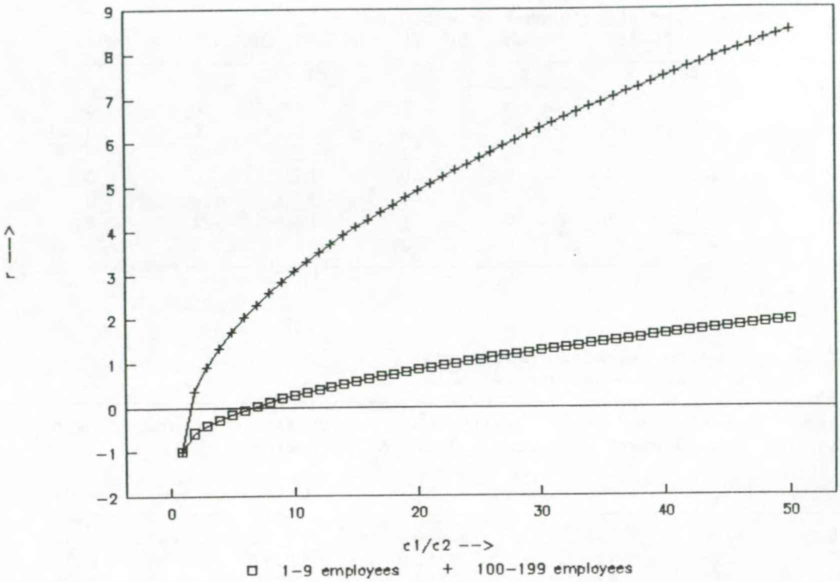
Table 1. True firm size by registered firm size for industry/construction

true size	registered size: number of employees 1- 9	10- 19	20- 49	50- 99	100-199	200-499	500+
1- 9	84	9	2	2	0	0	0
10- 19	7	66	20	3	0	0	0
20- 49	1	3	40	12	1	1	0
50- 99	1	0	2	40	15	0	0
100-199	0	2	3	6	37	12	0
200-499	2	1	0	1	10	38	8
500+	0	0	0	2	3	6	26

Table 2. Sample means, standard deviations of a typical expenditure variable

registered size	mean $\bar{Y}_j$	standard deviation S_j	registered stratum size N_j	coefficient of variation V_j	alpha α_j	beta β_j
1- 9	170	361	55664	2.12	8.94	1.58
10- 19	215	424	5366	1.97	2.30	1.91
20- 49	313	874	6521	2.79	6.45	3.53
50- 99	284	706	2118	2.48	2.05	1.48
100-199	847	643	1129	0.76	0.88	1.61
200-499	1024	4168	551	4.07	61.98	57.46
500+	1451	3510	189	2.42	7.40	6.21

Figure 1. r as a function of c1/c2



Finally, the optimum values of n_j are given in the above example, given a fixed budget C. We assume that the proportion c_1/c_2 is equal to 15. According to table 3 we find that for the category "1-9" the value of r_j is negative. Before we can calculate the optima for n_j we have to calculate c^*_{2j} for this category. The values of n_j are given as percentages of $\Sigma_j n_j$ (these percentages are constant for varying values of C). The values of m_j are also given as percentages of $\Sigma_j n_j$, the total sample size of the complete interviews. In this example, the proportion of limited interviews is rather small, which is due to the high coefficients of variation. It is to be expected that for more homogeneous target populations it is more profitable to conduct the limited interviews.

Table 3. Values of r_j, c^*_{2j} and optimum n_j, m_j for $c_1=5$, $c_2=1$

registered size	r_j	c^*_{2j}	n_j %	m_j %
1- 9	-0.1586	0.76	14.6	
10- 19	0.8202		7.6	6.2
20- 49	0.4802		12.7	6.1
50- 99	0.7001		7.1	5.0
100-199	1.7128		4.7	8.0
200-499	0.9257		39.5	35.5
500+	0.8318		13.7	11.4

5. Conclusion

Limited information about sizes of target populations often can be obtained very cheaply, e.g. by a screening by telephone. It depends, however, on a number of parameters whether it is profitable to collect such information. An important parameter is the coefficient of variation in the target population; the higher this quantity, the less likely that limited information can be obtained profitably. A second important parameter is the relative size of the target population. The larger this size, the less useful it is to obtain limited information.

References

Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1975, Discrete multivariate analysis. MIT Press, Cambridge.

Cochran, W.G., 1977, Sampling Techniques, third edition. Wiley, New York.

Courant, R., 1970, Differential and Integral Calculus, vol II. Blackie, London.

Groves, R.M. and J.M. Lepkovski, 1985, Dual Frame, Mixed Mode Survey Designs. Journal of Official Statistics, vol. 1, no. 3, pp. 263-286.

Groves, R.M., 1989, Survey Errors and Survey Costs. Wiley, New York.

ontvangen 23- 1 -1991
geaccepteerd 3 - 4 -1992

Appendix

A.1 Second order derivatives in the stratified case

In this section we show that the stationary points defined by (25) and (26) correspond to a maximum of the loss function L. First, we observe that the second order derivatives $\partial^2 L/\partial n_j \partial n_k$, $\partial^2 L/\partial r_j \partial r_k$ and $\partial^2 L/\partial n_j \partial r_k$ for $j \neq k$ are equal to zero; in other words: the matrix of second order derivatives is block-diagonal. A sufficient condition for obtaining a maximum is that this matrix is positive definite, which is the case when each of the blocks is positive definite. Now let us look at the second order derivatives within the block corresponding to index j.

$$\frac{\partial^2 L}{\partial n_j{}^2} = \frac{2}{n_j{}^3} (\alpha_j + \beta_j/(1+r_j)) \ , \tag{A.1}$$

$$\frac{\partial^2 L}{\partial r_j{}^2} = \frac{2\beta_j}{n_j (1+r_j)^3} \ , \tag{A.2}$$

$$\frac{\partial^2 L}{\partial n_j \partial r_j} = \frac{\beta_j}{n_j{}^2 (1+r_j)^2} + \lambda c_{2j} \tag{A.3}$$

A necessary and sufficient condition for this block to be positive definite is that

$$\frac{\partial^2 L}{\partial n_j{}^2} \cdot \frac{\partial^2 L}{\partial r_j{}^2} - \left(\frac{\partial^2 L}{\partial n_j \partial r_j}\right)^2 > 0 \tag{A.4}$$

in the stationary points as defined by (25) and (26) (see e.g. Courant, 1970). Substitution of (A.1), (A.2), (A.3), (25) and (26) into (A.4) yield after (a lot of) calculation that the left hand side of (A.4) is equal to

$$4c_{2j} \lambda^2 \sqrt{\frac{\alpha_j c_{2j} (c_{1j} - c_{2j})}{\beta_j}} \ ,$$

which is greater than zero for every value of λ, provided that $c_{1j} > c_{2j}$ and $c_{2j} > 0$, which both are plausible conditions.

A.2. Solutions for the stratified case when some of the r_j are negative

When some of the r_j are negative, it was suggested at the end of section 3 that the costs c_{2j} should be changed to c^*_{2j}, such that the corresponding r_j would be zero. It remains, however, to be proved that this yields the optimum solution in terms of n_j and λ. In fact, the problem is to optimize (20), not only under the budget restriction, but also under the condition that for all j $r_j \geq 0$. First we define the sets $R_+ = \{j | r_j \geq 0\}$ and $R_- = \{j | r_j < 0\}$ when no conditions are imposed on r_j. Now we reformulate the optimization problem as to minimize

$$\ell(\mathbf{n},\mathbf{r}) = \sum_{j \in R_+} \{\alpha_j / n_j + \beta_j / (n_j (1+r_j))\} + \sum_{j \in R_-} (\alpha_j + \beta_j)/n_j \quad , \tag{A.5}$$

under the condition

$$\sum_{j \in R_+} (c_{1j} n_j + c_{2j} r_j n_j) + \sum_{j \in R_-} c_{1j} n_j = C \tag{A.6}$$

implying $r_j = 0$ for $j \in R_-$. The solution of this problem is:

$$n_j = \begin{cases} \sqrt{\dfrac{\alpha_j}{(c_{1j} - c_{2j})\lambda}} & \text{for } j \in R_+ \\ \sqrt{\dfrac{\alpha_j + \beta_j}{c_{1j}\lambda}} & \text{for } j \in R_- \end{cases} \tag{A.7}$$

and

$$\sqrt{\lambda} = \frac{1}{C}\left[\sum_{j \in R_+} \left(\sqrt{\alpha_j (c_{1j} - c_{2j})} + \sqrt{\beta_j c_{2j}}\right) + \sum_{j \in R_-} \sqrt{c_{1j} (\alpha_j + \beta_j)}\right] \tag{A.8}$$

It is easily verified that the original problem with c^*_{2j} instead of c_{2j} has the same solution. The remaining question is whether this solution is a maximum. Now let us look at the partial derivatives $\partial L / \partial r_j$ for $j \in R_-$ as given by (24). This derivative is monotonically increasing with r_j. For the n_j from the restricted solution we have

$$\left.\frac{\partial L}{\partial r_j}\right|_{r_j = 0} = \frac{-\beta_j}{\sqrt{\dfrac{\alpha_j + \beta_j}{\lambda c_{1j}}}} + \lambda c_{2j} \sqrt{\frac{\alpha_j + \beta_j}{\lambda c_{1j}}}$$

$$= - \frac{\alpha_j c_{2j} \sqrt{\lambda}}{\sqrt{c_{1j} (\alpha_j + \beta_j)}} \times \left[\frac{\beta_j}{\alpha_j}\left(\frac{c_{1j}}{c_{2j}} - 1\right) - 1\right] . \tag{A.9}$$

Now from $j \in R_-$ it follows that, according to (25), $(\beta_j/\alpha_j)(c_{1j}/c_{2j}-1)<1$, hence the second factor in (A.9) is negative, hence L increases in r_j for $r_j \geq 0$ and for the optimum values of n_j. This proves that at the border $r_j=0$ for $j \in R_-$ there is at least a local minimum.