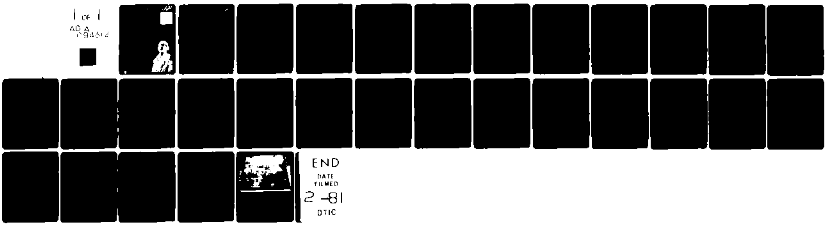
AD-A094 312 GEORGE WASHINGTON UNIV WASHINGTON D C PROGRAM IN LOG--ETC F/G 12/1
KALMAN FILTER TECHNIQUES FOR CONTROL OF REPEATED ECONOMIC SURVE--ETC(U)
SEP 80 W SMITH, Z BARZILY N00014-75-C-0729
UNCLASSIFIED SERIAL-T-428 NL

1 OF 1
AD A
094312



END
DATE
FILMED
2-81
DTIC

LEVEL II

12

AD A094312

THE
GEORGE
WASHINGTON
UNIVERSITY

STUDENTS FACULTY STUDY R
ESEARCH DEVELOPMENT FUT
URE CAREER CREATIVITY CO
MMUNITY LEADERSHIP TECH
NOLOGY FRONTI ESIGN
ENGINEERING APP ENC
GEORGE WASHIN NIV

81 1 30 001

SCHOOL OF ENGINEERING
AND APPLIED SCIENCE

LEVEL II

12

(6) KALMAN FILTER TECHNIQUES FOR CONTROL OF REPEATED ECONOMIC SURVEYS.

by

(10) Wray Smith
Zeev Barzily

(9) Scientific rept.

(14) SERIAL-T-

Serial-T-428
(11) 30 September 1980

(12) 28

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

(15) Program in Logistics
Contract N00014-75-C-0729
Project NR 347 020
Office of Naval Research

405337

NONE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER T-428	2. GOVT ACCESSION NO. AD-A094 312	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) KALMAN FILTER TECHNIQUES FOR CONTROL OF REPEATED ECONOMIC SURVEYS		5. TYPE OF REPORT & PERIOD COVERED SCIENTIFIC 6. PERFORMING ORG. REPORT NUMBER T-428
7. AUTHOR(s) WRAY SMITH ZEEV BARZILY		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0729
9. PERFORMING ORGANIZATION NAME AND ADDRESS THE GEORGE WASHINGTON UNIVERSITY PROGRAM IN LOGISTICS WASHINGTON, DC 20052		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH, CODE 434 DEPARTMENT OF THE NAVY ARLINGTON, VA 22217		12. REPORT DATE 30 SEPTEMBER 1980 13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) NONE 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC SALE AND RELEASE; DISTRIBUTION IS UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

INVENTORY ANALYSIS
MEASUREMENT ERROR
COST OF OBSERVATION
OPTIMAL CONTROL
RANDOM-YIELD SAMPLING
SURVEY TIMING

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we discuss the determination of sample sizes and inter-survey intervals in repeated economic surveys. Three models are discussed. The first two maintain the intersurvey intervals at a constant length while the third allows the intersurvey intervals to vary in length. In the first model we assume that the number of obtained observations in each survey is equal to the number of designated observations. The second and third models

(continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

NONE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

assume that the number of obtained observations is a random variable. The costs which are taken into account are the fixed and varying costs of surveying and a cost which is due to the use of estimates from the repeated surveys.

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

Program in Logistics

Abstract
of
Serial T-428
30 September 1980

KALMAN FILTER TECHNIQUES FOR CONTROL
OF REPEATED ECONOMIC SURVEYS

by

Wray Smith
Zeev Barzily

In this paper we discuss the determination of sample sizes and intersurvey intervals in repeated economic surveys. Three models are discussed. The first two maintain the intersurvey intervals at a constant length while the third allows the intersurvey intervals to vary in length. In the first model we assume that the number of obtained observations in each survey is equal to the number of designated observations. The second and third models assume that the number of obtained observations is a random variable. The costs which are taken into account are the fixed and varying costs of surveying and a cost which is due to the use of estimates from the repeated surveys.

Research Sponsored by
Office of Naval Research
Contract N00014-75-C-0729
Project NR 347 020

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

Program in Logistics

KALMAN FILTER TECHNIQUES FOR CONTROL OF REPEATED ECONOMIC SURVEYS*

by

Wray Smith
Zeev Barzily

1. Introduction and Background

Statistical data from household surveys and administrative records are widely used in formulas to allocate grant-in-aid funds or other governmental benefits to state and local jurisdictions. A review of U.S. programs appears in Gonzalez (1978). An explicit tradeoff between data collection costs in repeated surveys and an imputed cost of misallocation of resources due to the use of imprecise survey data in setting allocation levels was considered in Smith and Zalkind (1978) as a scalar problem in deterministic inventory theory. The present paper places such problems in a control theory framework accommodating vector linear models of nonstationary economic processes measured by noisy multi-item repeated surveys. Optimal rules are found for control of a sequence of surveys, where there is a fixed charge plus unit costs of

---

*An earlier version of this paper was presented at the 2nd Economics and Control Conference, Princeton University, Princeton, New Jersey, June 2-4, 1980.

surveying and an imputed quadratic loss associated with the imprecision of the resulting state estimates of the observed economic process.

In this paper we analyze one model in which the number of completed interviews in a survey is equal to the designated number of interviews and two models in which the number of completed interviews is a random variable. The models are analyzed with the aid of an equivalent sample size form of the Kalman filter that was derived in Smith (1979). A related inventory analysis of single-item repeated surveys under random-yield sampling is set forth in Smith and Zalkind (1980). Methods for time series analysis of repeated surveys are presented in Scott and Smith (1974). Problems of the choice of loss function for misallocation error are of practical importance, but they are not treated in the present paper, nor are problems of "distributive equity" in allocations across states or local jurisdictions, as discussed in Spencer (1979). In the models of this paper, we are estimating (tracking but not controlling) the state of the observed economic process, and we are controlling the survey measurement subsystem; see Meier, Peschon, and Dressler (1967). Also see Aoki and Li (1968), which discusses control problems with cost for observation, although not treating the fixed-plus-variable cost structure of the present paper.

2. Formulation of the Models

In this section we formulate the models to be discussed in the present paper. In the first two models (Sections 3 and 4) we are interested in the evolution of $\underset{\sim}{x}(j)$, a multivariate socioeconomic process (e.g., money income or proportion of a population in poverty), through discrete time points j , $j=0,1,2,\ldots$. Here $\underset{\sim}{x}(j)$ is an $m \times 1$ state vector, and we assume that it evolves according to the vector random walk

$$\underset{\sim}{x}(j+1) = \underset{\sim}{x}(j) + \underset{\sim}{w}(j+1) , \qquad (1)$$

where $\underset{\sim}{w}(j) \sim N(\underset{\sim}{0},\underset{\sim}{Q})$. That is, $\underset{\sim}{w}(j)$ has a multivariate normal distribution with a zero mean vector and known process disturbance covariance matrix $\underset{\sim}{Q}$. Surveys of the process $\underset{\sim}{x}(j)$ are taken once every T units of time at $k = T$, $2T$, $3T$, ..., rT, ... (In one of our stochastic models, T may vary.) We also assume that $\underset{\sim}{x}(0) \sim N(E[\underset{\sim}{x}(0)]\ ,\ \underset{\sim}{C}(0))$, where $E[\underset{\sim}{x}(0)]$ and $C(0)$ are known.

At each survey time $n_d(k)$ individuals are sampled, but only $N(k)$ completed interviews are obtained. We assume that $P[N(k) \geq 1] = 1$. The difference $n_d(k) - N(k)$ is due to refusals to respond or incomplete observations. We assume that $\underset{\sim}{y}(k)$, the m-dimensional survey measurement, is represented by the survey equation

$$\underset{\sim}{y}(k) \quad = \quad \underset{\sim}{x}(k) + (1/N(k)) \sum_{i=0}^{N(k)} \underset{\sim}{u}_i(k) \ , \tag{2}$$

where each $\underset{\sim}{u}_i(k)$ is an $m \times 1$ vector that is distributed normally with a zero mean and measurement noise covariance matrix $\underset{\sim}{R}$. The matrix $\underset{\sim}{R}$ is assumed to be known, time invariant, symmetric, and positive definite. Both $\underset{\sim}{w}(j)$ and the $\underset{\sim}{u}_i(k)$ are assumed to be serially uncorrelated.

We may find the best linear estimate of $\underset{\sim}{x}(rT+j)$, given $E[\underset{\sim}{x}(0)]$, $\underset{\sim}{C}(0)$, and the r measurements $y^r = (\underset{\sim}{y}(T), \underset{\sim}{y}(2T), \ldots, \underset{\sim}{y}(rT))$, as a linear combination of the observations y^r . We define the estimation error $\underset{\sim}{x}^e(k+j|k)$ as

$$\underset{\sim}{x}^e(k+j|k) \quad = \quad \underset{\sim}{x}(k+j) - \hat{\underset{\sim}{x}}(k+j|k) \ ,$$

where

$$\hat{\underset{\sim}{x}}(k+j|k) = E[\underset{\sim}{x}(k+j)|y^r] \ .$$

We also define the j-step-ahead estimation error covariance matrix $\underset{\sim}{C}(k+j|k)$ as the conditional expectation of the outer product of the j-step-ahead estimation error vector given the r measurements; namely,

$$\underset{\sim}{C}(k+j|k) = E[\underset{\sim}{x}^e(k+j|k)\ (\underset{\sim}{x}^e(k+j|k))'] \ . \tag{3}$$

We now state a general optimal filter theorem for the vector random walk models. Related proofs and derivations of Kalman-type filters for vector models may be found in Jazwinski (1970), Melsa and Cohn (1978), or Sage and White (1977).

Optimal Filter Theorem for Vector Random Walk Models. The optimal (minimum variance) filter for the discrete system (1), (2) consists of difference equations for the conditional mean $\hat{\underset{\sim}{x}}(k+j|k)$ and the estimation error covariance matrix $\underset{\sim}{C}(k+j|k)$, for $T=1,2,\ldots,T_{max}$:

between surveys,

$$\begin{aligned} \hat{\underset{\sim}{x}}(k+j|k) &= \hat{\underset{\sim}{x}}(k|k) \ , \quad \text{for} \quad k=T,2T,\ldots \ , \text{ and } \ j=1,\ldots,T \ , \\ \underset{\sim}{C}(k+j|k) &= \underset{\sim}{C}(k|k) + j\underset{\sim}{Q} \ ; \end{aligned} \tag{4}$$

at surveys

$$\begin{aligned} \hat{\underset{\sim}{x}}(k|k) &= \hat{\underset{\sim}{x}}(k|k-T) + \underset{\sim}{K}(k)\,[\underset{\sim}{y}(k) - \hat{\underset{\sim}{x}}(k|k-T)] \ , \\ \underset{\sim}{C}(k|k) &= [\underset{\sim}{I} - \underset{\sim}{K}(k)]\,\underset{\sim}{C}(k|k-T) \ , \end{aligned} \tag{5}$$

where $\underset{\sim}{K}(k)$ is the Kalman gain

$$\underset{\sim}{K}(k) = \underset{\sim}{C}(k|k-T)\,[\underset{\sim}{C}(k|k-T) + \underset{\sim}{B}(k)]^{-1} \ , \tag{6}$$

and $\underset{\sim}{B}(k)$ is the sample noise covariance matrix obtained from $\underset{\sim}{R}$ by dividing each of its elements by $N(k)$. The second equations in (4) and (5) are called the error covariance equations.

The first two models differ in the number of observations obtained. In Model A we assume that the number of observations in each survey is a known constant $n_d(k) = n_d$. In other words,

$$P[N(k) = n_d] = 1 \ .$$

Model B assumes that a constant scalar designated sample size n_d is to be used for all survey times, that the obtained sample size is a random variable with expectation θn_d , where $0 < \theta < 1$ is known, and that the inter-survey intervals T are nonvarying.

The third model (Model C) involves a continuous-time process--a Brownian motion process. Here we assume that the sample sizes obtained are i.i.d. random variables and we allow the intersurvey periods to vary as a function of the sample sizes obtained. The remaining assumptions and the analysis parallel those of the first two models.

In all three models we are interested in determining $n_d(k)$, or n_d, the designated sample size to be used at time k. We determine the $n_d(k)$ that minimizes a cost per unit time function J. We define J by

$$J = E[(1/T)(c_0 + c_1 n_d(k) + c_2 N(k)) + L_{avg}] , \tag{7}$$

where c_0 is the fixed start-up cost of ordering a survey, c_1 is a unit cost per designated item-interview (for m items), and c_2 is the additional unit cost per obtained item-interview. Letting a_1 be a loss weighting coefficient set by the decision maker, L_{avg} is defined by

$$\begin{aligned} L_{avg} &= a_1 \text{ trace } [(1/T) \sum_{j=0}^{T-1} \underset{\sim}{C}(k+j|k)] \\ &= a_1 \text{ trace } [\underset{\sim}{C}(k|k) + ((T-1)/2)\underset{\sim}{Q}] . \end{aligned} \tag{8}$$

The L_{avg} is an imputed loss attributed to the use of survey estimates of the state of the process and is proportional to the sum of the estimation error variances. By (6), $\underset{\sim}{C}(k|k)$ is dependent on the realization of $N(k)$ through $\underset{\sim}{K}(k)$.

3. Model A

We specify for Model A that there is a sequence of m-item surveys of one homogeneous population with obtained sample sizes $N(k) = n_d(k)$. The filter gain $\underset{\sim}{K}(k)$ of (6) may be written as

$$\underset{\sim}{K}(k) = \underset{\sim}{C}(k|k-T) [\underset{\sim}{C}(k|k-T) + \underset{\sim}{R}/n_d(k)]^{-1} .$$

Let us now introduce the concept of *equivalent sample size* and show how the error covariance equations in the optimal filter theorem may be usefully analyzed using this concept. Let the matrix $\underset{\sim}{N}_0(k+j|k)$ be defined by

$$\underset{\sim}{N}_0(k+j|k) = \underset{\sim}{R}^{1/2}\, \underset{\sim}{C}^{-1}(k+j|k)\, \underset{\sim}{R}^{1/2} , \quad j=0,1,\ldots,T-1 , \tag{9}$$

where $\underset{\sim}{R}^{1/2}\underset{\sim}{R}^{1/2} = \underset{\sim}{R}$. The Kalman gain may then be written as

$$\underset{\sim}{K}(k) = \underset{\sim}{R}^{1/2}\, \underset{\sim}{N}_0^{-1}(k|k-T)\, \underset{\sim}{R}^{1/2}\, [\underset{\sim}{R}^{1/2}\, \underset{\sim}{N}_0^{-1}(k|k-T)\, \underset{\sim}{R}^{1/2} + \underset{\sim}{R}/n_d(k)]^{-1} , \tag{10}$$

and thus we obtain a recursive relation in $\underset{\sim}{N}_0(k|k)$,

$$\underset{\sim}{N}_0(k|k) = n_d(k)\underset{\sim}{I} + \underset{\sim}{N}_0(k-T|k-T)\, [\underset{\sim}{I} + T\underset{\sim}{R}^{-1/2}\, \underset{\sim}{Q}\, \underset{\sim}{R}^{-1/2}\, \underset{\sim}{N}_0(k-T|k-T)]^{-1} . \tag{11}$$

If only one item is surveyed (m=1), then Q and R (and K) are scalars, the matrix $\underset{\sim}{N}_0(k|k-T)$ is replaced by the scalar $N_0(k|k-T)$, and

$$K(k) = n_d(k) / [N_0(k|k-T) + n_d(k)]$$

and

$$N_0(k|k) = n_d(k) + N_0(k|k-T) .$$

This development leads to equivalent sample size relations to replace the error variance equations in the optimal filter theorem for the scalar case of Model A:

between surveys,

$$N_0(k+j|k) = N_0(k|k)\, [1 + jQR^{-1}\, N_0(k|k)]^{-1} ;$$

at surveys,

$$N_0(k|k) = n_d(k) + N_0(k-T|k-T)\, [1 + TQR^{-1}\, N_0(k-T|k-T)]^{-1} .$$

Here $N_0(k+j|k)$ is the equivalent sample size remaining at time $k+j$, j time units after ordering new stock. If the system is in steady state, $N_0(k|k-T)$ or $N_0(k+T|k)$ may be interpreted in inventory terms as the "reorder point." Since $n_d(k)$ is the sample size of the survey conducted at time k, we may interpret $N_0(k|k)$ as the size of a survey that would be required to obtain the same degree of precision of estimate at time k as that provided by an optimally combined data set of size $n_d(k) + N_0(k|k-T)$.

Figure 1 depicts the pattern of deterministic decay and replenishment of equivalent sample size for the scalar case of Model A.

Returning to the case where $\underset{\sim}{Q}$ and $\underset{\sim}{R}$ are matrices, using (8), (5), and (9) in (7) we obtain

$$\begin{aligned} J(N_0(k|k),T) &= (1/T)[c_0 + c_1 m\, n_d(k)] + a_1 \text{ trace } [\underset{\sim}{N}_0^{-1}(k|k)\underset{\sim}{R} + ((T-1)/2)\underset{\sim}{Q}] \\ &= (1/T)\Big[c_0 + c_1 \text{ trace } [\underset{\sim}{N}_0(k|k) - N_0(k-T|k-T) \cdot \\ &\qquad (\underset{\sim}{1} + T\underset{\sim}{Q}\underset{\sim}{R}^{-1}\underset{\sim}{N}_0(k-T|k-T))^{-1}]\Big] \\ &\qquad + a_1 \text{ trace } [\underset{\sim}{N}_0^{-1}(k|k)\underset{\sim}{R} + ((T-1)/2)\underset{\sim}{Q}] \;. \end{aligned} \tag{12}$$

The optimal pair $(n_d^*(k), T^*)$ or, equivalently, $(\underset{\sim}{N}_0^*(k|k), T^*)$, can easily be determined, since for our survey systems there always exists a T, denoted T_{max}, beyond which it is not possible to lengthen the intersurvey interval T without violating $1 \le \text{trace}[\underset{\sim}{N}_0(k+T|k)]$. Thus we calculate $(n_d^*(k), T^*)$ by setting T at successive integer values, $T=1,2,\ldots,T_{max}$ and computing the value of the sample size $n_d(k)$ which minimizes J for each T. We then adopt the T corresponding to the minimum of J over T.

It can be shown that when k is large the error covariance matrices $\underset{\sim}{C}(k|k)$, $\underset{\sim}{C}(k+T|k+T)$, ... , will approach a limit, say $\underset{\sim}{C}$. Equivalently, $\underset{\sim}{N}_0(k|k)$ will approach a steady-state matrix $\underset{\sim}{N}_0$. If the system is nearing steady state, then $\underset{\sim}{C}(k|k-T)$ will also approach a limit. From the covariance equation in (4), we then have

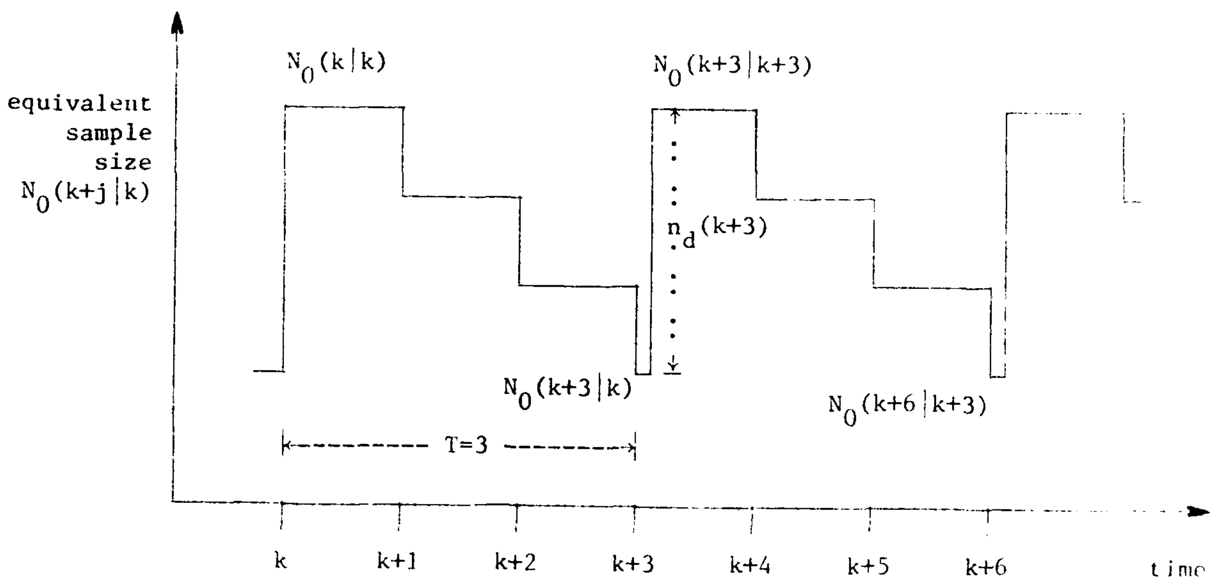


Figure 1.--Deterministic decay and replenishment of equivalent sample size for Model A with $m=1$.

$$C(k|k-T) \rightarrow C + TQ \; .$$

The steady-state Kalman gain K will be

$$K = (C + TQ)\,(C + TQ + R/n_d)^{-1} \; ,$$

so that

$$C = (I - K)\,(C + TQ) \; .$$

That is,

$$\begin{aligned} TQ &= K(C + TQ) \\ &= (C + TQ)\,(C + TQ + R/n_d)^{-1}\,(C + TQ) \; , \end{aligned}$$

or

$$C\,Q^{-1}\,C + T\,C - (T/n_d)\,R = 0 \; . \qquad (13)$$

We solve (13) for C, obtaining

$$C = -(T/2)Q + Q^{1/2}\,[(T/n)Q^{-1/2}\,R\,Q^{-1/2} + (T^2/4)I]^{1/2}\,Q^{1/2} \; . \qquad (14)$$

<u>A numerical example of Model A</u>

Suppose for $m=2$ the cost coefficients are

$$c_0 = \$10^6 \; , \quad c_1 = \$67.50 \; , \quad c_2 = 0 \; , \quad \text{and} \quad a_1 = 2\times10^9 \; ,$$

and the Q and R noise covariance matrices have the following numerical entries:

$$Q = \begin{bmatrix} .00005 & .00004 \\ .00004 & .00007 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 0.10 & 0.05 \\ 0.05 & 0.25 \end{bmatrix} .$$

Let D be the matrix of eigenvalues of $Q^{-1/2}\,R\,Q^{-1/2}$,

$$Q^{-1/2}\,R\,Q^{-1/2} = \begin{bmatrix} 3140 & -1978 \\ -1978 & 5018 \end{bmatrix} ; \quad \text{then} \quad D = \begin{bmatrix} 1889 & 0 \\ 0 & 6269 \end{bmatrix} .$$

We can construct an orthogonal matrix V whose columns are normalized eigenvectors of $Q^{-1/2}\,R\,Q^{-1/2}$, so that $Q^{-1/2}\,R\,Q^{-1/2} = V\,D\,V'$. We will also have need of the matrix $\underset{\sim}{S} = \underset{\sim}{V}'\, \underset{\sim}{Q}\, \underset{\sim}{V}$. For the numerical values above,

$$\underset{\sim}{V} = \begin{bmatrix} .8453 & -.5344 \\ .5344 & .8453 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{S} = \begin{bmatrix} .00009184 & .00002619 \\ .00002619 & .00002816 \end{bmatrix} .$$

We will find it convenient to make the change of variable $z = T/n_d$, where $0 < n_d$ is the sample size at each survey time. We assume the loss function appropriate to the end-use allocation formula is a quadratic loss

$$L_{avg} = a_1 \text{ trace } [(1/T) \sum_{j=0}^{T-1} (\underset{\sim}{C}+j\underset{\sim}{Q})] .$$

Using (14) we obtain

$$L_{avg} = a_1 \text{ trace } [-\underset{\sim}{Q}/2 + (z\underset{\sim}{D} + (T^2/4)\underset{\sim}{I})^{1/2}\, \underset{\sim}{V}'\underset{\sim}{Q}\underset{\sim}{V}] ,$$

where we have used the property that $\text{trace}[\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}] = \text{trace}[\underset{\sim}{B}\underset{\sim}{C}\underset{\sim}{A}]$, that if $\underset{\sim}{U}$ is an orthogonal matrix that diagonalizes the matrix $\underset{\sim}{A}$, then $\underset{\sim}{U}$ also diagonalizes $\underset{\sim}{A}^{1/2}$, or if $\underset{\sim}{A} = \underset{\sim}{U}\underset{\sim}{B}\underset{\sim}{U}'$, then $\underset{\sim}{A}^{1/2} = \underset{\sim}{U}\underset{\sim}{B}^{1/2}\underset{\sim}{U}'$, and that $\text{trace}[\underset{\sim}{A}] = \text{trace}[\underset{\sim}{U}'\underset{\sim}{A}\underset{\sim}{U}] = \text{trace}[\underset{\sim}{B}]$; see, for example, Bellman (1970).

We may now write our cost function J for the two-item repeated survey problem with scalar sample size n_d and sampling interval T as

$$\begin{aligned} J &= (1/T)\,[c_0 + c_1 m n_d] + a_1 \text{ trace } [-\underset{\sim}{Q}/2 + ((T/n_d)\underset{\sim}{D} + (T^2/4)\underset{\sim}{I})^{1/2}\, \underset{\sim}{S}] \\ &= c_0/T + c_1\, m/z + a_1\, [s_{11}(z\, d_{11} + T^2/4)^{1/2} + s_{22}(z\, d_{22} + T^2/4)^{1/2} \\ &\qquad - (q_{11} + q_{22})/2] . \end{aligned}$$

It can be shown that J is convex in n_d given T and convex in T given n_d . In practice, (n_d^*, T^*) will be found in a region in which J is convex in (n_d, T) . Fixing T successively at $T=1,2,\ldots,10$

years, we solve $\partial J/\partial z = 0$ for z , and hence for $n_d = 1/z$, by a numerical search procedure. We then compute J for each T using the minimizing n_d and pick the T and n_d combination with lowest average annual cost J . For the given cost coefficients, numerical results are displayed in Figure 2. The T and n_d combination for which the average cost is minimized is $T=3$ years and n_d = 5,26[illegible] units, with J = \$875,473 per year. As Figure 2 makes clear, the average cost curve is rather flat in the neighborhood of the optimal T^* and typically the survey administrator will not incur major additional costs by choosing $T^* + 1$ or $T^* - 1$ instead of T^* . It may also be seen that sampling too frequently is relatively more costly than sampling too seldom, assuming the correctness of the underlying random walk process model for the socioeconomic variables. An administrator who is concerned that the underlying process parameters may take unexpected jumps or exhibit turning points, which are not modeled by the simple time-invariant random walk models, would presumably opt for sampling more frequently than the optimal interval found by this method.

4. Model B

In this section we present a steady-state analysis of a vector random yield model in which a fixed scalar designated sample size n_d is used for all survey times, and T remains constant. In the use of a fixed order size, the development resembles the treatment of random supply in the inventory literature, but in our present analysis we have level-dependent deterministic decay rather than the independent stochastic demands of Karlin (1958). The stochastic nature of $N(k)$ enters the steady-state analysis in the case of a fixed n_d only through the expected obtained sample size,

$$E[N(k)] = \theta\, n_d \, .$$

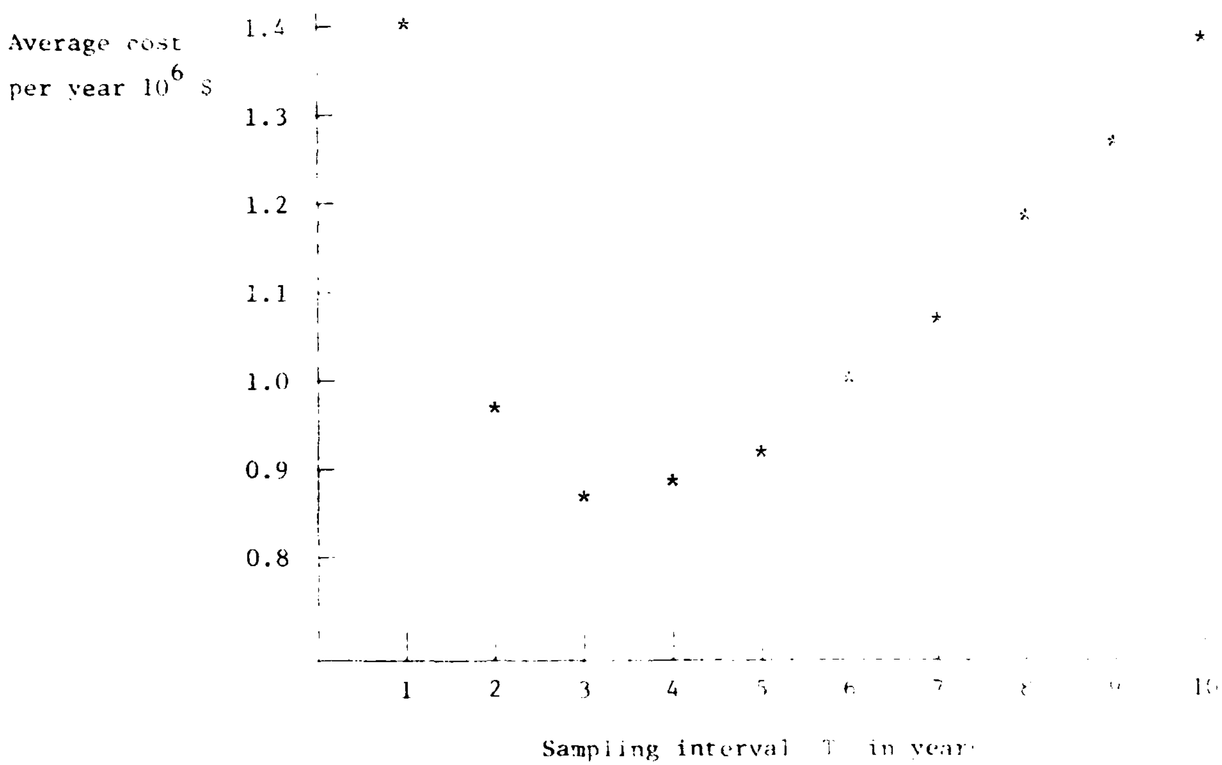


Figure 2.--Average cost versus sampling interval: Model A with m

In a random yield model, an ordering rule that requires us to order a fixed amount at every survey time is an "open loop" approach in the sense that not all of the information available or to become available from the evolving history of the survey system is used in choosing the sampling interval T and the designated sample size n_d. In the special case treated here, we do not accommodate the general case under random yield sampling where T may vary stochastically. Under the assumptions of Model B, the cost function [illegible] becomes, using (8),

$$\begin{aligned} L &= E[L'] = E[c_0 + c_1 n_d + c_2 N(k)] + a_1 \operatorname{trace} [C(k|k) + ((T-1)/2) Q] \\ &= c_0 + (c_1 + c_2) n_d + a_1 \operatorname{trace} [E[C(k|k)] + ((T-1)/2) Q] . \end{aligned} \qquad (15)$$

Note that $C(k|k)$ is now stochastically dependent on $N(k)$.

Since we assume that the stochastic system is in steady state, we define [illegible], $\hat{C} = C(k|k) = C(k|k-T)$, as the conditional expected error covariance

$$C = E[C(k|k) | [illegible], n_d, T, Q, R] ,$$

where [illegible] is a known constant and $n_d = n_d(k)$ for all k. Since Q and R are positive definite, so are C and $\hat{C}$. For scalar sample size n_d, the C we seek is found by methods similar to those employed in Section 3 for Model A, which is a fixed-yield vector model with scalar sample size. The Kalman relations must now be expressed in the form of relations among expected error covariances, expected Kalman gain, and expected obtained sample size. We now have

$$C = [I - K] (\hat{C} + TQ) ,$$

where

$$K = (\hat{C} + TQ) [\hat{C} + TQ + E[R/N(k)]]^{-1} .$$

This leads to the matrix equation

$$\hat{C} Q^{-1} \hat{C} + T C - T \underset{\sim}{R} E[1/N(k)] = 0 . \qquad (16)$$

This equation is of the form of Equation (13) except that it involves expectations rather than deterministic quantities.

Solving (16) for $\hat{C}$ we obtain, strictly paralleling (14),

$$\underset{\sim}{C} = -(T/2)\,\underset{\sim}{Q} + \underset{\sim}{Q}^{1/2}\,[T\,E[1/N(k)]\,\underset{\sim}{Q}^{-1/2}\,\underset{\sim}{R}\,\underset{\sim}{Q}^{-1/2} + (T^2/4)\underset{\sim}{I}]^{1/2}\,\underset{\sim}{Q}^{1/2} ,$$

so that, since $E[\underset{\sim}{C} + ((T-1)/2)\underset{\sim}{Q}] = \hat{\underset{\sim}{C}} + ((T-1)/2)\underset{\sim}{Q}$, we may substitute the value of $\hat{\underset{\sim}{C}}$ found above into the cost function J given by (15). The solution procedures for Model A may then be applied to find (n_d^*, T^*) .

5. Model C

Model C is an extension of random-yield Model B. We consider here a scalar process which evolves in continuous time t with survey measurements taken at time points t_k , $k=1,2,\ldots$. Here we allow the intersurvey periods T_k , $k=1,\ldots$ ($T_k = t_k - t_{k-1}$) , to vary as a function of the obtained sample size $N(t_{k-1})$. We assume that the scalar process is represented by the linear stochastic differential equation

$$dx(t) = x(t)dt + d\beta(t) , \quad t_0 < t , \tag{17}$$

where $x(t)$ is a unidimensional Brownian motion process with

$$E[(d\beta(t))^2] = Q\,dt .$$

The survey equation (2) now becomes

$$y(t_k) = x(t_k) + (1/N(t_k)) \sum_{i=1}^{N(t_k)} u_i(t_k) ,$$
$$k=1,2,\ldots \quad \text{and} \quad t_0 \leq t_k < t_{k+1} , \tag{18}$$

where $y(t_k)$ and $u_i(t_k)$ are defined as before. As set forth in Jazwinski (1970, Theorem 7.1), the optimal filter for the system (17) and (18) becomes the continuous-discrete filter:

between surveys,

$$d\hat{x}(t|t)/dt = \hat{x}(t|t) ,$$
$$dC(t|t)/dt = 2C(t|t) + Q , \quad t_k < t < t_{k+1} ; \tag{19}$$

at surveys,

$$\hat{x}(t_k|t_k) = \hat{x}(t_k|t_k^-) + K(t_k)\,[y(t_k) - \hat{x}(t_k|t_k^-)] \;,$$
$$C(t_k|t_k) = [I - K(t_k)]\,C(t_k|t_k^-) \;, \tag{20}$$

where the Kalman gain $K(t_k)$ is given by

$$K(t_k) = C(t_k|t_k^-)\,[C(t_k|t_k^-) + R/N(t_k)]^{-1} \;, \tag{21}$$

and t_k^- represents the time instant just before the survey conducted (and instantaneously processed) at t_k . As shown in Jazwinski (1970), if we integrate the process equation in (19) over intervals $[t_k, t_{k+1}]$, we may write

$$x(t_{k+1}) = x(t_k) + \int_{t_k}^{t_{k+1}} d\beta(\tau) \;,$$

$$= x(t_k) + w(k+1,k) \;,$$

where

$$w(k+1,k) = \int_{t_k}^{t_{k+1}} d\beta(\tau) \;.$$

By Jazwinski (1970, Theorem 4.1), $\{w(k+1,k)\}$ is a zero-mean, white Gaussian sequence with

$$E[(w(k+1,k))^2] = (t_{k+1} - t_k)Q \;,$$

and the continuous-discrete filter for Model C may be imbedded in a discrete filter paralleling Models A and B. The integrated form of (19) is

$$\hat{x}(t|t) = \hat{x}(t_k|t_k) \;,$$
$$C(t|t) = C(t_k|t_k) + (t-t_k)Q \;, \quad t_k < t < t_{k+1} \;. \tag{22}$$

We now analyze Model C with response rate $\phi(k)$. Suppose that the equivalent sample size on hand at time t_0 is $N_0(t_0)$. The equivalent sample size remaining at time t, $N_0(t|t_0)$, is

$$N_0(t|t_0) = N_0(t_0|t_0)\,[1 + (t-t_0)QR^{-1}N_0(t_0|t_0)]^{-1} . \tag{23}$$

When $N_0(t|t_0)$ just falls to a preselected value n_r (the reorder point), $1 < n_r$, a new survey of designated sample size n_d is ordered [illegible] resulting in [illegible] $N(t_j) \le n_d$. We assume that $N(t_j)$ is independent of $N(t_i)$, $i \neq j$, and hence $N_0(t_j|t_0) = n_r + N(t_j)$ is independent of $N_0(t_i|t_0) = n_r + N(t_i)$, $i \neq j$. In other words, [illegible] we have a renewal process with renewals occurring at the time points where $N_0(t|t_0)$ has just decayed to the preselected reorder point n_r, $1 < n_r$.

We now denote by $\phi(t_k)$ the response rate for the survey conducted at t_k. We assume that the $\phi(t_k)$ are i.i.d. random variables with a known distribution and that $0 < \phi(t_k) < 1$. We thus obtain

$$N_0(t_k|t_k) = n_r + \phi(t_k)n_d ,$$

and

$$E[N_0(t_k|t_k)] = n_r + \bar{\phi} n_d ,$$

where $\bar{\phi}$ is $E[\phi(t_k)]$.

In inventory terms, we thus have a problem of stochastic replenishment, with designated order size n_d, delivered order size $N(t_k)$, and updated equivalent sample size $N_0(t_k|t_k)$. Given the realization of $N_0(t_k|t_k)$, we then have level-dependent deterministic decay from $N_0(t_k|t_k)$ down to the reorder point n_r. The time required for this decay to take place, T_{k+1}, is found in terms of n_r, n_d, and $\phi(t_k)$ to be

$$T_{k+1} = (R/Q) \left[1/n_r - (n_r + \phi(t_k)n_d)^{-1}\right] . \tag{24}$$

We may also write

$$T_{k+1}^{-1} = QR^{-1} n_r \left[n_r + \phi(t_k)n_d\right] / \phi(t_k)n_d . \tag{25}$$

A typical pattern of level-dependent deterministic decay and stochastic replenishment is depicted in Figure 3.

The average ordering cost, $G_{avg}(\phi(t_k))$, over an interval of length $T_{k+1} = t_{k+1} - t_k$, is

$$\begin{aligned} G_{avg}(\phi(t_k)) &= T_{k+1}^{-1} \left[c_0 + (c_1 + c_2\phi(t_k))n_d\right] \\ &= QR^{-1} n_r (n_r + \phi(t_k)n_d) \left[c_0 + (c_1 + c_2\phi(t_k))n_d\right] / \phi(t_k)n_d . \end{aligned} \tag{26}$$

Since $C(t|t_k) = R / N_0(t|t_k)$, we may write the average quadratic loss $L_{avg}(\phi(t_k))$ over the interval T_{k+1} as

$$L_{avg}(\phi(t_k)) = (a_1/T_{k+1}) \int_{t_k}^{t_{k+1}} C(\tau|t_k)d\tau ,$$

which yields

$$L_{avg}(\phi(t_k)) = (a_1R/2) \left[1/n_r + 1/(n_r + \phi(t_k)n_d)\right] . \tag{27}$$

Using (8) and (15), we obtain the cost function J as

$$J = E_\phi \left[G_{avg}(\phi(t_k)) + L_{avg}(\phi(t_k))\right] . \tag{28}$$

It can be verified by taking second derivatives that J is convex in n_r , $1 < n_r$, for fixed n_d , $0 < n_d$, and in n_d , $0 < n_d$, for fixed n_r , $1 < n_r$. If we compute the determinant of the Hessian of $G_{avg} + L_{avg}$, we find that for some (but not all) combinations of values of n_r , n_d , and the cost coefficients the determinant is negative. Therefore, J is not convex in general. Nonetheless, in practice there

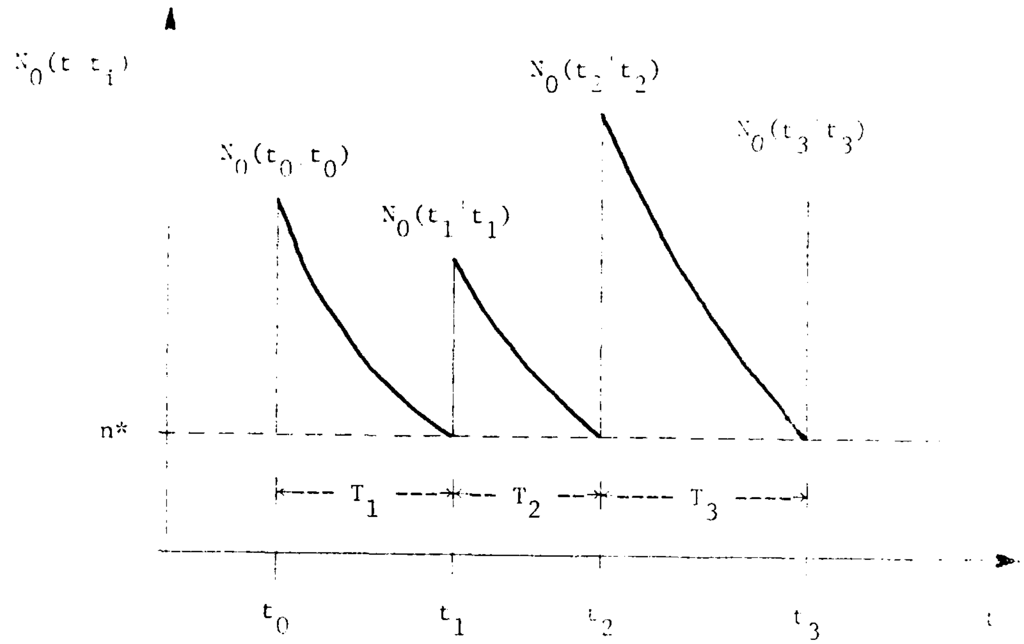


Figure 3.--Level-dependent decay and stochastic replenishment in continuous-time random yield Model C.

is an optimal (n_r^*, n_d^*) pair on the interior of a region in which J is convex and it may be found as an iterative solution to the pair of equations $\partial J/\partial n_r = 0$, $\partial J/\partial n_d = 0$.

Case of Uniform ϕ

In the case of a uniform ϕ on (a,b) ,

$$E[\phi] = \mu = (a+b)/2 \ ,$$

$$E[1/\phi] = \Lambda = (b-a)^{-1} \log(b/a) \ ,$$

and

$$E[(n_r + \phi n_d)^{-1}] = (1/n_d)(b-a)^{-1} \log[(n_r + b n_d)/(n_r + a n_d)] \ ,$$

so that

$$J = E[C_{avg}(\phi)] + E[L_{avg}(\phi)]$$

$$= QR^{-1}[c_0 n_r^2 \Lambda/n_d + c_1 n_r^2 \Lambda + c_2 n_r^2 + c_0 n_r + c_1 n_r \mu n_d + c_2 n_r \mu n_d] \qquad (29)$$

$$+ (a_1 R/2)\left[1/n_r + (1/n_d)(b-a)^{-1} \log[(n_r + b n_d)/(n_r + a n_d)]\right] .$$

Now J may be minimized over (n_r, n_d) by standard iterative methods closely paralleling those used for (s,q) inventory problems. A numerical example of Model C with a uniform ϕ is portrayed in Figure 4. As is shown in Smith (1980), a Markov chain set-up may also be used to find an optimal rule.

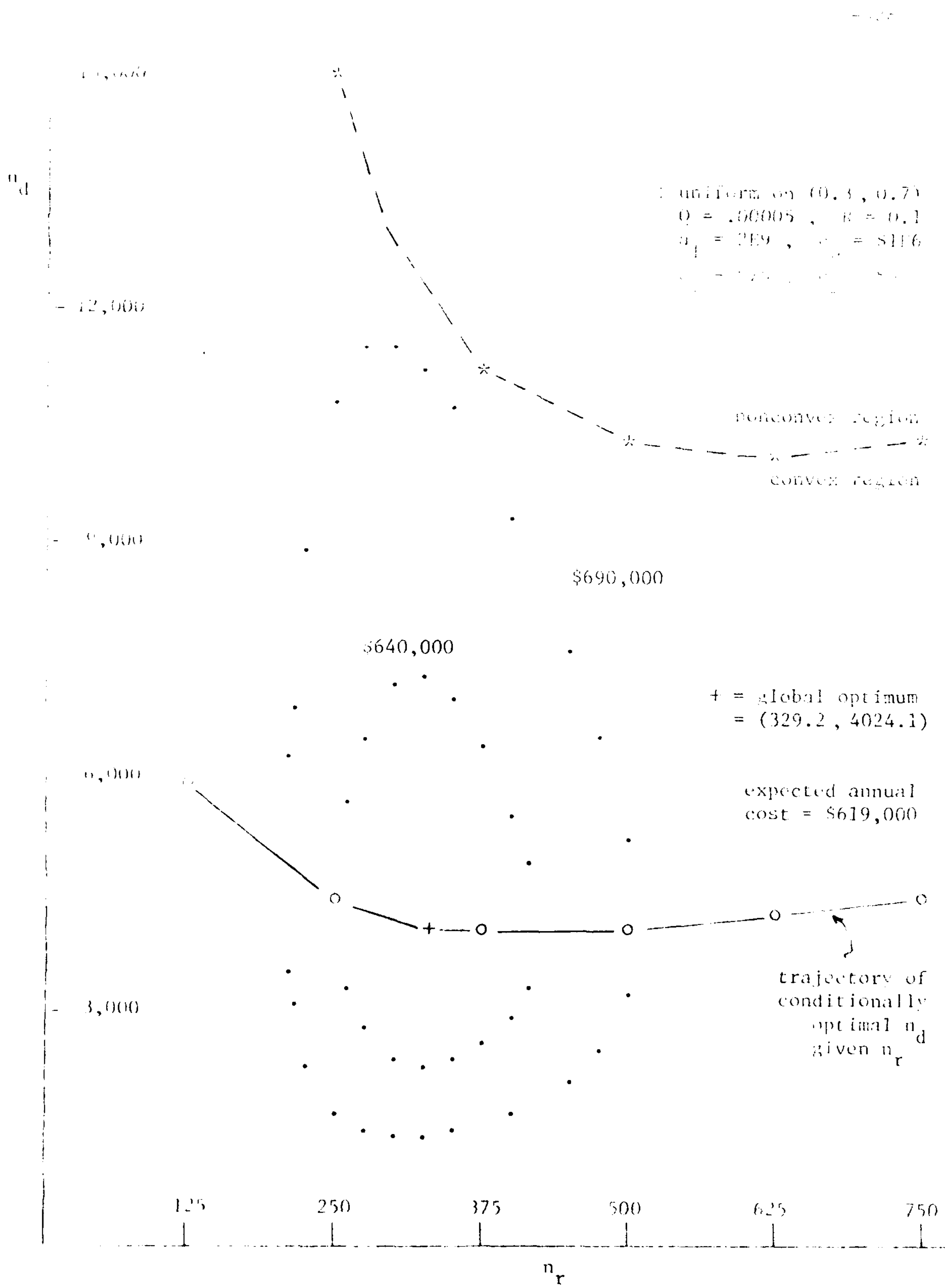


Figure 4.--Designated sample size n_d versus reorder point n_r: isocost contours for example of Model C.

AOKI, M. and M. T. LI (1968). Optimal discrete-time control with cost for observation. *IEEE Transactions on Automatic Control*, AC-14, 165-175.

[illegible] (1970). [illegible] New York.

[illegible] (1976). Effects of formulas and data used on the allocation of federal funds. In *Proceedings [illegible]*, American Statistical Association, Washington, D.C.

JAZWINSKI, A. (1970). *[illegible]*. Academic Press, New York.

KARLIN, S. (1958). Steady state solutions. In *[illegible]* (K. J. Arrow, S. Karlin, and H. Scarf, eds.), Chapter 14. Stanford University Press, Stanford, California.

MEIER, L., J. PESCHON, and R. M. DRESSLER (1967). Optimal control of measurement subsystems. *IEEE Transactions on Automatic Control*, AC-12, 528-536.

MELSA, J. L. and D. L. COHN (1978). *[illegible]*. McGraw-Hill, New York.

SAGE, A. P. and C. C. WHITE, III (1977). *Optimum Systems Control*, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

SCOTT, A. J. and T. M. F. SMITH (1974). Analysis of repeated surveys using time series methods. *Journal of the American Statistical Association*, 69, 674-678.

REFERENCES

[illegible], M. and M. [illegible] (1968). Optimal discrete-time control with cost for observation. *IEEE Transactions on [illegible]*, AC-1[illegible], 165-175.

[illegible], [illegible] (19[illegible]). [illegible] [illegible], New York.

[illegible] (1978). Effects of formulas and data used in the [illegible] of federal funds. In *Proceedings [illegible] 197[illegible]*, American Statistical Association, Washington, D.C.

[illegible], A. (1970). *[illegible]*. Academic Press, New York.

KARLIN, S. (1958). Steady state solutions. In *[illegible] of Inventory and Production* (K. J. Arrow, S. Karlin, and H. Scarf, eds.), Chapter 14. Stanford University Press, Stanford, California.

MEIER, L., J. PESCHON, and R. M. DRESSLER (1967). Optimal control of measurement subsystems. *IEEE Transactions on Automatic [illegible]*, AC-12, 528-536.

[illegible], J. L. and D. L. COHN (1978). *Decision and [illegible]*. McGraw-Hill, New York.

[illegible], A. P. and C. C. WHITE, III (1977). *Optimum Systems [illegible]*, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

SCOTT, A. J. and T. M. F. SMITH (1974). Analysis of repeated surveys using time series methods. *Journal of the American Statistical Association*, 69, 674-678.

SMITH, W. (1979). An equivalent sample size form of the [illegible]
Unpublished technical note.

SMITH, W. (1980). Sample size and timing decisions for repeated socio-economic surveys. D.Sc. dissertation, The George Washington [illegible].

SMITH, W. and D. [illegible] (197[illegible]). Statistical [illegible] formulas for allocation of funds. [illegible], American Statistical [illegible], Washington, D.C.

SMITH, W. and D. ZALKIND (1980). Inventory analysis of repeated surveys under random-yield sampling. Paper presented at Annual Meeting of the American Statistical Association, Section on Survey Research Methods, Houston, August 11-14, 1980.

SPENCER, B. (1979). Benefit-cost analysis of data used to allocate funds: general revenue sharing. Ph.D. dissertation, Yale University.

THE GEORGE WASHINGTON UNIVERSITY
Program in Logistics
Distribution List for Technical Papers

The George Washington University
   Office of Sponsored Research
   Library
   Vice President H. F. Bright
   Dean Harold Liebowitz
   Dean Henry Solomon

ONR
   Chief of Naval Research
     (Codes 200, 434)
   Resident Representative

OPNAV
   OP-40
   DCNO, Logistics
   Navy Dept Library
   NAVDATA Automation Cmd
   OP-964

Naval Aviation Integrated Log Support

NARDAC Tech Library

Naval Electronics Lab Library

Naval Facilities Eng Cmd Tech Library

Naval Ordnance Station
   Louisville, Ky.
   Indian Head, Md.

Naval Ordnance Sys Cmd Library

Naval Research Branch Office
   Boston
   Chicago
   New York
   Pasadena
   San Francisco

Naval Ship Eng Center
   Philadelphia, Pa.
   Washington, DC

Naval Ship Res & Dev Center

Naval Sea Systems Command
   PMS 30611
   Tech Library
   Code 073

Naval Supply Systems Command
   Library
   Operations and Inventory Analysis

Naval War College Library
   Newport

BUPERS Tech Library

FMSO

Integrated Sea Lift Study

USN Ammo Depot Earle

USN Postgrad School Monterey
   Library
   Dr Jack R. Borsting
   Prof C. R. Jones

US Marine Corps
   Commandant
   Deputy Chief of Staff, R&D

Marine Corps School Quantico
   Landing Force Dev Ctr
   Logistics Officer
   Commanding Officer
     USS Francis Marion (LPA-249)

Armed Forces Industrial College

Armed Forces Staff College

Army War College Library
   Carlisle Barracks

Army Cmd & Gen Staff College

Army Logistics Mgt Center
   Fort Lee

Commanding Officer, USALDSRA
   New Cumberland Army Depot

Army Inventory Res Ofc
   Philadelphia

Air Force Headquarters
   AFADS-3
   LEXY
   SAF/ALG

Griffiss Air Force Base
   Reliability Analysis Center

Gunter Air Force Base
   AFLMC/XR

Maxwell Air Force Base Library

Wright-Patterson Air Force Base
   Log Command
   Research Sch Log
   AFALD/XR

Defense Documentation Center

National Academy of Sciences
   Maritime Transportation Res Board Library

National Bureau of Standards
   Dr B. H. Colvin
   Dr Joan Rosenblatt

National Science Foundation

National Security Agency

Weapon Systems Evaluation Group

British Navy Staff

National Defense Hdqtrs, Ottawa
   Logistics, OR Analysis Establishment

American Power Jet Co
   George Chernowitz

General Dynamics, Pomona

General Research Corp
   Dr Hugh Cole
   Library

Logistics Management Institute
   Dr Murray A. Geisler

MATHTEC
   Dr Eliot Feldman

Rand Corporation
   Library

Carnegie-Mellon University
   Dean H. A. Simon
   Prof G. Thompson

Case Western Reserve University
   Prof B. V. Dean
   Prof M. Mesarovic
   Prof S. Zacks

Cornell University
   Prof R. E. Bechhofer
   Prof R. W. Conway
   Prof Andrew Schultz, Jr.

Cowles Foundation for Research in Economics
   Prof Herbert Scarf
   Prof Martin Shubik

Florida State University
   Prof R. A. Bradley

Harvard University
   Prof K. J. Arrow
   Prof W. G. Cochran
   Prof Arthur Schleifer, Jr.

Princeton University
   Prof A. W. Tucker
   Prof J. W. Tukey
   Prof Geoffrey S. Watson

Purdue University
Prof S. S. Gupta
Prof H. Rubin
Prof Andrew Whinston

Stanford University
Prof T. W. Anderson
Prof G. B. Dantzig
Prof F. S. Hillier
Prof D. L. Iglehart
Prof Samuel Karlin
Prof G. J. Lieberman
Prof Herbert Solomon
Prof A. F. Veinott, Jr.

University of California, Berkeley
Prof R. E. Barlow
Prof D. Gale
Prof Jack Kiefer
Prof Rosedith Sitgreaves

University of California, Los Angeles
Prof J. R. Jackson
Prof R. R. O'Neill

University of North Carolina
Prof W. L. Smith
Prof M. R. Leadbetter

University of Pennsylvania
Prof Russell Ackoff
Prof Thomas L. Saaty

University of Texas
Prof A. Charnes

Yale University
Prof F. J. Anscombe
Prof I. R. Savage

Prof Z. W. Birnbaum
University of Washington

Prof B. H. Bissinger
The Pennsylvania State University

Prof Seth Bonder
University of Michigan

Prof G. E. P. Box
University of Wisconsin

Dr Jerome Bracken
Institute for Defense Analyses

Prof H. Chernoff
Mass. Institute of Technology

Prof Arthur Cohen
Rutgers - The State University

Mr Wallace M. Cohen
US General Accounting Office

Prof C. Derman
Columbia University

Prof Masao Fukushima
Kyoto University

Prof Saul I. Gass
University of Maryland

Dr Donald P. Gaver
Carmel, California

Prof Amrit L. Goel
Syracuse University

Prof J. F. Hannan
Michigan State University

Prof H. O. Hartley
Texas A & M Foundation

Mr Gerald F. Hein
NASA, Lewis Research Center

Prof W. M. Hirsch
Courant Institute

Dr Alan J. Hoffman
IBM, Yorktown Heights

Prof John R. Isbell
State University of New York, Amherst

Dr J. L. Jain
University of Delhi

Prof J. H. K. Kao
Polytech Institute of New York

Prof W. Kruskal
University of Chicago

Mr S. Kumar
University of Madras

Prof C. E. Lemke
Rensselaer Polytech Institute

Prof Loynes
University of Sheffield, England

Prof Steven Nahmias
University of Pittsburgh

Prof D. B. Owen
Southern Methodist University

Prof E. Parzen
Texas A & M University

Prof H. O. Posten
University of Connecticut

Prof R. Remage, Jr.
University of Delaware

Prof Hans Riedwyl
University of Bern

Dr Fred Rigby
Texas Tech College

Mr David Rosenblatt
Washington, D. C.

Prof M. Rosenblatt
University of California, San Diego

Prof Alan J. Rowe
University of Southern California

Prof A. H. Rubenstein
Northwestern University

Dr M. E. Salveson
West Los Angeles

Prof Edward A. Silver
University of Waterloo, Canada

Prof M. J. Sobel
Georgia Inst of Technology

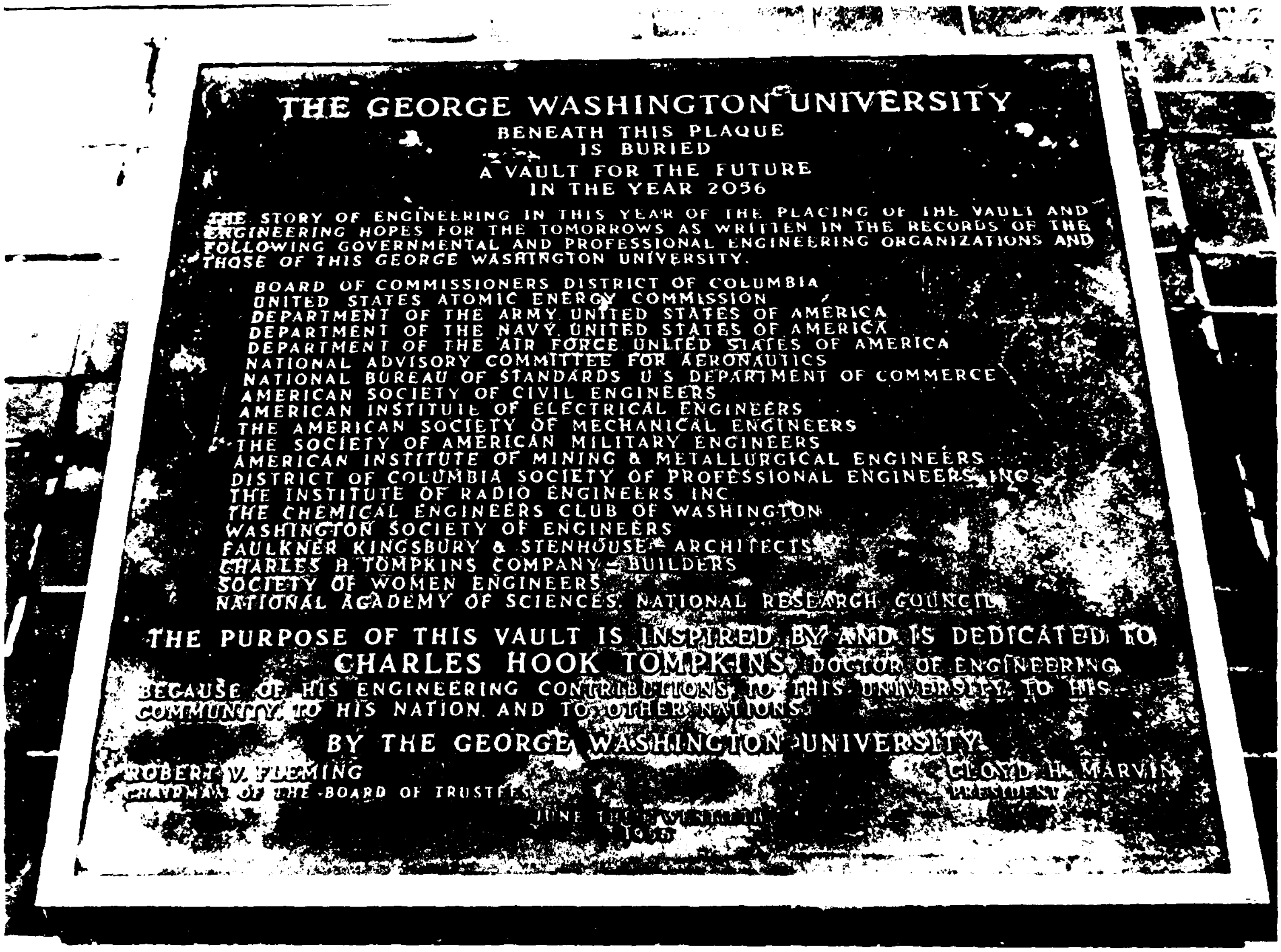
Prof R. M. Thrall
Rice University

Dr S. Vajda
University of Sussex, England

Prof T. M. Whitin
Wesleyan University

Prof Jacob Wolfowitz
University of South Florida

Prof Max A. Woodbury
Duke University



To cope with the expanding technology, our society must be assured of a continuing supply of rigorously trained and educated engineers. The School of Engineering and Applied Science is completely committed to this objective.