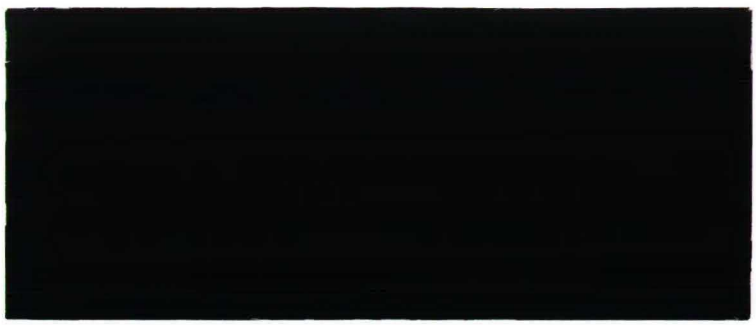
[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM



SEQUENTIAL BIFURCATION FOR OBSERVATIONS WITH RANDOM ERRORS

Bert Bettonvil

FEW 382

march 1989

SEQUENTIAL BIFURCATION FOR OBSERVATIONS WITH RANDOM ERRORS

Bert Bettonvil

Tilburg University/Eindhoven University of Technology

P O Box 90153

5000 LE Tilburg

Netherlands

Sequential Bifurcation is a method for factor screening, which is proven to be very efficient in case of observations without random errors. In this paper the method is extended to observations with random errors. The signal-to-noise ratio is taken as a measure of "importance". By means of Monte Carlo experiments the power, significance and number of observations are investigated.

Keywords: Experimental Design, Screening, Aggregated Variables, Binary Search, Simulation.

1. INTRODUCTION

We focus on problems with a great many (100,1000,10000?) input variables, out of which only a few are really important. A straightforward screening method would use at least as many observations as there are variables to be inspected. But an observation can be so time-consuming, that collecting so many data is prohibitive. Also see Kleijnen (1987).

We represent the model as a linear regression (meta) model. By imposing one special restriction we can reduce the number of runs. We consider the first-order (main effects) model

$$y(x_1,x_2,\ldots,x_N,e) = \beta_0+\beta_1x_1+\beta_2x_2+\ldots+\beta_Nx_N+e, \qquad (1.1)$$

in which we assume non-negative effects only: $\beta_l \geq 0$ ($l=1,\ldots,N$), and in which we assume for convenience that N is a power of two ($N=2^m$ for some $m \in \mathbb{N}$). To estimate a first-order model, a two-level experiment suffices; so we may take $x_l \in \{0,1\}$ for $l=1,\ldots,N$. In previous papers (Bettonvil 1988a,b,c) we developed a method, called Sequential Bifurcation or "SB", to handle the case of no errors: $e=0$. At present we shall treat the case $e \sim NID(0,\sigma^2)$, where $NID(\mu,\sigma^2)$ stands for normally independently distributed with mean μ and variance σ^2. We assume σ^2 to be known (in a future paper we shall drop this assumption). We want to find all "important" factors, calling a factor important, iff its regression parameter is large. In the error-free case, we called a regression parameter "large", iff it is greater than some given number $\delta \geq 0$; in case of errors ($\sigma^2>0$) a regression parameter is called large, iff it is greater than $\delta\sigma$, with δ some given positive number.

SB for observations without random error is described in a condensed way in section 2. Section 3 is the main part of this paper; here we introduce random error, propose an adaption of SB, and present results from Monte Carlo experiments. Our findings are summarized in section 4.

2. SEQUENTIAL BIFURCATION FOR OBSERVATIONS WITHOUT ERRORS

In this section, we give a brief description of Sequential Bifurcation in case the observations have no errors: e=0 in (1.1). For an extensive description, we refer to Bettonvil(1988c).

Sequential Bifurcation is a group screening method. The groups it considers are: first (in "stage 0") one group of size $N=2^m$, next (in stage 1) two groups of size $N/2=2^{m-1}$ each, then (in stage 2) four groups of size $N/4=2^{m-2}$ each, and so on. In general, in stage j (j=0,1,...,m) SB considers 2^j groups of size $N/2^j=2^{m-j}$ each. In stage m, we have $2^m=N$ groups of size $2^{m-m}=1$: we have reached the individual factors. The sum of the parameters in the k^{th} group at stage j (k=1,2,...,2^j; j=0,1,...,m) is called the aggregated effect of this group, or "the k^{th} parameter at stage j". It is denoted by $\beta_{k|j}$ and defined as

$$\beta_{k|j} := \sum_{l=(k-1)*2^{m-j}+1}^{k*2^{m-j}} \beta_l \qquad (j=0,1,\ldots,m;\ k=1,2,..,2^j), \qquad (2.1)$$

where ":=" means "is defined as". A direct consequence of (2.1) is

$$\beta_{k|j} = \beta_{2k-1|j+1} + \beta_{2k|j+1} \qquad (j=0,1,\ldots,m-1;\ k=1,2,\ldots,2^j), \qquad (2.2)$$

which can be proven by substitution.

Apart from β_0, all regression parameters are assumed to be non-negative. This means that all aggregated parameters are also non-negative, and that for each j (j=0,1,...,m) the sequence β_0, $\beta_0+\beta_{1|j}$, $\beta_0+\beta_{1|j}+\beta_{2|j}$, ... , $\beta_0+\beta_{1|j}+\beta_{2|j}+\ldots+\beta_{2^j|j}$ is non-decreasing. So, if we define

$$\beta^+_{i|j} := \beta_0 + \Sigma^i_{k=1}\,\beta_{k|j} \qquad (i=0,1,\ldots,2^j;\ j=0,1,\ldots,m) \tag{2.3}$$

(where $\Sigma^0_{k=1}$ is always 0), then the sequence $\beta^+_{i|j}$ is non-decreasing in i for each j. We need one more definition for the parameters, namely

$$\beta^+_i := \beta^+_{i|m} = \Sigma^i_{k=0}\,\beta_k \qquad (i=0,\ldots,2^m). \tag{2.4}$$

The SB design is such that all observations have the first i input variables at their high levels, and the remaining N-i input variables at their low levels; they are denoted as y_i ($i\in\{0,1,\ldots,N\}$). Because of (2.4), (1.1) and e=0 we have

$$y_i = \beta^+_i \quad \text{for all observations } y_i. \tag{2.5}$$

We distinguish between y_i and β^+_i for two reasons. First, β^+_i is defined for all i=0,1,...,N: there are N+1 β^+_i's, whereas we will not need N+1 observations y_i. Second, when random errors are introduced, β^+_i will be the expectation of y_i.

Alternatively, we write the observations as $y_{i|j}$, in which the input variables in groups 1|j, ... , i|j are at their high levels; the remaining input variables are at their low levels ($i\in\{0,1,\ldots,2^j\}$, $j\in\{0,1,\ldots,m\}$). But

the input variables in groups $1|j, \ldots, i|j$ are exactly the input variables $1,2,\ldots,i*2^{m-j}$, so

$$y_{i|j} = y_{i*2^{m-j}} \qquad \text{for all observations.} \tag{2.6}$$

Note that $y_{i|j}=\beta^{+}_{i|j}=\beta^{+}_{i*2^{m-j}}=y_{i*2^{m-j}}$. Why do we need so many notations for the same quantity? First of all, the equality between observations and sums of parameters will end as soon as we switch to observations with random errors. The other differences will be clarified presently.

Now suppose we know $\beta^{+}_{i-1|j}$ and $\beta^{+}_{i|j}$. Then $\beta_{i|j}=\beta^{+}_{i|j}-\beta^{+}_{i-1|j}$. If $\beta_{i|j}\leq\delta$, then none of the parameters in group i at level j can exceed δ: $\beta_{(i-1)*2^{m-j}+1}\leq\delta$, $\beta_{(i-1)*2^{m-j}+2}\leq\delta$, ... , $\beta_{i*2^{m-j}}\leq\delta$. We do not investigate this group any further. On the other hand, if $\beta_{i|j}>\delta$, we distinguish between $j=m$ and $j\neq m$. If $j=m$, we have found that $\beta_i>\delta$; so factor number i is important. If $j<m$ and $\beta_{i|j}>\delta$, then we proceed to the next stage, using (2.2): we investigate whether $\beta_{2i-1|j+1}$ and/or $\beta_{2i|j+1}$ exceed δ. For $\beta_{2i-1|j+1}$ we need $\beta^{+}_{2i-1|j+1}$ and $\beta^{+}_{2i-2|j+1}$; for $\beta_{2i|j+1}$ we need $\beta^{+}_{2i|j+1}$ and $\beta^{+}_{2i-1|j+1}$.

We already have $\beta^{+}_{2i-2|j+1}$ and $\beta^{+}_{2i|j+1}$, as $\beta^{+}_{i|j} = \beta^{+}_{i*2^{m-j}} = \beta^{+}_{2i*2^{m-j-1}} = \beta^{+}_{2i|j+1}$, and, in the same way, $\beta^{+}_{i-1|j} = \beta^{+}_{2i-2|j}$. So all we need is $\beta^{+}_{2i-1|j+1}$: we observe $y_{2i-1|j+1}=\beta^{+}_{2i-1|j+1}$, and now we can compute $\beta_{2i-1|j+1}$ and $\beta_{2i|j+1}$.

SB always starts with the observation of $y_{0|0}=\beta^{+}_{0|0}$ and $y_{1|0}=\beta^{+}_{1|0}$, and continues as described above, starting with j=0, i=1. For the number of observations SB uses, we refer to Bettonvil(1988a,b,c).

Example 2.1. Suppose we have 8 factors (m=3), with only $\beta_2>0$. *This implies that* $\beta_0^+=\beta_1^+<\beta_2^+=\beta_3^+=\ldots=\beta_8^+$. *Take* $\delta=0$. *SB starts by observing* $y_{0|0}=\beta_{0|0}^+=\beta_0^+$ *and* $y_{1|0}=\beta_{1|0}^+=\beta_8^+$, *and computing* $\beta_{1|0}=\beta_{1|0}^+-\beta_{0|0}^+$.

As $\beta_{1|0}>0$, *and* $\beta_{1|0}=\beta_{1|1}+\beta_{2|1}$, *we observe* $y_{1|1}=\beta_{1|1}^+=\beta_4^+$ *and note that* $\beta_{0|0}^+=\beta_{0|1}^+=\beta_0^+$, $\beta_{1|0}^+=\beta_{2|1}^+=\beta_8^+$. *We find that* $\beta_{1|1}=\beta_{1|1}^+-\beta_{0|1}^+>0$, *and that* $\beta_{2|1}=\beta_{2|1}^+-\beta_{1|1}^+=0$. *We need not bother about* $\beta_{2|1}$: *all its components are zero.*

As $\beta_{1|1}>0$, *and* $\beta_{1|1}=\beta_{1|2}+\beta_{2|2}$, *we observe* $y_{1|2}=\beta_{1|2}^+=\beta_2^+$ *and note that* $\beta_{0|1}^+=\beta_{0|2}^+=\beta_0^+$, $\beta_{1|1}^+=\beta_{2|2}^+=\beta_4^+$. *We find that* $\beta_{1|2}=\beta_{1|2}^+-\beta_{0|2}^+>0$, *and that* $\beta_{2|2}=\beta_{2|2}^+-\beta_{1|2}^+=0$. *We need not bother about* $\beta_{2|2}$: *both its components are zero.*

As $\beta_{1|2}>0$, *and* $\beta_{1|2}=\beta_{1|3}+\beta_{2|3}$, *we observe* $y_{1|3}=\beta_{1|3}^+=\beta_3^+$ *and note that* $\beta_{0|2}^+=\beta_{0|3}^+=\beta_0^+$, $\beta_{1|2}^+=\beta_{2|3}^+=\beta_2^+$. *We find that* $\beta_1=\beta_{1|3}=\beta_{1|3}^+-\beta_{0|3}^+=0$, *and that* $\beta_2=\beta_{2|3}=\beta_{2|3}^+-\beta_{1|3}^+>0$, *which completes the screening.*

3. OBSERVATIONS WITH RANDOM ERRORS

From now on, we assume that

$$y_{i|j} = \beta_{i|j}^+ + e_{i|j} \tag{3.1}$$

with $e_{i|j}\sim NID(0,\sigma^2)$, σ^2 known; where $NID(\mu,\sigma^2)$ means Normally Independently Distributed with mean μ and variance σ^2. An alternative formulation of (3.1) is

$$y_k = \beta_k^+ + e_k \tag{3.2}$$

where $k=i*2^{m-j}$. We want to find the important factors by means of SB.

How do we define "important" in case of random noise? We might concentrate on the probability that a small parameter is found to be significant. Instead we focus on the probability that a large parameter is indeed declared large: power (the complement of the β-error). E.g., consider two confidence intervals, the first running from 1.0 to 2.0, the second from -1.0 to 11.0, and (say) a parameter with magnitude 10.0 is considered to be large; a parameter with confidence interval [1.0,2.0] is then significant but unimportant, while a parameter with confidence interval [-1.0,11.0] is not significant, but may be important. We want our procedure to work such that, if a parameter is at least equal to $\delta\sigma$ (with given $\delta>0$), then the probability of this parameter being declared unimportant, is at most equal to some given constant $\varepsilon>0$: β error. Furthermore, we want to minimize the number of observations, as well as the number of parameters that is incorrectly declared important.

In the non-error case we needed m+2 observations to find out that a single factor is important (see appendix 1). However, to find out that a particular input variable is unimportant, we needed at most m+2 (and at least 2) observations. As soon as we obtained a small $\beta_{k|j}$, we concluded that all original parameters composing $\beta_{k|j}$ are small, and we stopped investigating these parameters. If we are dealing with observations with random errors, we proceed analogously. Our decision to declare a factor important, must be based on m+2 observations (and for each single factor we know beforehand which observations we need). The decision to declare a factor unimportant, however, should be based on as few observations as possible.

Consider the m+2 random variables $y_{i_0}, y_{i_1}, \ldots, y_{i_{m+1}}$ with $0=i_0<i_1<\ldots<i_{m+1}=2^m$, on which our decision whether or not to declare β_ℓ to

be large (that is, larger than $\delta\sigma$) should be based. When do we accept H_0^{ℓ}: $\beta_{\ell}>\delta\sigma$? (Note that this null-hypothesis concerns factor ℓ only.) Of course the observations $y_{i_0},y_{i_1},\ldots,y_{i_{m+1}}$ include y_{ℓ} and $y_{\ell-1}$ (see 3.1 and 3.3 with j=m). So we might consider $\hat{\beta}_{\ell}= y_{\ell}-y_{\ell-1} = \beta_{\ell}+e_{\ell}-e_{\ell-1}$. Unfortunately, either y_{ℓ} or $y_{\ell-1}$ is the last of the m+2 observations that becomes available, that is, y_{ℓ} or $y_{\ell-1}$ is an observation at stage m (for otherwise, we would not need m+2 observations for β_{ℓ}). This means that, to investigate all factors, we must have N+1 observations (2 at level 0, 2^{j-1} at level j, j=1,...,m; together $2+\Sigma_{j=1}^{m}2^{j-1} = 2+2^m-1 = 2^m+1 = N+1$). The number of observations would be of the same order as the number of factors, which we wanted to avoid. So, we cannot use $\hat{\beta}_{\ell}$.

We take another view of $y_{i_0},y_{i_1},\ldots,y_{i_{m+1}}$. In observations $y_0,\ldots,y_{\ell-1}$ factor number ℓ is "off" ($x_{\ell}=0$), in observations $y_{\ell},\ldots,y_N$ factor number ℓ is "on" ($x_{\ell}=1$). Let L be such that $i_L=\ell-1$. We may consider $y_{i_0},\ldots,y_{i_L}$ as (under)estimators of $\beta_0+\beta_1+\ldots+\beta_{\ell-1}$; $y_{i_L+1},\ldots,y_{i_{m+1}}$ as (over)estimators of $\beta_0+\beta_1+\ldots+\beta_{\ell-1}+\beta_{\ell}$. Hence, the (L+1)*(m-L+1) differences $y_{i_+}-y_{i_-}$ with $i_+\in\{i_L+1,\ldots,i_{m+1}\}$ and $i_-\in\{i_0,\ldots,i_L\}$ are all (over)estimators of β_{ℓ}.

Example 3.1. Suppose we are dealing with $2^3=8$ factors. To arrive at β_2 we observe $y_{0|0}=y_0$, $y_{1|0}=y_8$, $y_{1|1}=y_4$, $y_{1|2}=y_2$ and $y_{1|3}=y_1$ and consider successively

$$\hat{\beta}_{1|0} = y_{1|0}-y_{0|0} = \beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6+\beta_7+\beta_8+e_{1|0}-e_{0|0};$$
$$\hat{\beta}_{1|1} = y_{1|1}-y_{0|0} = \beta_1+\beta_2+\beta_3+\beta_4 \qquad +e_{1|1}-e_{0|0};$$
$$\hat{\beta}_{1|2} = y_{1|2}-y_{0|0} = \beta_1+\beta_2 \qquad +e_{1|2}-e_{0|0};$$
$$\hat{\beta}_{2|3} = y_{1|2}-y_{1|3} = \beta_2 \qquad +e_{1|2}-e_{1|3}.$$

Now we consider the minimum of these four (over)estimators of β_2*:*

$$\min\{y_{1|0}-y_{0|0}, y_{1|1}-y_{0|0}, y_{1|2}-y_{0|0}, y_{1|2}-y_{1|3}\}=$$

$$\min\{\beta_2+e_{1|0}-e_{0|0}+\beta_1+\beta_3+\beta_4+\beta_5+\beta_6+\beta_7+\beta_8,$$

$$\beta_2+e_{1|1}-e_{0|0}+\beta_1+\beta_3+\beta_4,$$

$$\beta_2+e_{1|2}-e_{0|0}+\beta_1,$$

$$\beta_2+e_{1|2}-e_{1|3}\}\geq$$

$$\beta_2+\min\{e_{1|0}-e_{0|0}, e_{1|1}-e_{0|0}, e_{1|2}-e_{0|0}, e_{1|2}-e_{1|3}\}\geq$$

$$\beta_2+\min\{e_{1|0}-e_{0|0}, e_{1|0}-e_{1|3}, e_{1|1}-e_{0|0}, e_{1|1}-e_{1|3}, e_{1|2}-e_{0|0}, e_{1|2}-e_{1|3}\}=$$

$$\beta_2+\min\{e_{1|0}, e_{1|1}, e_{1|2}\}-\max\{e_{0|0}, e_{1|3}\}.$$

If we know the distribution of $\min\{e_{1|0}, e_{1|1}, e_{1|2}\}-\max\{e_{0|0}, e_{1|3}\}$*, then we have a rule to decide whether or not* β_2 *is large. Note that* $e_{1|0}, e_{1|1}, e_{1|2}, e_{0|0}, e_{1|3}$ *are independent. We shall return to this distribution.*

In general we have m+2 observations available for each β_ℓ. Of these observations, L+1 do not contain β_ℓ, and m-L+1 do contain β_ℓ. We consider the m+1 differences $\hat{\beta}_{1|0}:=y_{i_{m+1}}-y_{i_0}$, ..., $\hat{\beta}_{\ell|m}:=y_\ell-y_{\ell-1}$, corresponding with the m+1 aggregated parameters, that β_ℓ belongs to: $\beta_{1|0}=\beta^+_{1|0}-\beta^+_{0|0}$, ..., $\beta_{\ell|m}=\beta^+_{\ell|m}-\beta^+_{\ell-1|m}$. The minimum of these differences is greater than

$$\beta_\ell+\min\{e_{i_{m+1}}, e_{i_m}, \ldots, e_{i_{L+1}}\}-\max\{e_{i_L}, \ldots e_{i_0}\}, \tag{3.3}$$

and, as is shown in appendix 2, this is greater than

$$\beta_\ell-\sigma\wedge_0 \qquad \text{with probability } p_0 \tag{3.4}$$

where the relation between p_0 and $\wedge_0$ is given by table I in Bechhofer (1954). We reproduce a small part of Bechhofer's table I in our table 3.1. The table entries record the values $\wedge_0$ for which

$$P(\wedge_0+\min\{e_{i_{m+1}},e_{i_m},\ldots,e_{i_{L+1}}\}-\max\{e_{i_L},\ldots e_{i_0}\}<0)=p_0,$$

for m=3, L=0,1,2,3; and for m=8, L=0,...,9, and p_0=.9995, .995, and .95.

table 3.1 about here

As soon as a $\hat{\beta}_{i|j}$ is smaller than $\sigma(\delta-\wedge_0)$, all parameters composing $\beta_{i|j}$ are smaller than $\delta\sigma$ with probability $1-p_0$.

Example 3.3. In example 3.2 we had 8 factors, and we were interested in β_2. *We saw that*

$$\beta_2+\min\{e_{1|0},e_{1|1},e_{3|3}\}-\max\{e_{0|0},e_{1|2}\} \leq$$
$$\leq \min\{y_{1|0}-y_{0|0},y_{1|1}-y_{0|0},y_{1|1}-y_{1|2},y_{3|3}-y_{1|2}\}.$$

Suppose we want to find all factors that are greater than 10σ with probability of at least .95 (that is, δ=10, ε=0.05). We have a set of 5 random variables, subdivided into sets of 2 and 3 variables. From Bechhofer's (1954) table I it follows that

$$P(3.2805+\min\{e_{1|0}/\sigma, e_{1|1}/\sigma, e_{3|3}/\sigma\}-\max\{e_{0|0}/\sigma, e_{1|2}/\sigma\} > 0) = .95.$$

So we stop as soon as any of $y_{1|0}-y_{0|0}$, $y_{1|1}-y_{0|0}$, $y_{1|1}-y_{1|2}$, $y_{3|3}-y_{1|2}$ *is smaller than* $(10-3.2805)\sigma=6.7195\sigma$.

We used a number of Monte-Carlo experiments to investigate the behaviour of the above procedure. We took N=256 variables (m=8), $\sigma^2=1$, and investigated the following cases:

(a) $\beta_1=\delta$, all other parameters zero;

(b) $\beta_{86}=\delta$, all other parameters zero;

(c) $\beta_{241}=\delta$. all other parameters zero;

(d) all parameters zero.

(Note: for β_1 we always take the "left branch in the bifurcation tree"; for β_{86} we go left, right, left, right, etc.; for β_{241} we go four times right, then four times left. By considering these parameters, we may get an impression of the influence of the "path" we follow).

We took ε=.05, .005, .0005; δ=10, 8, 6, 4, and repeated all 4*3*4=48 experiments 1,000 times. We recorded the number of times the large parameter was (correctly) found, the number of times any other parameter was (incorrectly) found, and the total number of observations. To facilitate the

comparability of the experiments, we made them all use the same random number stream (i.e. each experiment used the same seed for the random number generator). The results are shown in tables 3.2.1 through 3.2.4.

tables 3.2.1 through 3.2.4 about here

From tables 3.2.1 through 3.2.4 we learn the following.

(1) The prescribed power is reached is all cases: where we expect .95 we find .954, .962, and .951; where we expect .995 we find .993, .997, and .994; and where we expect .9995 we find 1.000.

(2) The experimental α-error (the number of incorrectly found small parameters, divided by the number of small parameters) ranges from 0.0 to .41; this maximum is reached in the experiment with $\beta_{86}=4$, $\epsilon=.0005$. The latter result may be regarded as inadmissible. Let us see what happened. According to our procedure, we stop investigating a branch in the bifurcation tree, when the difference between the minimum upper estimate and the maximum lower estimate is smaller than $\sigma(\delta-x)$, where x is given by Bechhofer's (1954) table I. For $\epsilon=.0005$ this table has entries ranging from 5.4432 to 5.7924; see our table 3.1. Now we have the situation that $x>\delta$ for $\delta=4$, so we go on investigating, even when the minimum upper estimate is

smaller than the maximum lower estimate, e.g. when $y_N < y_0$. If we want to avoid this peculiar situation, we have to demand that $x \leq \delta$, and by doing so, we introduce a relation between δ and ε, via Bechhofer's table I. If for $\delta=4$ we demand that $x \geq 4$, we drop the experiments $\delta=4$, $\varepsilon=.0005$ and $\delta=4$, $\varepsilon=.005$, and then the maximum experimental α-error is $4.418/255=.017$ for $\beta_{86}=\delta=4$, $\varepsilon=.05$.

(3) The number of runs is extremely small. Notice that if there were no noise, then we would need two observations to find out that no factor is important; we would need $m+2=10$ observations to find out that one factor is important. If we drop the experiments $\delta=4$, $\varepsilon=.0005$ and $\delta=4$, $\varepsilon=.005$, then the number of experiments in case all effect parameters are zero, ranges from 2.0 to 9.6; in case of one important parameter the number of experiments ranges from 9.9 to 26.4. One might wonder how it is possible that with one important variable the number of observations is less than 10. Note that this only occurs for $\varepsilon_1=.05$, where in 95% of the cases we must find the important factor (and use 10 observations), in 5% of the cases we may find no important factors (and might use only two observations). The mean minimum number of observations is then $0.95*10 + 0.05*2 = 9.6$.

(4) The case $\beta_1=\delta$ is treated slightly more efficient than $\beta_{86}=\delta$ and $\beta_{241}=\delta$, from the viewpoint of experimental α-error and number of observations.

5. CONCLUSION

Sequential Bifurcation is not only a very efficient tool for factor screening in the absence of random error; it is also capable of coping with

observations with random errors in a very efficient way. This efficiency depends on the signal-to-noise ratio.

No attention is yet payed to questions such as: how does SB handle

- numbers of input variables not equal to some power of two;
- small negative factor effects;
- normally distributed error with unknown variance;
- non-normally distributed error.

Research on these topics is to going on.

ACKNOWLEDGMENTS

I am grateful to Jack Kleijnen en Peter Sander for our fruitful discussions, and to S.S. Gupta, who pointed out the possibility of adopting Bechhofer's method to our problem.

APPENDIX 1: THE OBSERVATIONS GOING WITH A PARTICULAR FACTOR

Suppose that, in the deterministic case, factor number ℓ, say, is important. At each stage of the procedure exactly one aggregated parameter contains β_ℓ: $\beta_{i|j}(=\Sigma_{k=(i-1)*2^{m-j}+1}^{i*2^{m-j}} \beta_k)$ contains β_ℓ iff $(i-1)*2^{m-j} < \ell \leq i*2^{m-j}$. As there are $N=2^m$ factors, we have $1 \leq \ell \leq 2^m$, and the binary form of $\ell-1$ uses m digits:

$$\ell - 1 = \sum_{t=1}^{m} \ell_t 2^{m-t} \qquad \text{with } \ell_t \in \{0,1\},\ t=1,\ldots,m. \tag{A1.1}$$

For any $j \in \{0,1,\ldots,m\}$ we can write

$$\ell = 2^{m-j} * \Sigma_{t=1}^{j} \ell_t 2^{j-t} + \Sigma_{t=j+1}^{m} \ell_t 2^{m-j} + 1. \tag{A1.2}$$

Group i at stage j contains the input variables with numbers $(i-1)*2^{m-j}+1$ through $i*2^{m-j}$, so it contains factor number ℓ iff $i-1 = \Sigma_{t=1}^{j} \ell_t 2^{j-t}$: $\beta_{i|j}$ contains β_ℓ iff $i=1+\Sigma_{t=1}^{j} \ell_t 2^{j-t}$ $(j=0,1,\ldots,m)$. Examples: take j=0, then i=1; take j=m, then $i=\ell$.

To arrive at $\beta_{1|0}$ we use observations $y_{0|0}$ and $y_{1|0}$. Be $0 \le j < m$, and suppose $\beta_{i|j}$ contains β_ℓ. Then either $\beta_{2i-1|j+1}$ or $\beta_{2i|j+1}$ contains β_ℓ. To split $\beta_{i|j}$ into $\beta_{2i-1|j+1}$ and $\beta_{2i|j+1}$ we use observation $y_{2i-1|j+1}$. Now $i=1+\Sigma_{t=1}^{j} \ell_t 2^{j-t}$, so $2i-1=1+2\Sigma_{t=1}^{j} \ell_t 2^{j-t}$. According to (2.6), $y_{i|j}=y_{i*2^{m-j}}$, so $y_{2i-1|j+1}=y_{(2i-1)*2^{m-j-1}}=y_{(1+2\Sigma_{t=1}^{j} \ell_t 2^{j-t})*2^{m-j-1}}$. To arrive at β_ℓ we use the observations

$$y_0,\ y_N,\ \text{and } y_{(1+2\Sigma_{t=1}^{j} \ell_t 2^{j-t})*2^{m-j-1}}, \qquad j=0,1,\ldots,m-1, \tag{A1.3}$$

where the numbers ℓ_t $(t=1,\ldots,m)$ are defined by (A1.1).

Example A1. Take m=3; N=8. To arrive at β_2 we need the following observations:

at level 0 we use $\beta^{+}_{1|0}=y_{1|0}=y_8$ and $\beta^{+}_{0|0}=y_{0|0}=y_0$ to compute $\beta_{1|0}$;
at level 1 we use $\beta^{+}_{1|1}=y_{1|1}=y_4$ and $\beta^{+}_{0|1}=\beta^{+}_{0|0}=y_{0|0}=y_0$ for $\beta_{1|1}$;
at level 2 we use $\beta^{+}_{1|2}=y_{1|2}=y_2$ and $\beta^{+}_{0|2}=\beta^{+}_{0|0}=y_{0|0}=y_0$ for $\beta_{1|2}$;

at level 3 we use $\beta^+_{2|3=}\beta^+_{1|2=}y_{1|2=}y_2$ *and* $\beta^+_{1|3=}y_{1|3=}y_1$ *for* $\beta_{2|3=}\beta_3$.
*At each of the four levels we use two observations. In total, however, we do not use 4*2=8 observations, but only (m+2=)5, namely* y_0, y_1, y_2, y_4 *and* y_8.

We see that, no matter how many observations the whole screening uses, m+2 observations, specified by (A1.3), are used for a particular important factor. For a particular unimportant factor at most m+2 observations are used: the observations of (A1.3) are used until we reach a stage, where the aggregated parameter, the particular parameter is part of, is small. This may be level 0 (and then we use only two observations for all factors), but if we are unlucky, this can be level m.

APPENDIX 2: THE DISTRIBUTION OF UPPER LIMITS

Bechhofer (1954) considers k normal populations with common variance σ^2, and computes the probability that, after taking a sample of size N from each population and ordering the sample means, the t populations with the largest sample means correspond with the t populations with the largest population means. This probability depends on k, t and $\sqrt{N}\lambda$, where $\lambda=(\mu_t-\mu_{t+1})/\sigma$ (μ_i is the i^{th} largest population mean). In the sequel we shall use the symbol $\wedge$ instead of $\sqrt{N}\lambda$. For k=2,...,10, t=1,...,[k/2] (the problem is symmetric in t and k-t) as well as k=11,...,15 and some selected values of t, Bechhofer tabulates $\wedge$ against the probability of correct ranking:

$$P(\min\{x_1,\ldots,x_t\}>\max\{x_{t+1},\ldots,x_k\}$$

$$\mid \min\{\mu_1,\ldots,\mu_t\}-\max\{\mu_{t+1},\ldots,\mu_k\}\geq\lambda\sigma \ \&\ Vx_i=\sigma^2/N \ (i=1,\ldots,k)). \quad (A2.1)$$

By multiplying all variables by $\sqrt{N}/\sigma$, we see that (A2.1) is equal to

$$P(\min\{x_1,\ldots,x_t\}-\max\{x_{t+1},\ldots,x_k\}>0$$
$$\mid \min\{\mu_1,\ldots,\mu_t\}-\max\{\mu_{t+1},\ldots,\mu_k\}\geq\wedge \ \&\ Vx_i=1 \ (i=1,\ldots,k)). \quad (A2.2)$$

The least favorable configuration is the one with $\mu_1=\ldots=\mu_t$ and $\mu_{t+1}=\ldots=\mu_k$, and we may take $x_i=\wedge+e_i$ $(i=1.,,,.t)$; $x_i=e_i$ $(i=t+1,\ldots,k)$, and replace (A2.2) by

$$P(\wedge+\min\{e_1,\ldots,e_t\}-\max\{e_{t+1},\ldots,e_k\}>0 \mid Ve_i=1 \ (i=1,\ldots,k)). \quad (A2.3)$$

As said, this probabitity is tabulated against $\wedge$. Now we consider (3.3):

$$\beta_\ell+\min\{e_{i_{m+1}},\ldots,e_{i_{L+1}}\}-\max\{e_{i_L},\ldots,e_{i_0}\}, \quad (A2.4)$$

where $e_{i_0},\ldots,e_{i_{m+1}}$ are NID$(0,\sigma^2)$. (A2.4) equals

$$\beta_\ell-\sigma\wedge+\sigma(\wedge+\min\{e_{i_{m+1}}/\sigma,\ldots,e_{i_{L+1}}/\sigma\}-\max\{e_{i_L}/\sigma,\ldots,e_{i_0}/\sigma\}). \quad (A2.5)$$

The value $\wedge_0$ for which $\wedge+\min\{e_{i_{m+1}}/\sigma,\ldots,e_{i_{L+1}}/\sigma\}-\max\{e_{i_L}/\sigma,\ldots,e_{i_0}/\sigma\}>0$ with probability p_0 can be found in Bechhofer's table I, where for k we take m+2, and for t we take min{L+1,m-L+1}. So now we have

$$\beta_\ell+\min\{e_{i_{m+1}},\ldots,e_{i_{L+1}}\}-\max\{e_{i_L},\ldots,e_{i_0}\} > \beta_\ell-\sigma\wedge_0 \quad \text{wp } p_0. \quad (A2.6)$$

REFERENCES

Bechhofer, R.E. (1954), "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances", *Annals of Mathematical Statistics*, 25, 16-39.

Bettonvil, B. (1988a), "Sequential Bifurcation for Factor Screening", in *Operations Research Proceedings 1987*, eds. H.Schellhaas, P.van Beek, H.Iserman, R.Schmidt and M.Zijlstra, Springer Verlag, Heidelberg, 444-450.

Bettonvil, B. (1988b), "Factor Screening by Sequential Bifurcation", submitted for publication.

Bettonvil, B. (1988c), "Sequential Bifurcation: the Design of a Factor Screening Method", submitted for publication.

Jacoby, J.E. and Harrison, S. (1962), "Multi-Variable Experimentation and Simulation Models", *Naval Research Logistic Quarterly*, 9, 121-136.

Kleijnen, J.P.C. (1987), *Statistical Tools for Simulation Practitioners*, New York: Marcel Dekker.

Table 3.1. Limit values for various group sizes 2^m and probabilities p_0

p_0	m=3 L=0,L=3	m=3 L=1,L=2	m=8 L=0,L=9	m=8 L=1,L=8	m=8 L=2,L=7	m=8 L=3,L=6	m=8 L=4,L=5
.9995	5.1661	5.3127	5.4432	5.6425	5.7343	5.7788	5.7924
.995	4.2394	4.4138	4.5524	4.7878	4.8950	4.9468	4.9625
.95	3.0552	3.2805	3.4182	3.7198	3.8541	3.9184	3.9378

(Source: Bechhofer 1954, p. 30-34)

Table 3.2.1. Performance of SB with $\beta_1=\delta$.

	Power (1)			α-error times 255 (2)			Observations (3)		
	$\varepsilon=.05$	$\varepsilon=.005$	$\varepsilon=.0005$	$\varepsilon=.05$	$\varepsilon=.005$	$\varepsilon=.0005$	$\varepsilon=.05$	$\varepsilon=.005$	$\varepsilon=.0005$
$\delta=10$	.954	.993	1.000	.000	.000	.000	9.9	10.0	10.0
$\delta=8$	.954	.993	1.000	.000	.000	.005	9.9	10.0	10.2
$\delta=6$	.954	.993	1.000	.007	.150	2.038	10.2	12.1	19.9
$\delta=4$	.954	.993	1.000	2.831	30.673	97.243	22.4	68.9	138.5

(1) Average number of times β_1 is found; experimental power.

(2) Average number of incorrectly found parameters; the experimental α-error is this number divided by 255 (the number of unimportant factors).

(3) Average number of observations.

Table 3.2.2. Performance of SB with $\beta_{86}=\delta$.

	Power (1)			α-error times 255 (2)			Observations (3)		
	ε=.05	ε=.005	ε=.0005	ε=.05	ε=.005	ε=.0005	ε=.05	ε=.005	ε=.0005
δ=10	.962	.997	1.000	.000	.000	.000	9.9	10.0	10.0
δ=8	.962	.997	1.000	.000	.002	.016	9.9	10.0	10.3
δ=6	.962	.997	1.000	.019	.377	3.217	10.3	13.1	23.3
δ=4	.962	.997	1.000	4.418	36.523	105.366	26.4	77.1	146.9

(1) Average number of times β_{86} is found; experimental power.

(2) Average number of incorrectly found parameters; the experimental α-error is this number divided by 255 (the number of unimportant factors).

(3) Average number of observations.

Table 3.2.3. Performance of SB with $\beta_{241}=\delta$.

	Power (1)			α-error times 255 (2)			Observations (3)		
	ε=.05	ε=.005	ε=.0005	ε=.05	ε=.005	ε=.0005	ε=.05	ε=.005	ε=.0005
δ=10	.951	.994	1.000	.000	.000	.000	9.9	10.0	10.0
δ=8	.951	.994	1.000	.000	.002	.018	9.9	10.0	10.3
δ=6	.951	.994	1.000	.028	.397	2.921	10.3	12.9	21.1
δ=4	.951	.994	1.000	3.810	31.941	97.725	23.7	68.8	138.0

(1) Average number of times β_{241} is found; experimental power.

(2) Average number of incorrectly found parameters; the experimental α-error is this number divided by 255 (the number of unimportant factors).

(3) Average number of observations.

Table 3.2.4. Performance of SB with $\beta_1=\ldots=\beta_{256}=0$.

	α-error times 256 (2)			Observations (3)		
	ε=.05	ε=.005	ε=.0005	ε=.05	ε=.005	ε=.0005
δ=10	.000	.000	.000	2.0	2.0	2.0
δ=8	.000	.000	.001	2.0	2.0	2.1
δ=6	.001	.032	.759	2.2	3.1	7.7
δ=4	1.135	20.168	80.849	9.6	46.0	116.6

(2) Average number of incorrectly found parameters; the experimental α-error is this number divided by 256 (the number of unimportant factors).

(3) Average number of observations.

IN 1988 REEDS VERSCHENEN

297 Bert Bettonvil
Factor screening by sequential bifurcation

298 Robert P. Gilles
On perfect competition in an economy with a coalitional structure

299 Willem Selen, Ruud M. Heuts
Capacitated Lot-Size Production Planning in Process Industry

300 J. Kriens, J.Th. van Lieshout
Notes on the Markowitz portfolio selection method

301 Bert Bettonvil, Jack P.C. Kleijnen
Measurement scales and resolution IV designs: a note

302 Theo Nijman, Marno Verbeek
Estimation of time dependent parameters in lineair models
using cross sections, panels or both

303 Raymond H.J.M. Gradus
A differential game between government and firms: a non-cooperative approach

304 Leo W.G. Strijbosch, Ronald J.M.M. Does
Comparison of bias-reducing methods for estimating the parameter in dilution series

305 Drs. W.J. Reijnders, Drs. W.F. Verstappen
Strategische bespiegelingen betreffende het Nederlandse kwaliteits-concept

306 J.P.C. Kleijnen, J. Kriens, H. Timmermans and H. Van den Wildenberg
Regression sampling in statistical auditing

307 Isolde Woittiez, Arie Kapteyn
A Model of Job Choice, Labour Supply and Wages

308 Jack P.C. Kleijnen
Simulation and optimization in production planning: A case study

309 Robert P. Gilles and Pieter H.M. Ruys
Relational constraints in coalition formation

310 Drs. H. Leo Theuns
Determinanten van de vraag naar vakantiereizen: een verkenning van materiële en immateriële factoren

311 Peter M. Kort
Dynamic Firm Behaviour within an Uncertain Environment

312 J.P.C. Blanc
A numerical approach to cyclic-service queueing models

313 Drs. N.J. de Beer, Drs. A.M. van Nunen, Drs. M.O. Nijkamp
Does Morkmon Matter?

314 Th. van de Klundert
Wage differentials and employment in a two-sector model with a dual labour market

315 Aart de Zeeuw, Fons Groot, Cees Withagen
On Credible Optimal Tax Rate Policies

316 Christian B. Mulder
Wage moderating effects of corporatism
Decentralized versus centralized wage setting in a union, firm, government context

317 Jörg Glombowski, Michael Krüger
A short-period Goodwin growth cycle

318 Theo Nijman, Marno Verbeek, Arthur van Soest
The optimal design of rotating panels in a simple analysis of variance model

319 Drs. S.V. Hannema, Drs. P.A.M. Versteijne
De toepassing en toekomst van public private partnership's bij de grote en middelgrote Nederlandse gemeenten

320 Th. van de Klundert
Wage Rigidity, Capital Accumulation and Unemployment in a Small Open Economy

321 M.H.C. Paardekooper
An upper and a lower bound for the distance of a manifold to a nearby point

322 Th. ten Raa, F. van der Ploeg
A statistical approach to the problem of negatives in input-output analysis

323 P. Kooreman
Household Labor Force Participation as a Cooperative Game; an Empirical Model

324 A.B.T.M. van Schaik
Persistent Unemployment and Long Run Growth

325 Dr. F.W.M. Boekema, Drs. L.A.G. Oerlemans
De lokale produktiestructuur doorgelicht.
Bedrijfstakverkenningen ten behoeve van regionaal-economisch onderzoek

326 J.P.C. Kleijnen, J. Kriens, M.C.H.M. Lafleur, J.H.F. Pardoel
Sampling for quality inspection and correction: AOQL performance criteria

327 Theo E. Nijman, Mark F.J. Steel
Exclusion restrictions in instrumental variables equations

328 B.B. van der Genugten
Estimation in linear regression under the presence of heteroskedasticity of a completely unknown form

329 Raymond H.J.M. Gradus
The employment policy of government: to create jobs or to let them create?

330 Hans Kremers, Dolf Talman
Solving the nonlinear complementarity problem with lower and upper bounds

331 Antoon van den Elzen
Interpretation and generalization of the Lemke-Howson algorithm

332 Jack P.C. Kleijnen
Analyzing simulation experiments with common random numbers, part II: Rao's approach

333 Jacek Osiewalski
Posterior and Predictive Densities for Nonlinear Regression.
A Partly Linear Model Case

334 A.H. van den Elzen, A.J.J. Talman
A procedure for finding Nash equilibria in bi-matrix games

335 Arthur van Soest
Minimum wage rates and unemployment in The Netherlands

336 Arthur van Soest, Peter Kooreman, Arie Kapteyn
Coherent specification of demand systems with corner solutions and endogenous regimes

337 Dr. F.W.M. Boekema, Drs. L.A.G. Oerlemans
De lokale produktiestruktuur doorgelicht II. Bedrijfstakverkenningen ten behoeve van regionaal-economisch onderzoek. De zeescheepsnieuwbouwindustrie

338 Gerard J. van den Berg
Search behaviour, transitions to nonparticipation and the duration of unemployment

339 W.J.H. Groenendaal and J.W.A. Vingerhoets
The new cocoa-agreement analysed

340 Drs. F.G. van den Heuvel, Drs. M.P.H. de Vor
Kwantificering van ombuigen en bezuinigen op collectieve uitgaven 1977-1990

341 Pieter J.F.G. Meulendijks
An exercise in welfare economics (III)

342 W.J. Selen and R.M. Heuts
A modified priority index for Günther's lot-sizing heuristic under capacitated single stage production

343 Linda J. Mittermaier, Willem J. Selen, Jeri B. Waggoner, Wallace R. Wood
Accounting estimates as cost inputs to logistics models

344 Remy L. de Jong, Rashid I. Al Layla, Willem J. Selen
Alternative water management scenarios for Saudi Arabia

345 W.J. Selen and R.M. Heuts
Capacitated Single Stage Production Planning with Storage Constraints and Sequence-Dependent Setup Times

346 Peter Kort
The Flexible Accelerator Mechanism in a Financial Adjustment Cost Model

347 W.J. Reijnders en W.F. Verstappen
De toenemende importantie van het verticale marketing systeem

348 P.C. van Batenburg en J. Kriens
E.O.Q.L. - A revised and improved version of A.O.Q.L.

349 Drs. W.P.C. van den Nieuwenhof
Multinationalisatie en coördinatie
De internationale strategie van Nederlandse ondernemingen nader beschouwd

350 K.A. Bubshait, W.J. Selen
Estimation of the relationship between project attributes and the implementation of engineering management tools

351 M.P. Tummers, I. Woittiez
A simultaneous wage and labour supply model with hours restrictions

352 Marco Versteijne
Measuring the effectiveness of advertising in a positioning context with multi dimensional scaling techniques

353 Dr. F. Boekema, Drs. L. Oerlemans
Innovatie en stedelijke economische ontwikkeling

354 J.M. Schumacher
Discrete events: perspectives from system theory

355 F.C. Bussemaker, W.H. Haemers, R. Mathon and H.A. Wilbrink
A (49,16,3,6) strongly regular graph does not exist

356 Drs. J.C. Caanen
Tien jaar inflatieneutrale belastingheffing door middel van vermogensaftrek en voorraadaftrek: een kwantitatieve benadering

357 R.M. Heuts, M. Bronckers
A modified coordinated reorder procedure under aggregate investment and service constraints using optimal policy surfaces

358 B.B. van der Genugten
Linear time-invariant filters of infinite order for non-stationary processes

359 J.C. Engwerda
LQ-problem: the discrete-time time-varying case

360 Shan-Hwei Nienhuys-Cheng
Constraints in binary semantical networks

361 A.B.T.M. van Schaik
Interregional Propagation of Inflationary Shocks

362 F.C. Drost
How to define UMVU

363 Rommert J. Casimir
Infogame users manual
Rev 1.2 December 1988

364 M.H.C. Paardekooper
A quadratically convergent parallel Jacobi-process for diagonal dominant matrices with nondistinct eigenvalues

365 Robert P. Gilles, Pieter H.M. Ruys
Characterization of Economic Agents in Arbitrary Communication Structures

366 Harry H. Tigelaar
Informative sampling in a multivariate linear system disturbed by moving average noise

367 Jörg Glombowski
Cyclical interactions of politics and economics in an abstract capitalist economy

IN 1989 REEDS VERSCHENEN

368 Ed Nijssen, Will Reijnders
"Macht als strategisch en tactisch marketinginstrument binnen de distributieketen"

369 Raymond Gradus
Optimal dynamic taxation with respect to firms

370 Theo Nijman
The optimal choice of controls and pre-experimental observations

371 Robert P. Gilles, Pieter H.M. Ruys
Relational constraints in coalition formation

372 F.A. van der Duyn Schouten, S.G. Vanneste
Analysis and computation of (n,N)-strategies for maintenance of a two-component system

373 Drs. R. Hamers, Drs. P. Verstappen
Het company ranking model: a means for evaluating the competition

374 Rommert J. Casimir
Infogame Final Report

375 Christian B. Mulder
Efficient and inefficient institutional arrangements between governments and trade unions; an explanation of high unemployment, corporatism and union bashing

376 Marno Verbeek
On the estimation of a fixed effects model with selective non-response

377 J. Engwerda
Admissible target paths in economic models

378 Jack P.C. Kleijnen and Nabil Adams
Pseudorandom number generation on supercomputers

379 J.P.C. Blanc
The power-series algorithm applied to the shortest-queue model

380 Prof. Dr. Robert Bannink
Management's information needs and the definition of costs, with special regard to the cost of interest

381 Bert Bettonvil
Sequential bifurcation: the design of a factor screening method