Division of Research
Graduate School of Business Administration
The University of Michigan

August 1975

EMPIRICAL BAYES ANALYSIS OF LEASE AUCTIONS

Working Paper No. 116

by

R.W. Andrews

The University of Michigan

Introduction and Problem Statement

Many business and industrial transactions are decided through the use of sealed competitive bidding. This is common in contract bidding--e.g., construction jobs, defense contracts, plumbing and heating--in which the low-dollar bidder wins the contract and is entitled to perform the required services. It is also the case in acquiring property rights, for example, purchasing or leasing acreage for mineral or oil rights. In this case the high-dollar bidder wins the rights to the property.

Since Friedman's article [5] in 1954, much has been written about analyzing a competitive bidding situation. Most of what has been written is concerned with proposing a strategy for a bidder (see Arps [2], Brown [3], Friedman [5], and LaValle [7]). Notable exceptions are Crawford [4] and Pelto [10], who analyze bids without proposing a strategy. Stark [11] gives a comprehensive bibliography on competitive bidding.

Herein, the concern is with the analysis of a lease auction. Specifically, this report gives a procedure for estimating the bid-worth of a tract of land (or off-shore continental shelf) which will be used for oil and/or gas rights. Empirical Bayes as well as Bayes methods are employed and therefore other experiences of similar bidding situations are required. In the second section, a Bayes estimator with the associated Bayes risk is derived based on a multiplicative model for the bid worth of a lease. In section three, an Empirical Bayes estimator is proposed on the assumption that other experiences of similar bidding situations exist. The mean squared error (M.S.E.) of the Empirical Bayes estimator is compared with the M.S.E. of the maximum likelihood estimator and with the Bayes risk. Section three also investigates a case in which the prior distribution is assessed incorrectly. Since the Empirical Bayes procedure does not depend on an assessment of the prior distribution, we see that the Empirical Bayes estimator's M.S.E. is smaller than the risk that was realized using a Bayes procedure in conjunction with an incorrectly assessed prior distribution. In the concluding section data from the bid tabulations of U.S. offshore sales is analyzed and a strategy is suggested based on the sequence of Empirical Bayes estimators.

The development of this report first considers the following problem: There is a group of k tracts which will be let for bids. The bid worth of a tract is denoted by θ_j; for j = 1,2, ...k. We have the bids for all k tracts, where each bid is denoted by X_{ij}, the $i^{\underline{th}}$ bid on the $j^{\underline{th}}$ tract. Therefore, the problem is structured as

$$\begin{pmatrix} \theta_j \\ \tilde{X}_j \end{pmatrix}; \; j = 1, 2, \ldots. k$$

with

$$\tilde{X}_j = (X_{1j}, X_{2j}, \ldots. X_{n_j j})'.$$

In the second section we establish the distribution of the bids and find a connection between the parameters of the distribution of X_{ij} and the bid worth θ_j. The geometric mean of $\tilde{X}_j$, designated by

$$T_j = \left(\prod_{i=1}^{n_j} X_{ij} \right)^{1/n_j},$$

is sufficient for θ_j; therefore the problem is reduced to

$$(1.1) \qquad \begin{pmatrix} \theta_j \\ T_j \end{pmatrix}; j = 1, 2, \ldots k.$$

Using Empirical Bayes methods we estimate $\theta_1, \theta_2, \theta_3, \ldots \theta_k$, the bid worth of all the tracts. Enroute to the Empirical Bayes solution a Bayes solution is given which enables us to evaluate the accuracy of the Empirical Bayes procedure using Monte Carlo studies.

One of the benefits in solving the problem as stated above is that it leads to the next step, a bidding strategy. Throughout this report it is assumed that some of the $\tilde{X}_j$'s have an element, denoted by X_{0j}, which is the bid of the individual company or bidding combine which is attempting to assess $\theta_{k+1}, \theta_{k+2}, \ldots \theta_{k+r}$, the bid worth of the next r tracts. Using the Empirical Bayes solution, a sequence of bids $\tilde{X}_0 = (X_{0,k+1}, X_{0,k+2}, \ldots X_{0,k+r})$ is proposed.

A Bayes Estimator for Bid Worth

Consider the situation in which the bid worth, designated by θ, of an item or lease is formed by the product of m+1 different factors. If $D_1, D_2, \ldots, D_m$, and F represent these factors then

$$\theta = F \prod_{t=1}^{m} D_t .$$

The common factor, F, discounts the actual worth. For example, using the straight forward method of appraisal of individual tracts as given in Crawford [4], the discounted gross oil reserves would be determined by

$$\theta = FD_1D_2D_3D_4D_5D_6D_7$$

in which

D_1 = number of pays;

D_2 = average pay thickness;

D_3 = average length of reservoir;

D_4 = average width of reservoir;

D_5 = fraction of pore space occupied by hydrocarbons;

D_6 = fraction recovery;

D_7 = average worth per unit volume of oil produced.

Throughout this report we will be concerned with bid worth, and we assume that bid worth is a certain percentage of actual worth. This percentage, designated by F, is assumed the same for all bidders. The factor F is based on average deferment, interest rates, average rate of return before taxes, and dry hole risk.

Since a bidder cannot assess $D_1, D_2, \ldots, D_m$ exactly, these quantities must be estimated by $\hat{D}_1, \hat{D}_2, \ldots, \hat{D}_m$ which are assumed to be unbaised estimators. The resulting bid is

$$X = F \prod_{t=1}^{m} \hat{D}_t .$$

If the estimators $\hat{D}_1, \hat{D}_2, \ldots, \hat{D}_m$ are assumed independent, we have

(2.1) $E[X] = \theta.$

Extensive studies with bids for oil and mineral rights indicate that for a specific tract the bids are lognormally distributed [1], [4], [10]. If $X_1, X_2, \ldots X_n$ are n bids for a specific tract, we state that this sequence of random variables is independent and identically distributed lognormally with parameters μ and σ^2, and with σ^2 known. In practice, σ^2 can be estimated from the bids on all k tracts.

Using (2.1) we can write that for a given θ

$$X_1, X_2, \ldots X_n \sim \text{i.i.d. } \Lambda\,(\ln\theta - .5\sigma^2, \sigma^2),$$

and therefore

$$T|\theta \sim \Lambda\,(\ln\theta - .5\sigma^2, \sigma^2/n).$$

Since θ is formed by a multiplicative process, a lognormal prior, $G(\theta)$, is assigned,

$$\theta \sim \Lambda\,(\alpha, \beta^2).$$

We find that the resulting Bayes estimator of θ under squared error loss is

$$\hat{\theta}_B = T^{\left(\frac{n\beta^2}{\sigma^2+n\beta^2}\right)} \exp\left\{\frac{2\alpha\sigma^2+n\beta^2\sigma^2+\beta^2\sigma^2}{2\sigma^2+2n\beta^2}\right\}$$

with a resulting Bayes risk given by

$$\text{(2.2)} \qquad R(G) = \exp\left\{2\alpha + 2\beta^2\right\}\left[1 - \exp\left\{\frac{-\beta^2\sigma^2}{\sigma^2+n\beta^2}\right\}\right].$$

(In the next section, this Bayes risk is compared with the M.S.E. of the proposed Empirical Bayes estimator through the use of Monte Carlo studies.)

The resulting risk can sometimes be evaluated if the prior distribution is assessed incorrectly. For example, in using the log-normal prior, assume that β^2 is assessed exactly; however choose $\alpha + \varepsilon$ for the first parameter instead of the correct α. In this case the estimator we would use would be

$$\hat{\theta}_B^* = \hat{\theta}_B \exp\left\{\frac{\varepsilon\sigma^2}{\sigma^2 + n\beta^2}\right\} ,$$

with the resulting risk

$$\text{(2.3)} \qquad \text{Risk} = \exp\left(2\alpha+2\beta^2\right) \left[1 - \exp\left\{\frac{\varepsilon\sigma^2-\sigma^2\beta^2}{\sigma^2+n\beta^2}\right\}\left(2 - \exp\left\{\frac{\varepsilon\sigma^2}{\sigma^2+n\beta^2}\right\}\right)\right] .$$

This risk will be used to demonstrate the worth of the Empirical Bayes procedure when similar experiences are available.

An Empirical Bayes Estimator for Bid Worth

In bidding for oil and gas rights, usually a group of tracts is announced for sale. A collection of companies or consortiums bid for the rights to these tracts. After bids for these k tracts have been received, the geometric mean of the bids for each tract can be found, and the problem can be represented as given in (1.1) with

$$T_i | \theta_i \sim \Lambda(\ln\theta_i - .5\sigma^2, \sigma^2/n); \; i = 1,2,\ldots k.$$

The maximum likelihood estimator of θ_i adjusted for bias is

$$\hat{\theta}_i = T_i \exp\{.5\sigma^2 - (\sigma^2/2n)\}$$

and

(3.1) $$\hat{\theta}_j | \theta_i \sim \Lambda\,(\ln\theta_i - (\sigma^2/2n),\ \sigma^2/n).$$

Since we want to estimate θ_i, the bid worth of the $i^{\underline{th}}$ tract, for $i = 1,2,\ldots k$, we propose the Empirical Bayes estimator as given in Krutchkoff [6] and Lemon [8]:

(3.2) $$\tilde{\theta}_i = \frac{\sum_{j=1}^{k} \hat{\theta}_j\, f(\hat{\theta}_i | \hat{\theta}_j)}{\sum_{j=1}^{k} f(\hat{\theta}_i | \hat{\theta}_j)},$$

with

$$f(\hat{\theta}_i | \hat{\theta}_j) = (\sqrt{n}\,/\sigma\hat{\theta}_i\ \sqrt{2\pi})\ \exp\left\{-1/2 \left(\frac{\ln\hat{\theta}_i - \ln\hat{\theta}_j + \sigma^2/2n}{\sigma/\sqrt{n}}\right)^2\right\}.$$

Notice that in this problem we can use all the k experiences to find the Empirical Bayes estimator of the $i^{\underline{th}}$ experience, since the bids for all k tracts are collected at the same time. In this situation we can find an Empirical Bayes estimator of the first experience, i.e., θ_1, which would not be possible if the data were collected sequentially.

The Empirical Bayes estimator weights the classical estimators, $\hat{\theta}_j$'s, according to how close the data from the $j^{\underline{th}}$ experience is to the data from the $i^{\underline{th}}$ experience. The weighting used is the conditional likelihood of $\hat{\theta}_i$ given that $\theta_i = \theta_j$.

The following two definitions will enable us to comment on the general small sample properties of $\tilde{\theta}_i$:

$$(3.3) \qquad R = \frac{\text{M.S.E. of } \tilde{\theta}_i}{\text{M.S.E. of } \hat{\theta}_i} \text{ , and}$$

$$Z = \frac{\text{Var } (\hat{\theta}_i | \theta_i)}{\text{Var } (\theta_i)} \text{ .}$$

In the sequentially-collected data case, extensive Monte Carlo studies on (3.2), with a normal and binomial conditional distribution (Lemon [9]), have shown that the M.S.E. of $\tilde{\theta}_i$ is smaller than the M.S.E. of $\hat{\theta}_i$ for $k \geq 2$, i.e., if one or more other experiences exist. Furthermore, these studies have shown that for k increasing to approximately fifteen, R decreases. That is, the improvement levels off at approximately fourteen or fifteen other experiences. Figure 1 (see Figure 6 of Lemon [9]) demonstrates how R decreases as k increases. This plot was made using a normal conditional distribution. These extensive Monte Carlo studies also demonstrated that the amount of improvement of the Empirical Bayes estimator over the classical estimator is dependent on the prior distribution and the conditional distribution only as they effect Z, the ratio of the conditional variance to the prior variance.

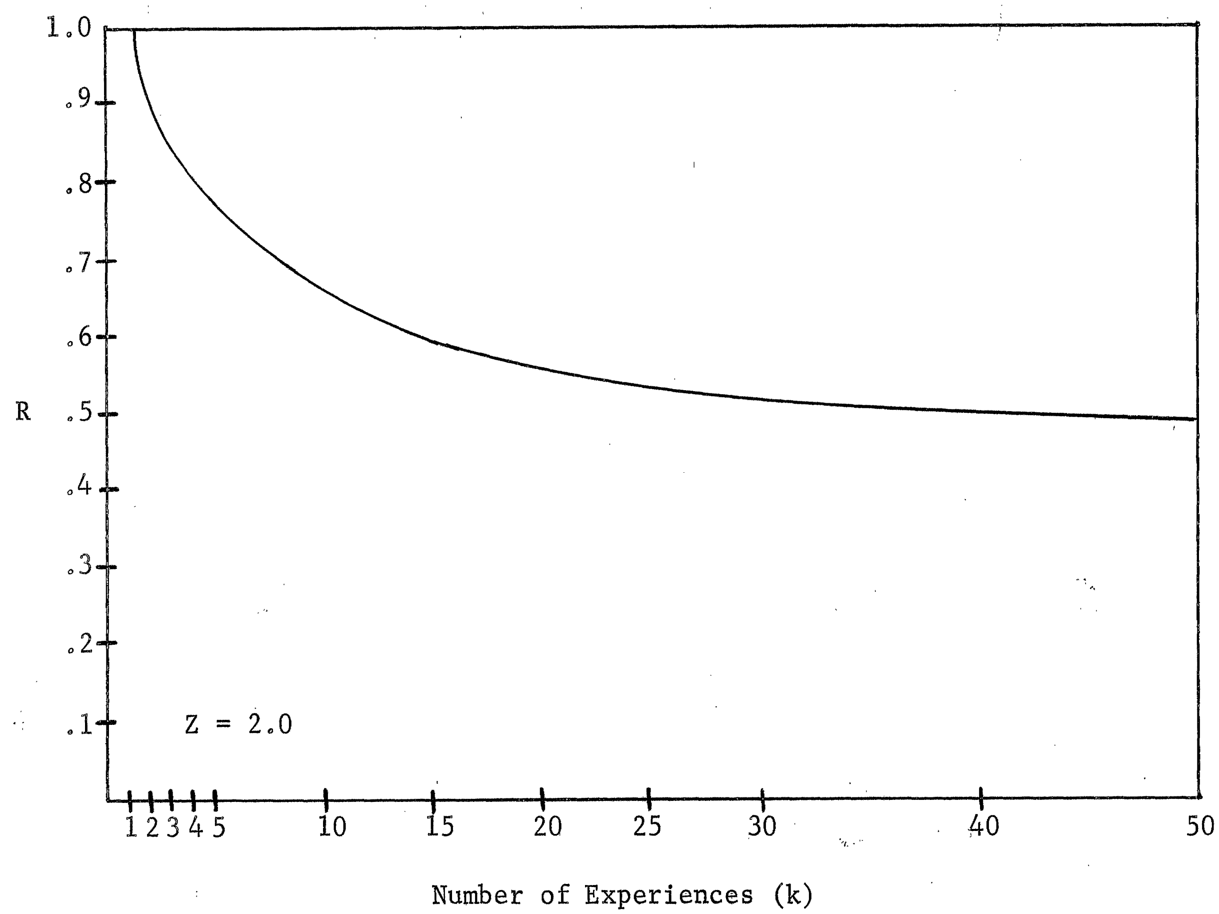

Fig. 1. Ratio of M.S.E.'s using a normal conditional distribution.

In further simulation studies, which were executed with a lognormal prior distribution and a lognormal conditional distribution of the form (3.1), the above comments held true with two exceptions. The plot of R versus the number of experiences was not always monotone decreasing since the value of the M.S.E. depended on the actual realized value of θ. Also, for the reduction in M.S.E. to be appreciable, the value of Z had to be greater than one. The output from a Monte Carlo simulation with 500 replications is given in Figure 2. For this run the estimated value of Z is 2.52. Z must be estimated since the conditional variance depends on the realized value of θ. Notice the large value of R at the thirty-eighth experience. This resulted because the realized value of θ for this experience was 2.17 standard deviations from the prior mean. This meant that the actual value of Z was 1.87. Since the Empirical Bayes estimator weights the other experiences, this also helps to account for the procedure's relatively poor showing against values of θ distant from the prior mean. Similar comments could be made for experiences five, twenty-nine, and forty-nine. Notice that the M.S.E. of the Empirical Bayes estimator is appreciably smaller than the M.S.E. of the classical estimator for almost all of the experiences.

Line B in Figure 2 is the ratio of the Bayes risk (2.2) to the expected mean squared error (E.M.S.E.) of $\hat{\theta}_j$, and line B* is the ratio of the realized risk of $\hat{\theta}^*_B$ (2.3) to the E.M.S.E. of $\hat{\theta}_j$. This realized risk was found by using $\hat{\theta}^*_B$, which is the Bayes estimator for the prior distribution Λ (5.4, .004), although the correct prior distribution should have been Λ (5.3, .004).

Table 1 gives the ratio B* for some values of ε, i.e., if Λ (5.3, .004) is the correct prior distribution and Λ (5.3+ε, .004) is used as the prior distribution, the ratio of the realized risk (2.3) to the E.M.S.E. of $\hat{\theta}_j$ is B*. Notice that in all but three experiences the Empirical Bayes M.S.E. is smaller than the realized risk if ε is .10. This indicates that if an error is made in assessing the prior distribution, the Empirical Bayes procedure can be substantially better than the procedure incorrectly assumed to be the Bayes procedure.

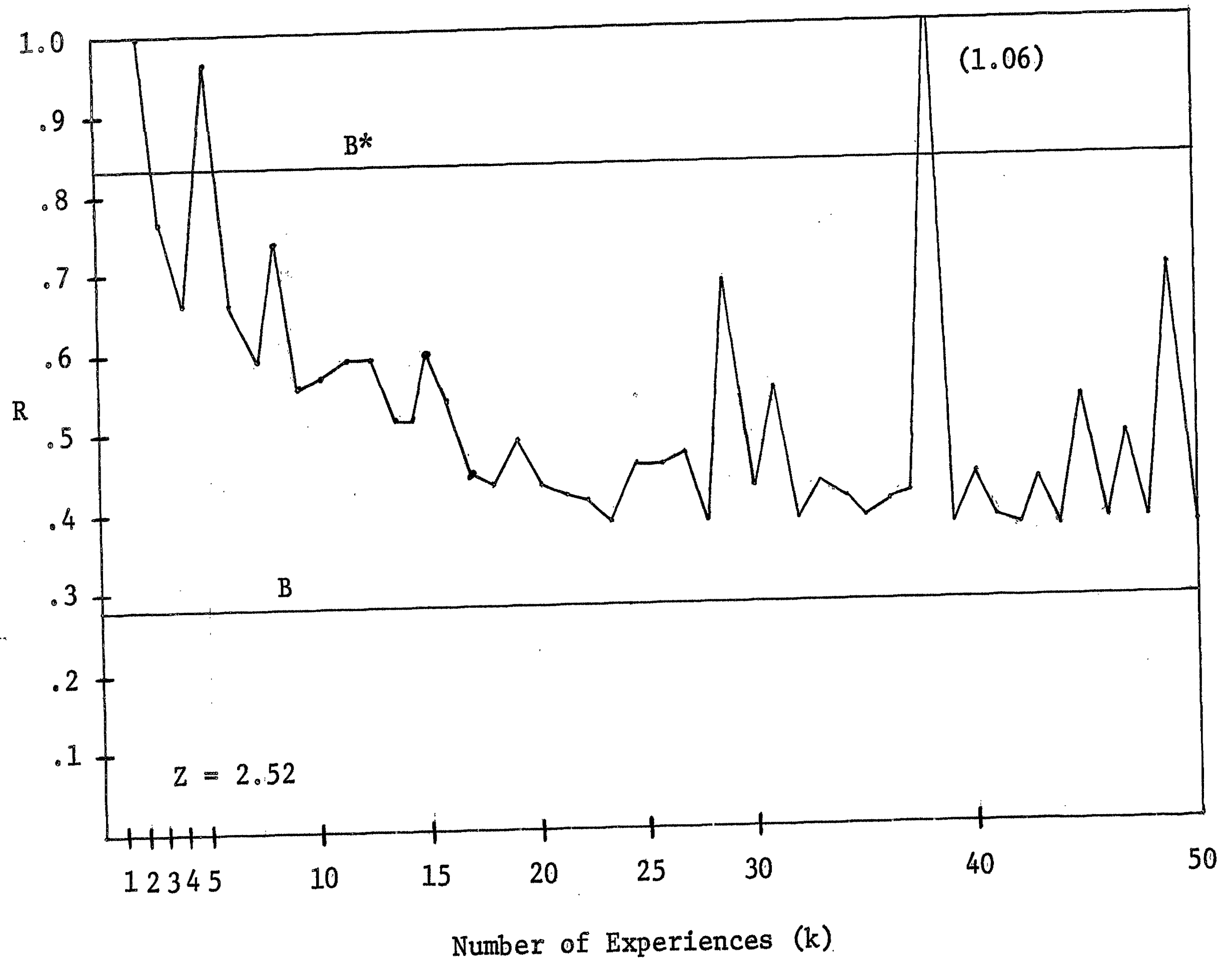

Fig. 2. Ratio of M.S.E.'s using a lognormal conditional distribution.

TABLE 1

B* = Realized Risk/E.M.S.E. of $\hat{\theta}_j$; $\theta \sim \Lambda$ (5.3, .004)

ε	-.20	-.10	-.05	0.0	.05	.10	.20
B*	2.04	.76	.41	.28	.42	.83	2.62

For k = 15, another Monte Carlo study was executed using (3.2) in which each $\bar{\theta}_j$ (j = 1,2...15) was estimated using the simulated observations in all fifteen experiences. This is representative of the outer continental shelf data which will be analyzed in the next section.

The value of Z for the data analyzed in the next section was estimated to be 1.49. Therefore, the estimated value of Z for this study was set at 1.49. For each of the fifteen experiences, Table 2 gives the R values (3.3) and the number of standard deviations that the realized θ is from the prior mean, designated by d. The largest value of R occurred when θ was 2.74 standard deviations from its mean. The ratio of the Bayes risk to the E.M.S.E. of the classical estimator was .387 for this study.

Outer Continental Shelf (OCS) Data with a Proposed Strategy

The Empirical Bayes estimation procedure was employed on tabulated bids of fifteen tracts [12]. The data is given in Table 3. For the value of σ^2 we used

$$\sigma^2 = k^{-1} \sum_{i=1}^{k} \frac{\sum_{j=1}^{n_j} (\ln X_{ij} - \overline{\ln X_j})^2}{n_j - 1}$$

with

$$\overline{\ln X_j} = n_j^{-1} \sum_{i=1}^{n_j} \ln X_{ij}.$$

TABLE 2

Reduction in M.S.E.; k = 15, Estimate of Z = 1.49

Experience No.	d	R
1	-.41	.59
2	-.31	.58
3	.00	.51
4	.40	.55
5	.51	.51
6	.50	.49
7	2.74	.88
8	-.11	.57
9	-.80	.76
10	1.32	.57
11	-.29	.59
12	-.21	.56
13	-.09	.54
14	-.60	.72
15	.17	.52

For the data in Table 3, $\sigma^2 = 1.81$. Since $E[\hat{\theta}_i|\theta_i] = \theta_i$ we can estimate the prior variance using [6]

$$\mathrm{Var}\theta = \mathrm{Var}\hat{\theta} - E[\mathrm{Var}(\hat{\theta}|\theta)].$$

For this data the approximate value of Z is 1.49 which indicates that the Empirical Bayes estimator will show some reduction in mean squared error but not as great as would have occurred if Z were larger.

Based on the sequence of Empirical Bayes estimates of bid worth, a strategy is proposed for the next set of tracts that are offered for sale. Since we have estimates of the θ_j's and since it has been substantiated that the bids, given θ, are lognormally distributed, we can determine a representative quantile for any bidder. A particular bidding firm should develop a bid in its usual manner; then, based on the previous empirical data, the firm should adjust this bid so that it will estimate the quantile position of the winning bid.

If u_q is the quantile of order q from a lognormal distribution with parameters (μ,σ^2), and v_q is the quantile of order q from the standard normal [1], then

$$u_q = \exp(\mu + v_q\sigma). \tag{4.1}$$

Consider the bids in Table 3 denoted with an A following each bid. These bids are from a specific bidding firm. Since

$$X_{ij} \sim \Lambda(\ln\theta_j - .5\sigma^2,\ \sigma^2),$$

we can solve for v_q in (4.1) by setting $\theta_j = \tilde{\theta}_j$ and $\sigma^2 = 1.81$. This can be accomplished for all of firm A's thirteen bids. The mean of these

TABLE 3

Tablulated Bids in Dollars Per Acre; $Z \approx 1.49$; $\sigma^2 = 1.81$

Tract No.	1	2	3	4	5	6	7	8
Bids	1003	2008	22010	13338	1225(B)	701(B)	10619	7541
	666(A)	1434	10619	10695(B)	1195(A)	580(A)	7905(A)	5405
	354	1369(A)	8919	1448(A)	382	274	5072(B)	3618(A)
	350(B)	894	4416(A)	1070	266	237	4581	1974
		583	3090	724	76		3172	1963
		577	2630(B)	301			2606	1402(B)
		350(B)	2230	266			1854	1303
		52	1073	53			1767	1073
			892				1133	73
			88				636	
Geometric Mean, T	536.4	619.1	2598.8	952.8	408.0	403.1	2871.1	1631.2
$\hat{\theta}$	1058.7	1368.6	5877.0	2106.3	842.7	795.7	6492.8	3651.9
$\tilde{\theta}$	1663.6	1629.0	5772.5	2627.8	1265.3	1349.9	5915.4	4439.7

Tract No.	9	10	11	12	13	14	15
Bids	36805	15932	1225(B)	908	8157	3706	14197
	19290(B)	3505(B)	681	701(B)	7616	2387	7184(B)
	7948	2008	389	244	7559(A)	1928(B)	2008
	5818(A)	1339(A)	76	198	5784(B)	1785(A)	1820(A)
	1360	631			2819	654	334
	1156	370			2681	307	35
	892	282			2087	177	
	88				2065	73	
					1111		
					367		
Geometric Mean, T	2824.6	1387.2	396.3	418.8	2855.7	728.1	1278.0
$\hat{\theta}$	6244.4	3017.5	782.2	826.6	6458.0	1609.6	2720.6
$\tilde{\theta}$	5751.7	3853.2	1335.4	1384.0	5908.7	1948.8	3581.6

thirteen v_q's is then used as an indication of the quantile for firm A and is designated $\overline{v}_q$. Similarly, the mean of the quantiles of winning bids can also be found and designated $\overline{v}_{WA}$. The A subscript denotes that we only considered the winning bids on those tracts for which firm A was bidding, in this case all tracts except numbers 11 and 12.

After all bids for a new sequence of tracts are developed by firm A, they are multiplied by

$$\exp\{\sigma(\overline{v}_{WA} - \overline{v}_A)\}$$

which adjusts the bid to the average quantile of the winning bids.

As a demonstration of this procedure we will reuse the data in Table 3 and indicate the results. This is a most favorable situation in that we are using the same bids for estimating the parameters and demonstrating the strategy.

In the case of firm A, $\overline{v}_A = .42$, $\overline{v}_{WA} = 1.17$, and $\exp\{\sigma(\overline{v}_{WA} - \overline{v}_A\} = 2.75$; so the bids firm A would have submitted using this strategy are:

Tract No.	1	2	3	4	5	6
A-Bid	1832	3765	12146	3983	3287	1595
	7	8	9	10	13	
	21741	9951	16002	36003	20790	
	14	15				
	4909	5005.				

This would have resulted in firm A winning the rights to eight tracts. In addition to the number of winning bids, to indicate how well a specific firm, say firm C, has bid we define:

$$W = \frac{\text{Sum of winning bids for firm C}}{\text{Sum of the } 2^{nd} \text{ place bids on the tracts that C won}}.$$

For A's revised bidding strategy, W = 1.94.

This same procedure was applied to the fifteen bids from firm B (see Table 3) with the resulting bids being:

Tract No.	1	2	3	4	5	6
B-Bid	764	764	5739	23338	2673	1530
	7	8	9	10	11	
	11068	3059	42093	7648	2673	
	12	13	14	15		
	1530	12622	4207	15676.		

These bids show that firm B would have won the rights to seven more tracts using this strategy. The value of W for B increased only slightly from 1.28 to 1.30 with the addition of these seven winning bids. The ratio of the sum of all fifteen winning bids to the sum of all the second place bids was 1.75.

Concluding Remarks

For a specific firm, there is usually a fixed total amount of capital that is available for a group of tracts [2], say Y dollars. Therefore, it is suggested that the initial bids be formulated under the restriction that $Y \exp\{\sigma(\overline{v}_A - \overline{v}_{WA})\}$ be the total amount exposed. Then the resulting bids will total Y dollars.

Since the success of this entire procedure depends heavily on accurate estimates of the θ_j's, this report has emphasized that the M.S.E. of the Empirical Bayes estimator, $\tilde{\theta}$, is smaller than the M.S.E. of the classical estimator. Simulation studies have shown that the M.S.E. of $\tilde{\theta}$ will be smaller if Z is larger. Therefore, this procedure will yield better results for larger values of Z.

References

[1] Aitchison, J. and Brown, J. A. C. The Lognormal Distribution. New York: Cambridge University Press, 1957.

[2] Arps, John J. "A Strategy for Sealed Bidding." Journal of Petroleum Technology, 17 (September 1965): 1033-9.

[3] Brown, K. C. "Bidding for Offshore Oil: Toward an Optimal Strategy." Journal of the Graduate Research Center, Southern Methodist University Press, 36 (October 1969): 1-71.

[4] Crawford, P. B. "Texas Offshore Bidding Patterns." Journal of Petroleum Technology, 22 (March 1970): 283-9.

[5] Friedman, L. "A Competitive Bidding Strategy." Operations Research, 4 (February 1956): 104-12.

[6] Krutchkoff, R. G. "Empirical Bayes Estimation." The American Statistician, 26 (December 1972): 14-16.

[7] LaValle, I. H. "A Bayesian Approach to an Individual Player's Choice of Bid in Competitive Sealed Auctions." Management Science, 13 (March 1967): 584-97.

[8] Lemon, G. H. and Krutchkoff, R. G. "An Empirical Bayes Smoothing Technique." Biometrika, 56 (1969): 361-5.

[9] Lemon, G. H. "Smooth Empirical Bayes Estimators: With Results for the Binomial and Normal Situations." Proceedings of the Texas Tech Symposium on Statistics (August 1969), pp. 110-40.

[10] Pelto, C. R. "The Statistical Structure of Bidding for Oil and Mineral Rights." Journal of the American Statistical Association, 66 (September 1971): 456-60.

[11] Stark, R. M. "Competitive Bidding: A Comprehensive Bibliography." Operations Research, 19 (1971): 484-90.

[12] U.S. Department of the Interior, Bureau of Land Management, New Orleans Office, Outer Continental Shelf Statistical Summary 1973-1975.