# DECISION THEORY AND SELECTING CANDIDATES :

# WITH REFERENCE TO HONG KONG

Yue Chim Richard Wong*

Economics Department
The Chinese University of Hong Kong
Shatin, Hong Kong
Tel.: (852) 603-6424 Fax.: (852) 603-5245

April 1985

## I. Introduction

The motivation for writing this paper came as a result of the recent decision by the Hong Kong Examinations Authority to release only the reported letter grades of the Higher Level Examination (HLE) and the Advanced Level Examination (ALE) to the two universities for the purpose of admitting candidates. Until recently the two universities had full access to the raw test scores of the HLE and ALE. As we shall see in the rest of the paper, this amounts to reducing the information that can be used to assess candidates. Given these new circumstances, two issues should be studied. First, how can the universities maximize the utility of the reduced information. Second, what will be the likely response of the universities to the reduction of candidate information.

This paper is concerned with the first issue. When raw test scores were available to the universities, applicants were ranked according to the sum of their standard deviation scores (to be defined later) for a specified number of subjects. We shall show that this is indeed the best approximation available, and is used in American examinations like the Scholastic Aptitude Test (SAT), the Test of English as a Foreign Language (TOEFL), and the Graduate Record Examination (GRE). However, when raw test scores became no longer available, the universities appeared to have assigned ad hoc values to different letter grades, and then to use the summed values to rank students. We show that such ad hoc values can introduce more noise into the signals and lead to a worsening of the information available. We also propose a method to maximize the informativeness of the letter grades from the HLE and ALE.

The second issue is clearly as important. Although it is not our focus here, some preliminary thoughts can be outlined. How will the two universities respond to this challenge? Clearly a rational response must involve the search for alternative sources of information about the candidate. It may be instructive to consider the situation of the secondary schools in Hong Kong at the time when they admit students into the sixth- form. The secondary schools have never had access to the raw test scores of the Certificate of Education Examination (CEE). They simply used the letter grades as the basis for choosing candidates. As the candidate pool grew with the extension of educational opportunities, but grades continued to be reported in the same manner, the informativeness of these grades worsened, and the schools began to emphasize other criteria for admission. In the more sought after schools, the requirements in terms of CEE grades for candidates who spent the pre-sixth form years at a different school became increasingly demanding relative to those who had studied at the same school. This is most understandable because the schools have more information about their own students than those from other schools. Such private information became increasingly valuable as the public information available in CEE grades continued to deteriorate in informativeness. Over time the proportion of own students in the sixth-form continued to rise in the more sought after schools. Indeed this was one of the reasons that contributed to the increasing pressure for government to intervene in the admissions policy of the secondary schools, which of course took place, although there were other precipitating reasons as well.

Unlike the secondary schools, the universities have no students of their own prior to admissions. They have, therefore, little or no private information. Since students are an essential input into the educational process, additional uncertainty about their quality must be detrimental to the interest of the universities. Economic theory suggests that uncertainty about input quality may result in backward intergration of the firm. We are not suggesting that the universities are or will start running their own secondary schools. This seems highly unlikely. The universities can also invest in acquiring more information about candidate quality, but this can be extremely expensive. Furthermore, if other criteria begins to play an important role, then the universities are liable to be criticized for unfair discrimination. A charge which heavily subsidized universities are particularly vulnerable to. It appears that the choices are quite limited for the two universities. Unless the universities can have access to better information about a candidate's examination scores, the importance of maximizing the utility of existing information becomes paramount and this is the focus of this paper.

Decision theory is used here to characterize the process of selecting quality candidates. Information about quality is transmitted imperfectly through observable signals. Our discussion focuses on the role of test scores as the signalling device for selecting students for admission to a higher level of schooling, with particular reference to the situation in Hong Kong. The problem of multiple test scores for each candidate can result in an overwhelming volume of information so that the decision maker has to find methods of aggregating information. This introduces noise into the information which can lower the expected payoff of the decision maker according to Blackwell's Theorem. The link between expected payoff changes and information contamination is the major reason for using decision theory to study the problem. It permits a rigorous statement as to why introducing noise into the information obtained by the decision maker is usually undesirable. Decision theory is also useful in defining key concepts so that our thinking will not be troubled by semantic differences.

The most important issue we consider are ways to aggregate information that are less likely to introduce too much noise. In particular we show how certain aggregation methods used in Hong Kong and perhaps elsewhere can be improved. The discussion is at times more formal than is desirable, but is essential for a clear and precise understanding of the nature of the issues treated in the paper. Some effort has been made so that the essential ideas are not lost to the general reader. Although the paper uses Hong Kong as an expository case, it can clearly be applied with some modifications to other places.

## II. The Decision Model

### (a) Description

Assume a decision maker who attempts to maximize expected payoffs through taking some actions. For him the feasible actions are either to admit a candidate or to reject him. The payoff to the decison maker is assumed to be positively related to the quality of the admitted candidate and negatively related to that of the rejected candidate.[1] The decision maker does not observe the quality of the candidate, but has a set of prior (subjective) probabilities about the distribution of candidate quality in the sample. To decide on a candidate the decision maker relies on observable scores obtained by the candidate in an examination. The score is positively correlated with the quality of the candidate. The strength of the correlation indicates the informativeness of the examination. It is possible to have a score for each subject that was taken in the same examination.[2]

Six items have been introduced so far: (1) candidate quality, (2) test scores, (3) feasible actions, (4) payoff function, (5) prior probabilities of candidate quality, and (6) correlation between quality and score. The decision maker takes these as given and proceeds to choose a decision set that will bring him the maximum payoff.

A formal statement is as follows: let the quality space, the score space, and the action space each be finite.

| | |
|---|---|
| $Q = \{q_1 , ..., q_j\}$ | set of candidate quality sequenced in descending order |
| $S = \{s_1 , ..., s_k\}$ | set of score values sequenced n descending order |
| $A = \{a_1 , ..., a_I\}$ | set of feasible actions; assume in our case that I = 2, let $a_1$ denote a candidate is admitted and $a_2$ denote a candidate is rejected |
| $U = [u_{ij}]$ | an I x J matrix; $u_{ij}$ is the payoff to the decision maker from taking action $a_i$ and getting a $q_j$ candidate |
| $P = [p_{jj}]$ | an J x J diagonal matrix; $p_{jj}$ is the prior probability assigned to quality $q_j$ |
| $Z = [z_{jk}]$ | an J x K Markov matrix; $z_{jk}$ is the conditional probability of observing score value $s_k$, given a $q_j$ candidate |
| $D = [d_{ki}]$ | an K x I Markov matrix; $d_{ki}$ is the conditional probability of choosing action $a_i$, given that score value $s_k$ is observed |

D is a decision function. The problem for the decision maker is to choose a D, given Q, S, A, U, P, and Z, such that he maximizes the expected payoff

$$\max_D \sum_i \sum_j \sum_k u_{ij} p_{jj} z_{jk} d_{ki} = \max_D \text{ tr } UPZD$$

(b) A Simple Example

Suppose the following:

(i) Two qualities, $q_1$ is the high quality and $q_2$ is the low quality.
(ii) Two scores, $s_1$ is the high score and $s_2$ is the low score.
(iii) Two actions, $a_1$ is admission and $a_2$ is rejection.
(iv) Payoff depends only on the quality of the admitted candidate, each $q_1$ is worth 10 and each $q_2$ is worth -20.

$$U = \begin{vmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{vmatrix} = \begin{vmatrix} 10 & 0 \\ -20 & 0 \end{vmatrix}$$

(v) The prior probability of $q_1$ is .3 and of $q_2$ is .7,

$$P = \begin{vmatrix} p_{11} & O \\ O & p_{22} \end{vmatrix} = \begin{vmatrix} .3 & 0 \\ 0 & .7 \end{vmatrix}$$

(vi) The relationship between score and quality is given by the conditional probabilities

$$Z = \begin{vmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{vmatrix} = \begin{vmatrix} \Pr\langle s_1 | q_1 \rangle & \Pr\langle s_2 | q_1 \rangle \\ \Pr\langle s_1 | q_2 \rangle & \Pr\langle s_2 | q_2 \rangle \end{vmatrix} = \begin{vmatrix} .7 & .3 \\ .2 & .8 \end{vmatrix}$$

(vii) The decision matrix has the following elements

$$D = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix} = \begin{vmatrix} \Pr\langle a_1 | s_1 \rangle & \Pr\langle a_2 | s_1 \rangle \\ \Pr\langle a_1 | s_2 \rangle & \Pr\langle a_2 | s_2 \rangle \end{vmatrix}$$

Since $d_{11} = 1 - d_{12}$ and $d_{21} = 1 - d_{22}$, the problem then is to find $d_{11}$ and $d_{21}$ so as to maximize tr UPZD. It is straight forward to show in this example that

$$\text{tr UPZD} = 6.3d_{11} - 2.7d_{21} - 3.3$$

Clearly the maximum is achieved by setting $d_{11} = 1$ and $d_{21} = 0$. That is admitting all those with a high score, $s_1$, and rejecting all those with a low score, $s_2$.

Suppose now that the decision maker considers rejecting a high quality candidate as a tremendous loss, then

$$U' = \begin{vmatrix} 10 & -1000 \\ -20 & 0 \end{vmatrix}$$

The maximizing solution will be to set $d_{11} = 0$ and $d_{21} = 0$. That is, only those with a low score, $s_2$, will be admitted.

Consider now that the decision maker has a different set of prior probabilities, say

$$P = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}$$

that is all candidates are perceived to be of uniformly high quality. It is easy to see that given such an assumption there is no difference to the decision maker which candidates to admit, since, all of them are perceived to be of the same quality. So long as high quality candidates generate positive payoffs, then all will be admitted.

Both U and P characterize the preference of the decision maker. Evidently differences among decision makers, as defined in terms of U and P, can result in different solutions for D, given the information structure Z.

## III. The Information Structure

(a) Minimal Restrictions

The matrix Z of conditional probabilities is the information structure. It gives the conditional probabilities of observing score $s_k$ given a $q_j$ candidate. Clearly the decision to admit or reject a candidate depends on the matrix Z. For simplicity we shall impose the following "minimal restrictions":

(i) A high quality candidate has a higher conditional probability of achieving a high score than (1) a low score, and (2) a low quality candidate.

(ii) A low quality candidate has a higher conditional probability of achieving a low score than (1) a high score, and (2) a high quality candidate.

Hence, the Z matrix can be partitioned into four regions. Recall that $q_j$ and $s_k$ have been sequenced in descending order.

$$Z = \begin{vmatrix} A & C \\ & \\ B & D \end{vmatrix}$$

Region A = $z_{jk} > z_{j+1,k}$ and $z_{jk} > z_{j,k+1}$ .

Region B = $z_{jk} > z_{j+1,k}$ and $z_{jk} < z_{j,k+1}$ .

Region C = $z_{jk} > z_{j,k+1}$ and $z_{jk} < z_{j+1,k}$ .

Region D = $z_{jk} < z_{j,k+1}$ and $z_{jk} < z_{j+1,k}$ .

An example for the 2 x 2 case is

$$Z = \begin{vmatrix} .8 & .2 \\ .3 & .7 \end{vmatrix}$$

Suppose the decision maker chooses to select only candidates of high quality and sets himself a quota equal to a constant fraction F of the total candidate pool. Given the structure of the information matrix above, this would result in a simple characterization of the decision maker's actions. All candidates with scores $s_k$ for $k < k^*$ will be rejected. $k^*$ is determined by the inequality below

$$\sum_{k=1}^{k^*-1} \sum_{j=1}^{J} p_{jj} \, z_{jk} < F$$

Some of the candidates with score $s_{k^*}$ may be admitted randomly if the strict inequality holds, so as to fill up the remaining slots to satisfy the quota F. If the equality sign holds then no candidate with score $s_{k^*}$ will be admitted. The simplicity of this decision rule will help us focus the discussion in the next section on information garbling.[3]

It may be desirable to specify more about the Z matrix than is assumed here, but this is in general rather difficult. Even these "minimal restrictions" can be easily violated. This will be the case if we assume that quality is multi-dimensional and not uni-dimensional. In our initial setup, the scores are ordinal measures that are monotonically related to an uni-dimensional quality. This permits candidates to be ranked unambiguously. If quality has multiple attributes, then candidates can only be ranked along one dimension, but not for all dimensions. In general, there is no reason to specify a priori whether each quality component is either correlated or uncorrelated with another quality component. One can think of this as an examination consisting of separate tests on a number of different subjects. A candidate's quality is measured in terms of a vector of scores which is not reducible to a single valued measure. Under these specifications, the information matrix Z will not satisfy the "minimal restrictions" above.

Clearly one must retain the uni-dimensional assumption about quality in order to rank students unambiguously. To handle the problem of multiple scores in an examination, one can assume that the score attained for each subject is simply a random draw from the same distribution of scores; the latter is determined by the quality of the candidate. That is we assume that the conditional probability of observing a particular score in a given subject is independent of his scores in other subjects. Hence, the conditional probability of observing any particular vector of scores is simply the product of the probability of each of the observed subject scores. It is worth emphasizing that the independence assumption is for the conditional probabilities and not the unconditional probabilities. Scores across subjects will clearly be correlated. The independence assumption is implicitly maintained when decision makers attempt to aggregate information, which is treated in part IV below.

(b) Garbling Up Information

Many public examinations do not report raw test scores. Candidates are given transformed scores on the basis of a mapping of S into S', where the number of elements in S' is often less then in S. For example, grades or quantiles are given instead of scores that range from 0 to 100. This alters the information structure. The original Z matrix is now replaced by Z'. Z' is a Markov matrix of order J x K' and is obtained by post-multiplying Z by some Markov matrix X of order K x K'. Since, K'<K then by construction the new Z' matrix will have fewer columns than Z. The following example illustrates how one can convert an information structure with 3 scores into one with 2 grades by aggregating the last two scores.

$$\underset{\substack{Z \\ (3 \times 3)}}{\begin{vmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{vmatrix}} \underset{\substack{X \\ (3 \times 2)}}{\begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{vmatrix}} = \underset{\substack{Z' \\ (3 \times 2)}}{\begin{vmatrix} Z_{11} & Z_{12} + Z_{13} \\ Z_{21} & Z_{22} + Z_{23} \\ Z_{31} & Z_{32} + Z_{33} \end{vmatrix}}$$

This is equivalent to making the assumption that the information conveyed by some of the signals are equivalent. If this is not the case. Then clearly some of the information is lost by making such assumptions. The significance of this for our purposes is revealed in Blackwell's Theorem.[4] It states that for any payoff function which the decision maker can have, his expected payoff cannot be increased by reducing the informativeness of the signals. That is garbling up information can never be in the interest of the maximizing decision maker.

Let us consider the decision rule established in the previous section. Two scores $s_k'$ and $s_k''$ are combined into a single score. If we have either $k'<k^*$, $k''<k^*$ or $k'>k^*$, $k''>k^*$ then the choice of candidates will not be affected by such information garbling. The expected payoff to the decision maker will be unaffected. The point here is that information garbling hurts only when decision outcomes are affected. Consider next an illustration of this result.

Over time education opportunities have expanded. The increasing size of the candidate pool has led many to assert that it has increased the proportion of low quality candidates. If the decision maker continues to admit the same fraction of the candidates as before, then over time the average quality of the admitted candidate must decline. This is not a novel point. However, if the same set of scores are used over time to record examination results and the distribution of these scores are artificially maintained to be constant over time, then there is considerable garbling of information at all levels. This happens in all public examinations in Hong Kong. All examination results are reported as grades that range from A to H; that is a total of eight grades. Over time the proportion of each grade in the total pool has remained largely unchanged. This means that each observed grade now contains a higher proportion of lower quality candidates than was previously the case. In particular, the cutoff grade, $s_{k*}$, for determining admission or rejection suffers from such information garbling and becomes less discriminating than previously. This loss of information creates an independent source for lowering the expected payoff of the decision maker and may, therefore, generate a demand for improving the accuracy of the examination or for refining the reporting of results, both of which are equivalent in our formal structure. In this connection we note the recent introduction of plus and minus grades by the Hong Kong Examinations Authority.

IV. Aggregating Information

(a) Multiple Subjects

In many examinations we have separate scores for different subjects. Suppose there are L subjects each with K scores. This would give the decision maker a total of $K^L$ different possible signals, which are conditional probabilities of different arrangements of the scores by subjects. Even for modest values of K and L this amounts to a phenomenal amount of information that has to be evaluated. Formally we shall have a Markov matrix Z of order J x $K^L$. Some simplification rules to aggregate information is clearly necessary. One common rule is to assume that the information coveyed through a particular score is equivalent regardless of the subject in which it was obtained. This reduces the Markov matrix Z to order J x KL. For K=3 and L=2, the number of possible signals is reduced from 9 to 6. The Z matrix now has 3 columns less. Some garbling of information has taken place.

For higher values of K and L some further reduction in the number of signals is typically deemed desirable despite the loss of informativeness. One obvious way is to assign values to each score, and to sum these values across subjects. This is a powerful way of aggregating (and so losing) information.

Suppose there are 2 subjects and 3 scores. Let the values assigned to the score be $s_1$=1, $s_2$=2, $s_3$=3. At first blush one may mistakenly conclude that these values are elements in the payoff matrix U. But this is clearly false, since, payoff depends on candidate quality and not scores. What summing values across subjects does is just another way of making the assumption that some signals transmit the same amount of information as other signals. This permits us to further reduce the number of columns in the Z matrix. To see this, for K=3 and L=2, we have a total of six signals which are the combinations: $(s_1, s_1)$, $(s_1, s_2)$, $(s_1, s_3)$, $(s_2, s_2)$, $(s_2, s_3)$, $(s_3, s_3)$. Using the proposed values for summing up, signals $(s_1, s_3)$ and $(s_2, s_2)$ will have the same aggregated value, which amounts to asserting that these two signals convey the same information. In this example the number of signals has been reduced from 6 to 5.

In our framework, finding a one-dimensional measure to rank students is equivalent to the more general problem of aggregating information. That is assuming different signals are similar; garbling up information.[5]

(b) When Is This Less Worrisome?

How justifiable is information aggregation depends on the (1) payoff function, (2) prior probabilities, (3) information structure, and (4) decision function.

If payoff functions assign almost equal weights to admitting (or rejecting) candidates of different qualities. Then it would not matter very much how information is garbled up. This is the decision maker with egalitarian tastes. If prior probabilities assigned by the decision maker are such that he believes that all candidates have the same quality, then again it would not matter very much how information is garbled up. This is the decision maker with egalitarian views.

A decision maker would also be less concerned by information garbling if it did not affect significantly his actions. For example, garbling up information at the low end of the score distribution would be of less concern to a decision maker who admits candidates at the high end.[6]

(c) Some Aggregating Methods

Consider a distribution of raw test scores with values ranging from 0 to 100. A great deal of evidence indicates that for large candidatures their distribution are almost normal. Note that this tells us nothing about the underlying distribution of quality, which is not defined in our framework. The prior probabilities are assigned by the decision maker and are therefore subjective. Perhaps it is best to consider the normally distributed scores as an artifact of the way most tests are structured. Hence, scores have meaning only as ordinal ranks. The shape of their distribution is immaterial unless we wish to compare scores across different subjects. Furthermore, one can always structure tests so that the scores may be distributed differently, for example, uniformly or bimodally.[7]

An aggregation method performs a transformation which allows us to rank candidates who have scores from multiple subject tests. Consider some such methods.

(i) A raw test score is transformed for each subject into a measure of the number of standard deviations from the mean of that subject. These standard deviation scores are summed across different subjects. If the raw test scores are each distributed approximately normally, such a transformation garbles up very little information, because all it does is to map the actual distribution of raw test scores for each subject into the almost identical normal distribution. Furthermore, the sum of normal variates is also normal. SAT, GRE, and TOEFL scores are similarly reported. These American examinations select a mean value of 500 and a standard deviation of 100, giving a range from 200 to 800 (See figure 1).

Figure 1

Normal distribution with $\mu = 500$ and $\sigma = 100$

(ii) A raw test score is transformed for each subject into a measure of the percentile ranking in that subject. This is equivalent to an order preserving mapping from the distribution of raw test scores into a uniform distribution (See figure 2). This results in considerable garbling of information when percentile scores are summed across subjects and used to compare different candidates. If candidates are chosen from the upper part of the distribution, then summed percentile scores would rank an individual with a greater dispersion of raw test scores lower than an individual with the same average raw test score but has a smaller dispersion. Hence, less qualified candidates are being selected because of information garbling. Percentile scores are also reported for the SAT, GRE, and TOEFL.

But clearly the use of standard deviation scores will be more appropriate when we wish to aggregate across subjects.

Figure 2

Mapping a normal distribution into a uniform distribution

(iii) For the British GCE (General Certificate of Education) and the Hong Kong CEE, HLE, and ALE, raw test scores are not reported. Instead, eight letter grades in descending order of excellence from A to H are given. Typically there is a decrease in the fraction of candidates in each grade category towards the extremes. Grade reporting is equivalent to a transformation of the distribution of raw test scores into a segmented distribution function with eight linear segments (See figure 3). The end points connecting each segment are determined by the cut off points for each grade and are located on the distribution function of raw test scores. Since raw test scores are not reported this distribution is not observed. An approximation can be constructed if we know the fraction of candidates in each grade and assume that the distribution of raw test scores is normal. This can be done for every subject. In figure 3 we constructed such a hypothetical distribution (denoted as I) using the normal distribution function. Note that the progressive reduction in the fraction of candidates in grades towards the tails reduces the amount of information garbling.

It is reasonable to conjecture that the approximation can be quite good even for eight reported grades. To actually sum grades across subjects, we must assign values for each grade. These values can be the mean scores of all candidates in each grade and can be imputed from the normal distribution function. It should be evident that this method is superior to using percentile scores, but somewhat inferior to standard deviation scores.

Figure 3

Two Segmented Distributions Associated
With Using Grades As Signals

It is worth emphasizing that the assigning of values to grades should not be ad hoc. Consider the case where the assigned values are A=1, B=2, H=8, we are now constructing a segmented distribution function, where the end points of the segments have the same cut off points along the horizontal axis for different grades, but along the vertical axis they occur at probabilities 1/8, 2/8, ....., 7/8, i.e., as multiples of the reciprocal of the number of grades. This is displayed in figure 3 as the distribution function denoted II. We have mapped the raw test scores into an arbitrary distribution function. Information can be seriously garbled as a consequence. This is apparently the procedure adopted by the secondary schools at the sixth form level and the universities in Hong Kong when admitting candidates.

## V. Concluding Remarks

This paper has demonstrated that the expected payoff of the decision maker is directly related to the informativeness of signals in revealing candidate quality. As such it provides an important link as to why information garbling should be of concern to the decision maker. The selection of candidates based on scores from multiple tests poses a severe problem for the decision maker, because it may result in an enormous amount of signals which would be impractical to evaluate without using some method to aggregate those information. Since this garbles up information, it is necessary that the chosen method minimizes such garbling. It was shown that expressing raw test scores as standard deviations from the mean to be an ideal measure. Alternatively, assigning appropriate values to letter grades is shown to be less ideal than using standard deviation scores, but preferable to assigning ad hoc values. Our results have important policy implications for all decision makers who use examination results to select candidates, and for examination authorities who determine how such results are reported.

## FOOTNOTES

*The paper was written while the author was a visiting scholar at the National Opinion Research Center, University of Chicago. I am indebted to the Hewlett Foundation and the Ming Yu Foundation for financial support.

[1]In some situations the expected payoff can be assumed to be unaffected by the quality of the rejected candidate. We also assume that the payoff function is separable with respect to each different kind of candidate quality. This amounts to saying that the payoff to one type of quality does not depend upon the amount of another type of quality.

[2]We may assume that each score is an independent realization of the quality of the candidates. The scores will be correlated with one another, but the conditional scores given the quality of the candidate will be uncorrelated with one another. This assumption is not necessary but is implicit in many attempts to make meaningful comparisons of scores across subjects.

[3]The term garbling was apparently first used by Marschak and Miyasawa (1968).

[4]Blackwell's theorem is a summary of results of several writers. The original work is Blackwell (1951). A simple elegant proof is in an unpublished paper (2 pages) by Nancy L. Stokey. Blackwell's theorem can easily be demonstrated with examples and will not be performed here.

[5]McGuire (1972) shows that any search for a one-dimensional measure of informativeness is in vain and proceeds to develop ways of reducing the dimensional problem and of making partial comparisons of information structures.

[6]This can be tricky, recall the earlier examples with two different payoff matrices U and U'.

[7]The utility of structuring examinations so that scores are distributed normally is an interesting subject in itself, but is not our concern here. For us the scores and the Z matrix are givens. Whether it is desirable to have a different Z matrix is a separate question, and is equivalent to asking how examinations should be structured.

# REFERENCES

1. Blackwell, D. (1951), "Comparisons of Experiments," in J. Neyman (ed.), Proceedings of the Second Berkeley Symposium on *Mathematical Statistics and Probability*, 2, 93-102.

2. Marschak, J. and K. Miyasawa (1968), "Economic Comparability of Information Systems," *International Economic Review*, 9, 2, 137-174.

3. McGuire, G. B. (1972), "Comparisons of Information Structures," in C. B. McGuire and R. Radner (ed.), *Decision and Organization*, North-Holland, 101-130.

4. Stokey, N.L. (undated), "*A Simple Proof of Blackwell's Theorem*."