Managing Risk in a Four-Digit Number Game
Author(s): Chung-Piaw Teo and Siew Meng Leong
Source: *SIAM Review*, Vol. 44, No. 4 (Dec., 2002), pp. 601-615
Published by: Society for Industrial and Applied Mathematics
Stable URL: http://www.jstor.org/stable/4148316
Accessed: 05-02-2016 01:05 UTC

SIAM REVIEW
Vol. 44, No. 4, pp. 601–615

Managing Risk in a Four-Digit Number Game*

Chung-Piaw Teo†
Siew Meng Leong‡

Abstract. The four-digit number game is a popular game of chance played in Southeast Asia. The players in this game choose a four-digit number and place their bets on it. In this paper, we study the design of a control mechanism for managing bets in this game. Our objective is to design a control mechanism to decide whether bets should be accepted or rejected. We propose a nonlinear optimization model for this problem and provide the mathematical justification for the control mechanism used by several operators in this region. We also suggest a simple improved control mechanism. Using data provided by a company in the region, we show that our control mechanism can accept more money per draw, while the risk exposure of the proposed mechanism can be considerably smaller than the current system.

Key words. portfolio optimization, gambling, four-digit number game, control mechanism

AMS subject classifications. 90C90, 91B30

PII. S0036144500366988

1. Description of the Four-Digit Number Game.

The four-digit number game is a game of chance. The objective of the game is to forecast a combination of four digits with a view to matching the winning combinations of numbers drawn. This popular game is played in several places around the world. For instance, in Florida, the game "PLAY 4" is a daily four-digit lottery in which the player chooses, or has the terminal randomly "Quick Pick," any or all of four numbers from 0 through 9. Each evening, an official drawing determines the winning PLAY 4 number for that day. If the player's four-digit number matches the official four-digit winning number, the player may win up to \$5,000, depending on the type of play purchased, the order of the four numbers drawn, and the amount wagered [1].

In Southeast Asia—for example, Singapore [2] and Malaysia [3]—the four-digit game is played in a slightly different way. It is usually played through a network of computer terminals installed at licensed agency outlets. The on-line terminals are linked to a central system by dedicated telephone lines. Players choose and bet on four-digit numbers selected from 0000 to 9999 to match against a set of 23 four-digit numbers drawn to win prizes. The selected numbers are purchased by marking on a

*Received by the editors January 31, 2000; accepted for publication (in revised form) June 2, 2002; published electronically October 30, 2002.
http://www.siam.org/journals/sirev/44-4/36698.html

†Department of Decision Sciences, NUS Business School, National University of Singapore (bizteocp@nus.edu.sg). This author was supported in part by a fellowship from the Singapore-MIT Alliance Program in High-Performance Computation for Engineered Systems.
‡Department of Marketing, NUS Business School, National University of Singapore (bizlsm@nus.edu.sg).

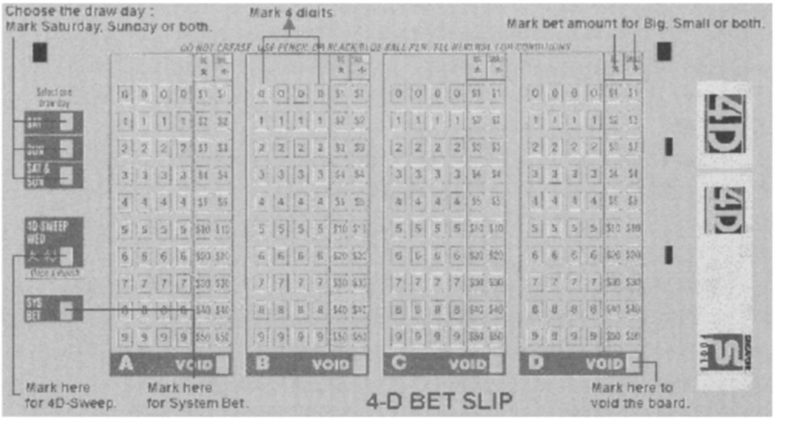

Fig. 1.1 *Betting slip used in Singapore.*

bet slip which is then fed into the computer terminals. A four-digit ticket is then issued. (See Figure 1.1 for a sample betting slip used in Singapore.)

The winning numbers are drawn on Wednesday, Saturday, and Sunday of every week, i.e., three draws per week. Bets for each draw are collected over a period of seven days preceding the draw. In each draw, winning four-digit numbers of 3 top prizes, 10 starter prizes, and 10 consolation prizes are determined. The customers can choose to place the bets for their numbers under the "BIG" or "SMALL" category. The prizes differ for the different categories of bets. For BIG bets, the four-digit number purchased is matched against the 23 four-digit winning numbers (1st, 2nd, and 3rd prizes, 10 starters, and 10 consolations) drawn to determine the winners. For SMALL bets, the four-digit number purchased is matched against the top 3 four-digit winning numbers (1st, 2nd, and 3rd prizes) drawn to determine the winners.

In Malaysia, for example, the prize structure for BIG and SMALL is

Prize	SMALL	BIG
1st	RM3,000	RM2,000
2nd	RM2,000	RM1,000
3rd	RM1,000	RM500
10 starter	—	RM200
10 consolation	—	RM60

The prize structure for the game played in Singapore is similar:

Prize	SMALL	BIG
1st	S$3,000	S$2,000
2nd	S$2,000	S$1,000
3rd	S$900	S$500
10 starter	—	S$250
10 consolation	—	S$65

At the time of this writing, U.S.$1 is equivalent to S$1.8 and RM3.8. The prizes are payable to the holders of the tickets bearing the set of four-digit numbers that exactly matches the four-digit number drawn for that prize. Both BIG and SMALL bets cost a minimum of S$1.00 per four-digit number for one draw. Each dollar on the winning number entitles the winner to one prize. For example, in Singapore, if a player bets S$2 on BIG for a number that wins the first prize and a starter prize in the same draw, the operator has to pay the player S$2000 $\times$ 2 + S$250 $\times$ 2 = S$4500.

The slip also contains various panels for bet selection. These are mainly of the following types (with slight variations in different countries):

- *Ordinary.* The basic form of entry to the four-digit game is the Ordinary Entry in which the customer selects any four-digit number ranging from 0000 to 9999.
- *System.* System Entry facilitates entry of multiple Ordinary Entries using one four-digit number. Essentially, it is the purchase of all permutations of a four-digit number. System Entry is not available for four-digit numbers where all the digits are the same, e.g., 0000, 1111, 2222, 3333, 8888, etc. Such numbers must be purchased under the Ordinary Entry.

The filled bet slip must be presented to the licensed agent who enters the data into a terminal and issues a ticket(s). The cost of a System Bet is the number of permutations of each four-digit number multiplied by the bet amount (either BIG, or SMALL, or both) multiplied by the number of draws selected.

1.1. Main Issues. The most interesting aspect of this game concerns the management of "hot numbers," which are numbers heavily bet by customers. To minimize the risk exposure, the game operators in Singapore and Malaysia (one legal operator in each country) limit the amount of bets accepted. If the amount available for sale (for each four-digit number) is less than requested, the system will only accept up to the sales limit and reject the excess amount. See Figure 1.2 for a sample output from the computer terminal when this arises.

As the betting system is centralized, the customers cannot approach any other retailer to bet on the same number. In Malaysia, for instance, the betting limit implemented is RM5,000 for BIG and RM3,000 for SMALL, respectively [3]. In Singapore, similar betting limits are imposed on the amount of bets for BIG and SMALL (Singapore Pools, private communication).

How does the operator determine whether to accept or reject the bets received? Currently, the operator will review periodically the betting profiles (i.e., all bets received, whether accepted or rejected) for the game and determine a cut-off level for the amount of bets to accept for each four-digit number on BIG and SMALL. The cut-off level is determined usually via computer simulation and managerial input. However, there is little mathematical justification to the choice of cut-off levels. In fact, it is not known whether there is a better way to manage the four-digit game. While the rejected bets are occasionally channeled to bets on other numbers in the same game, some revenue is often lost to illegal operators of the four-digit game. The main issue confronting the operator is to determine whether there is a better control mechanism that will accept more bets (to recover a portion of the lost revenues), but will not increase the risk exposure for the operator.

We address some of these issues in this paper. In particular, we provide mathematical justification for the current control mechanism and propose a new control mechanism which improves upon the current system. Interestingly, the naive approach to "accept all bets," which is good in the expected sense (i.e., it maximizes the expected return), fails to be a useful control mechanism due to the perceived danger of the system losing an "unacceptable" amount of money in the draws (the risk is too high). In this paper, we will focus only on the control mechanism for Ordinary Bets. Management of System Bets should not have any significant impact on the performance of the control policy, as the volume of system bets pales in comparison to the

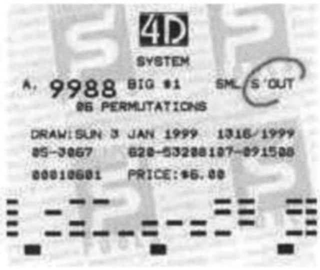

Fig. 1.2 *Output from the computer terminal.*

volume of Ordinary Bets in most cases. Nevertheless, the control of the System Bets can be easily incorporated into the control mechanism discussed in this paper.

2. Control Mechanism. It is clear that the prize structure is designed such that the expected return for the customer, for each dollar bet on BIG and SMALL, will be smaller than the bet amount received ($1). Hence, the "Accept All" policy is clearly the optimal policy if the objective is to maximize expected return. Unfortunately, the risk associated with this policy renders it unacceptable. As an illustration, we use a series of transaction data provided by an operator of this game to examine the two different accepted betting profiles obtained with the current policy and the Accept All policy.

We use the complete transaction data specifying

- the draw number,
- the time the betting slip is fed into the computer terminal at the outlet,
- the four-digit numbers marked in the betting slip,
- amount bet on each four-digit number,
- whether the bet is for the game BIG and/or the game SMALL, and
- whether the bet is accepted or rejected by the operator, using the current control policy.

The transaction data was collected over a two-week period. At the time of data collection, only two draws per week (Saturday and Sunday) were conducted. Hence, we have complete transactional data for bets received on four typical draws.

Note that the transaction data does not capture the fact that some of the bets received could have originated from earlier rejected bets (i.e., customers switching to a new set of numbers after their earlier bets had been rejected). We assume in our simulation that all the bets indicated in the transaction data (whether accepted or rejected by current policy) are genuine bets and will be accepted under the Accept All policy. We then simulate 500 draws of the 23 prizes, under the assumption that each digit appears with equal probability in all the 23 prizes; i.e., for each of the 500 draws, we select 23 winning four-digit numbers, uniformly over the set of 10,000 four-digit numbers.

The increase in bets accepted, using the Accept All policy, is around 0.025% of the total bets accepted. However, using this policy, it is possible for the system to *lose* up to 50% of the total bets accepted! This happens 2 to 3 times out of 500 simulations. For the system using current control policy, amazingly, the loss is *at most* 12% of the total bets received, based on results from 500 experiments. Thus the Accept All policy is potentially bad for the system, although it has a (slightly) higher expected return.

The challenge is thus to design a better control mechanism that accepts as many bets as possible, but at the same time, minimizes the risk exposure of the operator. To this end, we first distinguish between two classes of control mechanism: static versus dynamic.

In a dynamic control mechanism, the decision to accept fully, partially, or to completely reject a bet is decided based on the real-time betting information. For instance, starting with a predetermined control limit, one can dynamically raise the sales limit provided enough bets are received on all the numbers. One possible dynamic control mechanism is to mimic the rules used in the game "TETRIS." If there are at least one dollar bets on each of the 10,000 numbers, the control limit can be increased by a dollar without incurring additional risk. In the "TETRIS" analogy, the row is deleted so that the player has more room to maneuver the falling blocks. Though appealing, the dynamic control mechanism is unfortunately not considered useful to manage the four-digit game, mainly because of the following:

- The accept/reject decision has to be made almost instantly, given the volume of bets received per day. (The system reads close to several million betting slips per draw over 6 to 7 days for a draw.) A dynamic control mechanism may slow down the system considerably, if it is computationally very intensive.
- The dynamic control mechanism may classify a number as "hot" on one day, only to accept more bets on that number on the following day (based on changes of the betting profiles received). This is deemed undesirable by the operator as it encourages the players to return repeatedly to try their luck in betting on the "hot" number. This will result in unnecessary system delays and will almost surely lengthen the lines at the licensed outlets during the peak periods.
- The dynamic control mechanism cannot eliminate completely the need to rely on betting forecasts to make an accept/reject decision. If the decision at any time is made purely on the bets received so far, then the dynamic control mechanism will certainly reject a lot of bets that looked bad initially (increase in risk). However, together with the later bets, it may lead to a less risky betting profile for the system.

In view of the above, most operators have adopted a static control mechanism to manage the pool of money received for each draw. In a static control mechanism, a *forecast* of the total amount of bets on number i, i=0000 to 9999, is assumed to be given. The control mechanism, *prior* to the start of the draw, will determine a bound B_i and S_i for the maximum amount of BIG and SMALL bets, respectively, on number i to be accepted by the system. A number that hits the upper bound in the process is known as a "hot" number. Currently, most of the operators in the region are using a constant upper bound (denoted by U_B) for all bets on the game BIG, and another constant upper bound (U_S) for all bets on the game SMALL; i.e., $B_i = U_B$, $S_i = U_S$ for all i. We call this a *uniform control* policy. Note that when a bet arrives and pushes the total amount of bets received above the upper bound, the system will "truncate" the bet and accept only up to the sales limit imposed by the upper bound.

3. Analysis of Current Control Mechanism. To understand the rationale for the current control policy, we first provide a mathematical setting of the problem using the well-known mean-variance approach in investment analysis (see Steinbach [9] and the references therein for a comprehensive review). The mean-variance approach has also been used extensively in a variety of settings, such as data envelopment analysis (cf. Post [7]), robust optimization (cf. Ben-Tal and Nemirovski [4]), etc. Its application to this setting, however, appears to be new.

Let a'_{0000}, a'_{0001}, a'_{0002}, $\ldots$, a'_{9999} and b'_{0000}, b'_{0001}, b'_{0002}, $\ldots$, b'_{9999} be the *total* amount of dollars placed on the numbers 0000, 0001, 0002, ..., 9999 in the games BIG and SMALL, respectively.

For the rest of this section, we assume that $(a'_i, b'_i)_{i=0000}^{9999}$ are fixed. The payout from the betting system is denoted by p_i (resp., q_i) for the ith prize, $i = 1, 2, \ldots, 23$, in the BIG (resp., SMALL) category. Note that $q_4 = \cdots = q_{23} = 0$.

Let $\mathcal{C}$ denote the class of static policies. $a_k(C), b_k(C)$ are the amount of bets accepted by the policy C for the number k on BIG and SMALL, respectively, given that a'_k, b'_k are the total amount of bets that the system will receive for the number k on BIG and SMALL. For instance, for the uniform control policy with cut-off level of U_B and U_S for the games BIG and SMALL, respectively, $a_k(C) = \min(U_B, a'_k)$, $b_k(C) = \min(U_S, b'_k)$ for all k.

Let $X_1(C), X_2(C), \ldots, X_{23}(C)$ be the payout from the 23 prizes, under the control policy C in $\mathcal{C}$. For $i = 1, 2, 3, \ldots, 23$,

$$X_i(C) = p_i a_k(C) + q_i b_k(C) \text{ with probability } 1/10{,}000,$$
$$\text{for each } k \text{ in } 0000, 0001, \ldots, 9999.$$

Let K denote the amount of money the operator hopes to accept per draw. The optimal static control design problem can be formulated as the following:

$$\boxed{\begin{array}{ll} (SD)\ \text{min} & \sum_{k=1}^{23} Var(X_k(C)) \\ \text{subject to} & \sum_{i=0000}^{9999} (a_i(C) + b_i(C)) \geq K; \\ & 0 \leq a_i(C) \leq a'_i, i = 0000, \ldots, 9999; \\ & 0 \leq b_i(C) \leq b'_i, i = 0000, \ldots, 9999; \\ & a_i(C),\ b_i(C) \text{ integral for all } i; \end{array}}$$

Note that for all practical purposes, solving the above optimization problem to determine the optimal $a_i(C), b_i(C)$ (i.e., the optimal cut-off limit for each number i) is not likely to be useful, since the inputs to the model, a'_i, b'_i, are normally obtained from forecast and cannot be determined precisely prior to the implementation of the control policy. The value K, an input from the game operator, depends on its risk preferences and also partially on the forecast of the amount of bets for a particular draw. It is thus doubtful that the numerical solution to the above mean-variance model can be implemented directly in practice. Instead, we will derive structural results on the behavior of the optimal policy and show how these properties can be implemented in real time to obtain a practical control mechanism for the four-digit number game.

To illustrate the key idea, we first focus on the design of an optimal static control mechanism for a simplified version of the problem, by considering only the game on BIG. The problem (SD) reduces to

$$\boxed{\begin{array}{ll} \text{(SD-BIG)}\ \text{min} & \sum_{k=1}^{23} Var(X_k(C)) \\ \text{subject to} & \sum_{i=0000}^{9999} (a_i(C)) \geq K; \\ & a_i(C) \leq a'_i, i = 0000, \ldots, 9999; \\ & a_i(C) \text{ integral for all } i; \end{array}}$$

We prove next that the uniform control policy used by most operators is in fact close to optimal for this simplified problem. This provides a partial theoretical justification for the choice of the current control policy.

For a fixed betting profile, let $V(U)$ denote the variance of the payout function when the cut-off level is set at U. $a_i(C)$ in this case is denoted by $a_i(U) \equiv \min(U, a_i')$.

THEOREM 3.1. *The optimal control policy C^* for (SD-BIG) is close to that of the uniform control type; i.e., there exists a U such that*

$$|a_i(C^*) - a_i(U)| \leq 1.$$

Proof. Let $n = 10000$.

$$\sum_{k=1}^{23} Var(X_k(C)) = \sum_{k=1}^{23}\left(\sum_{i=1}^{n}(a_i(C)p_k)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C)p_k}{n}\right)^2\right).$$

Consider an optimal policy C^*. Let $A = \max_i(a_i(C^*))$. Without loss of generality, let $a_1(C^*) = A$. Suppose there exists a number, say k, such that $a_k(C^*) < a_k'$. Then by rejecting a dollar bet on 1, but accepting an additional dollar bet on k, the total amount of money received will be the same. If $A = a_k(C^*) + 1$, then by symmetry, this transformation does not change the variance of the total payout function. Hence we assume $A > a_k(C^*) + 1$. Let C' denote the corresponding policy. The variance changes to

$$\begin{aligned}
&\sum_{l=1}^{23}\left(\sum_{i=1}^{n}(a_i(C')p_l)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C')p_l}{n}\right)^2\right)\\
&= \sum_{l=1}^{23}\left(\sum_{i=1}^{n}(a_i(C^*)p_l)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C^*)p_l}{n}\right)^2\right)\\
&\quad + \sum_{l=1}^{23}\left(\frac{{p_l}^2}{n} + 2\frac{a_k(C^*){p_l}^2}{n} + \frac{{p_l}^2}{n} - 2\frac{a_1(C^*){p_l}^2}{n}\right)\\
&= \sum_{l=1}^{23}\left(\sum_{i=1}^{n}(a_i(C^*)p_l)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C^*)p_l}{n}\right)^2\right)\\
&\quad + \sum_{l=1}^{23}\frac{2{p_l}^2}{n}\Big(1 + a_k(C^*) - a_1(C^*)\Big)\\
&< \sum_{l=1}^{23}\left(\sum_{i=1}^{n}(a_i(C^*)p_l)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C^*)p_l}{n}\right)^2\right).
\end{aligned}$$

This is a contradiction. □

The betting profile received under C^* is thus close to that under a uniform control policy, with level set at $\max_i(a_i(C*))$, except possibly for some numbers where the level is set at $\max_i(a_i(C^*)) - 1$. For all practical purposes, it is therefore reasonable to focus on finding the best uniform control policy to manage the bets received for the system. This reduces a problem with possibly 10,000 variables to a single variable problem, since the question now boils down to finding the optimal cut-off level U. To this end, we first note that it is straightforward to prove the following proposition.

PROPOSITION 3.2. *$V(U) \leq V(U+1)$ for all $U > 0$.*

It is now easy to construct the optimal uniform control policy for (SD-BIG). Recall that $a_i(U) = \min(U, a_i')$. Let

$$U^* = \operatorname{argmin}_U \left(U : \sum_{i=0000}^{9999} a_i(U) \geq K \right).$$

It follows that the optimal uniform control policy for problem (SD-BIG) has limit set at U^*. All other uniform control policies have either $\sum_{i=0000}^{9999} a_i(U) < K$ (i.e., infeasible) or $V(U) \geq V(U^*)$ (i.e., inferior).

3.1. Determining U^* in Practice. The previous result showed that the optimal U^* can be determined if we know K and a_i' for all i. Since these values are usually not known prior to the draw, we need a different approach to determine U^* in practice. This problem is also compounded by the fact that the operator normally has no idea how to set the right K without knowing its risk preferences. One approach, as currently practiced by an operator, is to use past transaction data to build a forecasting model for the parameter a_i', and to use a simulation method based on the forecasting model to determine the expected revenue with different choices of U^*. The manager of the game will then decide how to set U^*, given the output from the simulation experiment. Managerial input and judgment is thus crucial for the determination of U^* using this approach.

In this section, we outline a different approach that can be used to determine U^* approximately without any managerial input.

Let U denote the limit imposed. U is determined in practice by examining historical data to forecast the amount of money that will be bet in each draw. How do we determine the optimal U? In particular, what is the effect on revenue collected by the game operator if the limit is raised from U to $U+1$?

To understand the importance of the selection of U, we note that a betting profile with an equal amount of dollar bets received on all 10,000 four-digit numbers presents the least risk to the game operator (in fact, zero risk!), since the payout does not depend on which 23 winning numbers are selected. On the other hand, a betting profile with all money stacked on a single four-digit number presents the greatest risk to the game operator, since although it gets to keep all the money most of the time, it will lose a lot of money occasionally, when the sole four-digit number is selected as a prize winning number. Given the histogram of the amount of bets the operator expects to receive, the cut-off limit U will essentially determine whether the betting profile accepted is closer to the first type (zero risk) or the second type (greatest risk). If U is set so low that all four-digit numbers will receive at least \$$U$ bets, then the betting profile accepted will be of the first type. However, the amount of money accepted by the game operator will be very low. On the other hand, if U is set too high, then the operator runs the risk of accepting a profile closer to the second type. Thus the determination of the appropriate U largely depends on the volume of bets the game operator expects to receive for a particular draw.

In the rest of this section, we provide a heuristic to determine the appropriate U. We analyze this question by computing the associated probabilities, with the aid of the symbolic computational program MAPLE.

We consider only the amount of money bet on BIG first. Suppose we raise the limit from U to $U+1$, and accept α additional dollars spread over α numbers; i.e., at level U, we assume that there are α hot numbers. In other words, we expect exactly α out of the 10,000 four-digit numbers will receive total bet dollars that exceed U. With

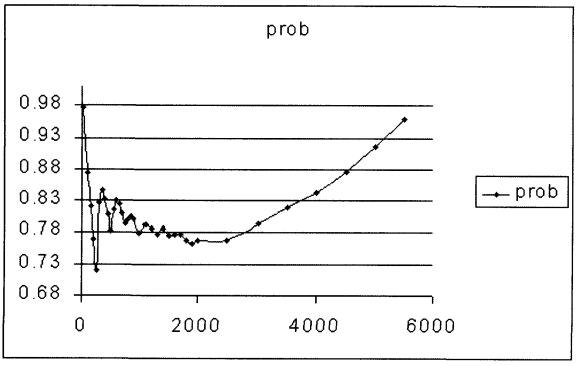

Fig. 3.1 *Probability of winning more by raising the limit by one dollar.*

probability $\alpha/10{,}000$, the kth prize will fall into one of these α numbers. Otherwise, with probability $1-\alpha/10{,}000$, the payout of the kth prize will be zero. Let $f(\alpha,x)$ denote the probability generating function of the total amount of prize money *awarded, with* α dollars accepted, \$1 each on α numbers:

$$\begin{aligned} f(\alpha,x) = \frac{1}{10000^{23}}&(\alpha x^{2000}+(10000-\alpha))\times(\alpha x^{1000}+(10000-\alpha)) \\ &\times(\alpha x^{500}+(10000-\alpha))\times(\alpha x^{250}+(10000-\alpha))^{10} \\ &\times(\alpha x^{65}+(10000-\alpha))^{10}. \end{aligned}$$

Note that the probability that the payout will be at most $\alpha-1$ (i.e., P(System makes more money by raising the limit)) is given by

$$f(\alpha,x)\mathrm{mod}(x^{\alpha})|_{x=1}.$$

The above is obtained by expanding the expression $f(\alpha,x)$ and removing all terms with x^k when $k\geq\alpha$. Note that the expression is independent of U.

Figure 3.1 plots the probability as a function of α. The x-axis denotes the number of hot numbers at level U. The y-axis denotes the corresponding probabilities of winning more by the modification. The computation suggests that for every limit U, by raising the limit by one dollar, *the probability that the system will make more money is at least* 0.7*!*

To determine the cut-off level, and to incorporate the desire to accept as many bets as possible, the cut-off level cannot be too low (so that the risk of losing is low), since it will cut off a lot of bets too. On the other hand, if the cut-off level is too high, then the risk exposure to the system will be unfavorable. We note from the probability computation that the chance of making more money is lowest in the region of $\alpha=200$ to 250. A natural policy to determine the cut-off level is thus

Cut-Off Level = the level that around 200–250 numbers will receive bets that exceed that level.

Intuitively speaking, the cut-off level (i.e., betting limit) should be set such that around 200–250 numbers will become "hot" numbers in the draw. By selecting this level, we will be able to cut off a portion of the betting profile where the risk of losing is among the highest. Interestingly, based on the real data provided, the current cut-off level a regional operator is using is within this range. In fact, for the data provided, there are 219 numbers that received bets greater than the current cut-off level in a particular draw. Again, the experience the operator gained over the years seems to have contributed to this sensible choice of the cut-off level.

In the next section, we will use this heuristic rule to determine the cut-off level of a new control mechanism and to compare the performance of the new control policy with the existing one.

4. An Improved Static Control Mechanism. In the previous section, we showed why the current control policy is optimal if the games BIG and SMALL are managed separately. However, since the outcomes for both games depend on the same draw, the current control policy is not able to exploit the information on the amount of bets received for BIG and SMALL jointly to improve its return and reduce its risk exposure. Intuitively, by pooling the risk from the games SMALL and BIG together, we should be able to accept more bets without any increase in the total risk exposure in the game. This is also highlighted in Figure 4.1, where some of the numbers were designated "hot" in the game BIG but were far away from their control limit in the game SMALL. By raising the limit for such numbers in the game BIG, the risk of a huge payout is mitigated by the fact that not many bets have been received for such numbers in the game SMALL.

In the rest of this section, we exploit this observation and show how the two games can be integrated, and we construct a new static control policy to improve on the performance of the current system. Recall that

$$\begin{aligned} &Var\left(\sum_{k=1}^{23} X_k(C)\right) \\ &= \sum_{k=1}^{23}\left(\sum_{i=1}^{n}(a_i(C)p_k + b_i(C)q_k)^2\frac{1}{n} - \left(\sum_{i=1}^{n}\frac{a_i(C)p_k + b_i(C)q_k}{n}\right)^2\right). \end{aligned}$$

Let C^* be an optimal control policy to the problem (SD). Let

$$K_{(BIG)} = \max_l\left(\sum_{k=1}^{23}\frac{2}{n}p_k^2 a_l(C^*) + \sum_{k=1}^{23}\frac{2}{n}p_k q_k b_l(C^*)\right),$$

say, $K_{(BIG)} = (\sum_{k=1}^{23}\frac{2}{n}p_k^2 a_j(C^*) + \sum_{k=1}^{23}\frac{2}{n}p_k q_k b_j(C^*))$ for some j.

Suppose $a_i(C^*) < a_i'$, $i \neq j$. Some bets on number i have been rejected under the policy. If we reduce the bets on BIG received for the number j ($j \neq i$) by a dollar, and increase the bets on BIG accepted for i by one dollar, the net change in the variance is

$$\frac{2\sum_{k=1}^{23}p_k^2}{n}(1 + a_i(C^*) - a_j(C^*)) + \frac{2\sum_{k=1}^{23}p_k q_k}{n}(b_i(C^*) - b_j(C^*)).$$

Suppose further that the policy C^* is optimal. Then we must have

$$\begin{aligned} &\sum_{k=1}^{23}\frac{2}{n}p_k^2(1 + a_i(C^*)) + \sum_{k=1}^{23}\frac{2}{n}p_k q_k(b_i(C^*)) \\ &\geq \sum_{k=1}^{23}\frac{2}{n}p_k^2(a_j(C^*)) + \sum_{k=1}^{23}\frac{2}{n}p_k q_k(b_j(C^*)). \end{aligned}$$

Otherwise, the above transformation will lead to a new policy with lower variance and an equal amount of bets received.

In the optimal control policy C^*, we have, if $a_i(C^*) < a'_i$,

$$\sum_{k=1}^{23} \frac{2}{n} p_k^2 (a_i(C^*)) + \sum_{k=1}^{23} \frac{2}{n} p_k q_k (b_i(C^*)) \leq K_{(BIG)},$$

but

$$\sum_{k=1}^{23} \frac{2}{n} p_k^2 (1 + a_i(C^*)) + \sum_{k=1}^{23} \frac{2}{n} p_k q_k (b_i(C^*)) \geq K_{(BIG)}.$$

This gives rise to a natural acceptance criterion for the bets on BIG:

> Let a_i, b_i be the current amount of bets accepted by BIG and SMALL on number i.
> Accept a $\$\Delta$ bet on BIG for number i only if
> $\sum_{k=1}^{23} \frac{2}{n} p_k^2 (a_i + \Delta) + \sum_{k=1}^{23} \frac{2}{n} p_k q_k (b_i) \leq K_{(BIG)}$.

To decide whether to accept the bets on SMALL for the number i, let

$$K_{(SMALL)} = \max_l \left(\sum_{k=1}^{23} \frac{2}{n} p_k q_k a_l(C^*) + \sum_{k=1}^{23} \frac{2}{n} q_k^2 b_l(C^*) \right).$$

We use a similar logic to construct the following acceptance region:

> Let a_i, b_i be the current amount of bets accepted by BIG and SMALL on number i.
> Accept a $\$\Delta$ bet on SMALL for number i if
> $\sum_{k=1}^{23} \frac{2}{n} q_k p_k (a_i) + \sum_{k=1}^{23} \frac{2}{n} q_k^2 (b_i + \Delta) \leq K_{(SMALL)}$.

As in the current system, the values of $K_{(BIG)}$ and $K_{(SMALL)}$ can be obtained by looking at the historical betting profiles of

$$\sum_{k=1}^{23} \frac{2}{n} p_k^2 (a'_i) + \sum_{k=1}^{23} \frac{2}{n} p_k q_k (b'_i)$$

and

$$\sum_{k=1}^{23} \frac{2}{n} q_k p_k (a'_i) + \sum_{k=1}^{23} \frac{2}{n} q_k^2 (b'_i).$$

Recall that a'_i and b'_i are the total amount of bets that the game operator expects to receive for the number i on the games BIG and SMALL, respectively. As usual, these numbers can be estimated by looking at past sales data.

In line with the observation we have concerning the current system, we will choose $K_{(BIG)}$ such that it is the level where around 200–250 of the numbers $(\sum_{k=1}^{23} \frac{2}{n} p_k^2 a'_i + \sum_{k=1}^{23} \frac{2}{n} p_k q_k b'_i)$ will surpass. Similarly, $K_{(SMALL)}$ is also taken to be the level where around 200–250 of the numbers $(\sum_{k=1}^{23} \frac{2}{n} q_k p_k a'_i + \sum_{k=1}^{23} \frac{2}{n} q_k^2 b'_i)$ will surpass.

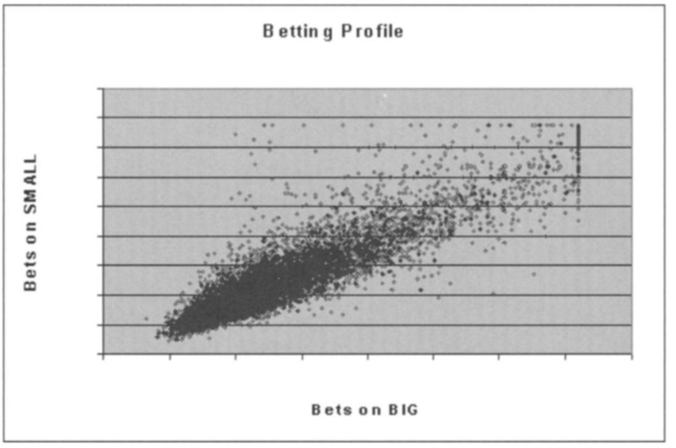

Fig. 4.1 *Betting profile under current policy.*

We can construct a control policy that strictly enforces *both* linear constraints at the same time when making an acceptance/rejection decision. We have also implemented a control mechanism with this feature. It beats the current system in terms of profits and is less risky. The increase in profits, however, is less than the one suggested above, as it is a more conservative approach. The difference between the two new policies is minimal as the gradients $\sum_{k=1}^{23} \frac{2}{n} p_k^2 / \sum_{k=1}^{23} p_k q_k$ and $\sum_{k=1}^{23} \frac{2}{n} q_k p_k / \sum_{k=1}^{23} q_k^2$ are relatively close to one another.

In the rest of this paper, we will only report the empirical results for the control policy that checks only one linear constraint while making an acceptance/rejection decision. Readers should also bear in mind that unlike the current system, we did not attempt to "truncate" the bets received while implementing the new control policy. Additional increments in profits could be expected with this modification to the new policy.

4.1. Empirical Results. We used the profiles for the bets received on a particular draw provided by a regional operator to construct the two bounds $K_{(SMALL)}$ and $K_{(BIG)}$. The numbers with betting levels that exceed these bounds are 202 and 201, respectively, in accordance with our choice of the cut-off level.

Graphically, the feasible region accepted by the new control policy is very different from that of the current system. For instance, the scatter plots of the feasible regions for bets on BIG and SMALL, on a particular draw, are shown in Figures 4.1 and 4.2 for betting profiles under the current and proposed policy, respectively.

Figure 4.1 shows that there are several "hot" numbers for the game SMALL that have received far fewer bets on the game BIG than the allowable limit. Instead of rejecting the bets on these numbers for SMALL, we will utilize the slack in the game for BIG to accept additional bets. This is the intuitive reason why the proposed mechanism is able to accept more bets than the current system.

We ran 500 simulations to determine the payouts based on the two betting profiles accepted using the current and proposed systems. The following tables summarize our main observations.

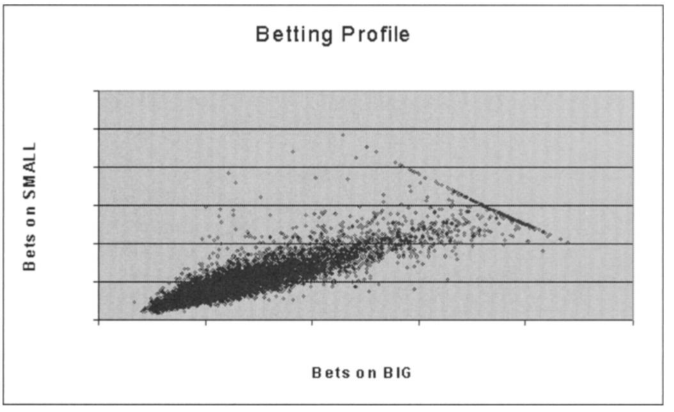

Fig. 4.2 *Betting profile under proposed policy.*

(1) *Which control policy accepts more money?*

	(Proposed - Current)/(Total rejected under current system)
Draw 1	4.34%
Draw 2	5.05%
Draw 3	3.99%
Draw 4	4.16%

In all four draws, the proposed policy accepts more money than the current policy. In fact, it accepts up to 4%–5% of the bets rejected under the current system in all four draws.

(2) *Which control policy makes more money?*

To address this question, we subtract from the amount of money accepted the payouts from the 500 simulations. The next two tables show the number of times each policy wins relative to the other, and the average improvement in additional profit using the proposed policy.

	Current policy	Proposed policy
Draw 1	74	426
Draw 2	84	416
Draw 3	75	425
Draw 4	88	412

Number of wins (out of 500)

Hence in 80% of all the experiments, the new policy makes more money than the current policy.

(3) *Which control policy is more risky?*

To address this question, we list the maximum loss from the two different policies, based on 500 simulations in the following table:

	Max loss in Current - Max loss in Proposed
Draw 1	\$41,834
Draw 2	\$172,921
Draw 3	−\$298,477
Draw 4	\$92,681

Difference in maximum loss from 500 simulations

Except for Draw 3, the maximum loss from the new policy is *smaller* than that from the current policy, even though the new policy accepts more bets!

5. Concluding Remarks. In this paper, we address risk management issues in a popular number game in Southeast Asia. In particular, we provide theoretical justification for the current system employed. By pooling the risks in the games BIG and SMALL together, we show how the current policy can be fine-tuned to improve its profits and to reduce its risk exposure. As a by-product of this analysis, we show that the nonlinear optimization problems (SD) and (SD-BIG) can be solved exactly or approximately with nice linear control policies.

The problem proposed in this paper can be studied in several other ways. Stochastic programming or robust programming techniques can be used to estimate the probability distribution of the payout function directly (see, for instance, the recent work of Ben-Tal and Nemirovski [4]). Also, it is debatable whether variance is the right way to measure the risk to the game operator, since it does not discriminate between upside risk and downside risk. A better way to measure risk in this game may be the "value-at-risk" approach (cf. Jorion [6]) or the more tractable "conditional value-at-risk" approach (cf. Rockafellar and Uryasev [8]). However, since our focus here is on finding structural properties of the optimal control policy, and not on solving the optimization model numerically, it is doubtful that these approaches will generate new and additional insights into the problem.

Another possible way to extend the work is to challenge the assumption that all digits in the game will be drawn with equal probability. Some statistical analysis and testing for uniformity in lottery hot numbers have been done before (cf. Johnson and Klotz [5]), although their issues and focus were somewhat different. As already evident in the data tracked by the operator in Malaysia (cf. [3]), this assumption may not be entirely true. A count of the number of times a digit is drawn in a winning number shows that certain digits have a higher chance of being drawn. Similarly, in Singapore, using information from the Singapore Pools web site, we have collected data for the winning numbers up to the draw on Feb. 6, 2000, over a period of about 13 years. A total of 32,830 occurrences was collected. The following table shows the number of times a digit was selected in the winning numbers:

Digit i	Observed frequency	Observed probability
0	13,137	0.1000
1	13,056	0.0994
2	13,202	0.1005
3	13,190	0.1004
4	13,401	0.1020
5	13,332	0.1015
6	12,880	0.0981
7	13,132	0.1000
8	13,122	0.0999
9	12,868	0.0980

A simple χ^2 test using the above data is able to reject the null hypothesis that the digits are selected uniformly, with a 95% confidence level. Hence the empirical data suggests that the digits are not selected uniformly.

Whether the above empirical observation can be exploited to design a better control mechanism remains a challenging problem.

Acknowledgments. We would like to thank the referees for commenting extensively and for pointing out certain errors and omissions in an earlier draft of this paper. Their comments helped us improve the presentation of this paper substantially. We would also like to thank Mr. Ting See Ming for help rendered in the course of this research.

REFERENCES

[1] *Florida Lottery games PLAY 4 web site*, http://www.flalottery.com.

[2] *Singapore Pools web site*, http://www.singaporepools.com.sg.

[3] *Magnum Four-Digit Malaysia web site*, http://www.twosteps.com/magnumfour-digit/.

[4] A. Ben-Tal and A. Nemirovski, *Robust solutions of linear programs*, Oper. Res. Lett., 25 (1999), pp. 1–13.

[5] R. Johnson and J. Klotz, *Estimating hot numbers and testing uniformity for the lottery*, J. Amer. Statist. Assoc., 88 (1993), pp. 662–668.

[6] P. Jorion, *Value at Risk*, McGraw-Hill, New York, 1997.

[7] T. Post, *Performance evaluation in stochastic environments using mean-variance data envelopment analysis*, Oper. Res., 49 (2001), pp. 281–292.

[8] R. T. Rockafellar and S. Uryasev, *Optimization of conditional value-at-risk*, J. Risk, 2 (2000), pp. 21–41.

[9] M. C. Steinbach, *Markowitz revisited: Mean-variance models in financial portfolio analysis*, SIAM Rev., 43 (2001), pp. 31–85.