ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Berninghaus, Siegfried

Working Paper

Activity dispersions and variable planning horizons

Diskussionsbeiträge - Serie A, No. 183

Provided in Cooperation with:
Department of Economics, University of Konstanz

Suggested Citation: Berninghaus, Siegfried (1984) : Activity dispersions and variable planning horizons, Diskussionsbeiträge - Serie A, No. 183, Universität Konstanz, Fakultät für Wirtschaftswissenschaften und Statistik, Konstanz

This Version is available at:
https://hdl.handle.net/10419/75066

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Universität Konstanz

Fakultät für Wirtschaftswissenschaften und Statistik

Siegfried Berninghaus

Activity Dispersions And Variable Planning Horizons

Diskussionsbeiträge

Postfach 5560
D-7750 Konstanz

Serie A — Nr. 183
Juni 1984

ACTIVITY DISPERSIONS AND VARIABLE PLANNING HORIZONS

Siegfried Berninghaus

Serie A - Nr. 183

Juni 1984

Serie A: Volkswirtschaftliche Beiträge

Serie B: Finanzwissenschaftliche Arbeitspapiere

Serie C: Betriebswirtschaftliche Beiträge

Introduction

It was M. Rothschild [7] who first demonstrated in his pioneering article the importance of the two-armed-bandit paradigm for applications in the "Economics of Incomplete Information".
In his model firms are setting prices as players who have to pull one arm of a two-armed-bandit. They learn by experience as "Bayesian statisticians". With the help of his model Rothschild is able to explain the existence of price-dispersions even in an economy with identical firms but with differing experience.
In the present paper we try to exploit a special case of the two-armed-bandit to give further arguments for the existence of dispersion of firm's activities: We consider the case of a two-armed-bandit where the success probability of one arm is known (see for example Bradt et.al. [3], Bellman [1], De Groot [4]). This seems to be a plausible assumption for decision making in situations where a decision maker has to choose between pursuing a (probabilistically) well known activity and experimenting with a new activity about whose success probability he has only an a priori estimation.
In contrast to the Rothschild model we assume that firms have finite planning horizons expressing their "degree of myopic behaviour". In a simple partial model we analyze the relation between the dispersion of planning horizons and the existence and persistence of acitivity dispersion. An economic interpretation of the abstract model yields conditions for the existence of quality-dispersion and price-quality-dispersions.
The present paper is organized as follows: In the first part an abstract model of two-armed-bandit with one arm known is developed. Explicit calculations of optimal strategies are made for an example of a priori distributions of a special type (see De Groot [4]). The simplification

is made mainly to depict clearly the underlying economic ideas. The economic implication of the abstract model is worked out in detail in Chapter II. In the last chapter it is demonstrated that a formal generalization of the model to arbitrary a priori distributions is straightforward.

I. The formal model

(a) At first the formal structure of the model is presented in the language of gambling with two one-armed-bandits (≡ "Two-armed-bandit"):
In its "wellknown" formulation (e.g. Bradt, et.al. [3], Robbins [6]) we have a slot-machine with an X_1-arm and an X_2-arm. When either arm is pulled the machine pays off one unit (of money for instance) with the success probabilities Π_1, resp. Π_2. These success probabilities are unknown to the player but he has an a priori estimation of (Π_1, Π_2) which is represented by a probability density function ξ.
One is allowed $(N-1)$ $(N < \infty)$ plays and a strategy is desired which will maximize the expected payoffs over N stages.
Therefore the following probability model is appropriate for a formal representation of the game:
Let the probability space of the N-stage game be denoted by $(\Omega, \mathcal{B}, P)$ where

- $\Omega = \Omega_1 \times \Omega_2$,
- $\Omega = S^N$, $S = \{o,1\}$, $i \in \{1,2\}$
- $\mathcal{B} = \mathcal{B}_1 \otimes \mathcal{B}_2$ (product σ-algebra), $\mathcal{B}_i = \bigotimes_{n=1}^{N} \mathcal{B}_{in}$,

  ($\mathcal{B}_{in}$ denotes the set of all subsets of S),
- $P = P_{\Pi_1} \otimes P_{\Pi_2}$ (product measure)[1], $P_{\Pi_i} = \bigotimes_{n=1}^{N} \Pi_i$.

Let the <u>winnings on each arm</u> be represented by two sequences of random variables $\{X_{1n}\}_{n=1}^{N}$, $\{X_{2n}\}_{n=1}^{N}$ where

$$X_{in} : \Omega \to S ,$$

$$X_{in}(\omega) := \tilde{X}_{in}(\omega_i) := \mathrm{proj}_n(\omega_i)^{2)}$$

The <u>sequence of information</u> of a player is described by the sequence of σ-algebras $\{\mathcal{F}_n\}_{n=1}^N$ where $\mathcal{F}_n \subset \mathcal{B}$ for all n ($\mathcal{F}_n$ is generated by the n-dimensional cylinder sets in $\mathcal{B}$). It is assumed that $\mathcal{F}_o := \{\phi, \Omega\}$ represents null-information before the game begins.
The <u>policy</u> (strategy) of a player is represented by a sequence $\{\sigma_n\}_{n=1}^N$ of $\mathcal{F}_{n-1}$-measurable functions

$$\sigma_n : \Omega \to \{1,2\}.$$

Consequently the choice of an arm $i (i \in \{1,2\})$ at stage $n \geq 2$ depends on the information available up to this stage. The <u>a posteriori probability</u> of (Π_1, Π_2) after n trials is given by a density function $\xi(x_{i_1,1}, \ldots, x_{i_n,n})$ where $(x_{i_1,1}, \ldots, x_{i_n,n})$ denotes a realized sequence of successes with ξ as prior density. By Bayes theorem one obtains

$$\xi(x_{i_1,1}, \ldots, x_{i_n,n}) = \frac{\Pi_1^{n_1}(1-\Pi_1)^{N_1-n_1}\,\Pi_2^{n_2}(1-\Pi_2)^{N_2-n_2}\,\xi(\Pi_1,\Pi_2)}{\int_{[o,1]\times[o,1]} \Pi_1^{n_1}(1-\Pi_1)^{N_1-n_1}\,\Pi_2^{n_2}(1-\Pi_2)^{N_2-n_2}\,\xi(\Pi_1,\Pi_2)\,d\Pi_1 d\Pi_2} \qquad (I.1)$$

N_i denotes the number of trials on arm i, n_i the number of successes.
The aim of the player is to find a strategy $\sigma = \{\sigma_i\}_{i=1}^N$ such that the objective function

$$\sum_{n=1}^{N} E(X_{\sigma_n,n}) \qquad (I.2)$$

is maximized.

(b) To derive the structure of the optimal policy it is often helpful to use dynamic programming techniques. For the two-armed-bandit problem the following <u>functional equation</u> describes the optimal choice of an arm at each stage. Let $V^m(\xi)$ denote the "worth" of the a priori density with m pulls remaining when an optimal strategy is followed, we have

$$- V^m(\xi) = \max_{i\in\{1,2\}} W_i^m(\xi) \quad , \quad m = 1,\dots,N-1 \qquad (I.2)$$

$$- W_i^m(\xi) = E_\xi(X_{i,N-m}) + E_\xi(V^{m-1}(\xi(X_{i,N-m}))) =$$

$$= E_\xi(\Pi_i) + E_\xi(\Pi_i)V^{m-1}(\xi_i(1)) +$$

$$+ (1-E_\xi(\Pi_i))V^{m-1}(\xi_i(0)),$$

$$- V^0(\xi) = 0$$

With $E_\xi(\Pi_i) = \int_0^1 \xi_i(\Pi_1,\Pi_2)\, d\,\Pi_1 d\,\Pi_2$, and $\xi_i(1)$ for instance denotes the a posteriori probability after a success on arm i ("$X_{i,N-m} = 1$"). Suppose the a posteriori probability at stage n can be described by ξ then an <u>optimal policy</u> chooses arm i iff

$$W_i^{N-n}(\xi) > W_j^{N-n}(\xi)$$

$i \neq j$, $i,j \in \{1,2\}$. In case of equality of $W_i^m(\cdot)$ and $W_j^m(\cdot)$ we adopt the convention that arm 2 is chosen. Let us define

$$\Delta_{N-n}(\xi) = W_1^{N-n}(\xi) - W_2^{N-n}(\xi)$$

Then arm 1 is chosen iff $\Delta_{N-m}(\xi) > 0$, arm 2 is chosen iff $\Delta_{N-m}(\xi) \leqq 0$.

(c) For the economic application given in the following chapters a special (simplified) form of the two-armed-bandit is needed: $\Pi_2(\in(0,1))$ is known to the player. For the derivation of the following results we will give a unified method of proof utilizing only Berry's recurrence relation (see Berry [2]). A well-known result concerning the structure of the optimal policy in this case is given by the following

<u>Lemma 1</u>: (Bradt et.al., [3], Yakowitz, [8]).
<u>The optimal policy σ^* for the above problem is characterized by the following property: Define a random variable n^* by</u>

$$n^* = \min \{n \mid \sigma_n = 2\}$$

<u>then</u> $\sigma^*_m = 2$ <u>for</u> $m > n^*$.

<u>Proof</u>: As Π_2 is known a priori information can be represented by a positive density function $\xi_1(\cdot)$ for arm 1 alone. The proof utilizes the following recurrence relation for $\Delta_n(\xi_1)$ which has been derived by D. Berry (in Berry, [2], p.876)

$$\Delta_n(\xi_1) = E_{\xi_1}(\Pi_1)\ \Delta^+_{n-1}(\xi_1(1)) + \tag{1}$$
$$+ (1-E_{\xi_1}(\Pi_1)\ \Delta^+_{n-1}(\xi_1(0)) + (\Delta^-_{n-1}(\xi_1)$$

where $\Delta^+_n = \max\{0,\Delta_n\}$, $\Delta^-_n = \min\{0,\Delta_n\}$.

Now let's assume there exists in contrast to the statement of the lemma an integer $K > 0$ such that

$$\Delta_{N-n^*-m}(\xi_1) \leqq 0 \text{ for all } m \geqq 0,\ m < K$$

and

$$\Delta_{N-n^*-K}(\xi_1) > 0.$$

As ξ_1 is not changed by observation we have

$$\Delta_{N-n^*-K+1}(\xi_1) = E_{\xi_1}(\Pi_1)\Delta^+_{N-n^*-K}(\xi_1(1)) + \tag{2}$$

$$+(1-E_{\xi_1}(\Pi_1))\Delta^+_{N-n^*-K}(\xi_1(0)) + \Delta^-_{N-n^*-K}(\xi_1) \leqq 0.$$

By definition of K and because of the implication

$$\Delta_{N-n^*-K}(\xi_1) > 0 \Rightarrow \Delta_{N-n^*-K}(\xi_1(1)) > 0$$

(See Berry [2], p. 889) it follows from (2) that

$$E_{\xi_1}(\Pi_1) = 0$$

a contradiction to the positivity of ξ_1.

For the following a further simplification of the two-armed-bandit problem is made:

- the a priori probability of Π_1 (Π_2 is known) is given by

$$\Pi_1 = \bar{\Pi}(>\Pi_2)^{3)} \quad \text{with probability } \xi \in (0,1)$$

$$\Pi_1 = 0 \quad \text{with probability } (1-\xi).$$

Clearly this assumption is restrictive. It is made in order to simplify the calculations. In the last chapter it is shown that a generalization does not change the direction of the results.

Under this assumption a characterization of the optimal policy is given by

<u>Lemma 2</u>: (See De Groot, [4])

<u>The optimal policy σ^* is characterized as follows:</u>

<u>There exists a</u> $r^* \in \{1,\ldots,N\}$ <u>such that</u>

$$\sigma_n = \begin{cases} 1 & n < r^* \\ 2 & n \geqq r^* \end{cases} \quad \underline{\text{if}}\ X_{1,n} = 0\ \underline{\text{for}}\ n < r^*$$

$\sigma_n^* = 1$ <u>for all</u> n <u>if</u> $X_{1,n} = 1$ <u>for at least one</u> $n' < r^*$.

<u>Proof</u>: The property of σ^* stated in the theorem follows at once from Lemma 1 and the following statements

a) suppose n' is the time point where the first success with arm 1 occurs, then $\Delta_n(1) > 0$ for all n with $1 \leqq n \leqq N-n'$ (note: $\xi_1(1) = 1$).

b) $\Delta_n(\xi_1) > \Delta_{n-1}(\xi_1(0))$ as far as $\Delta_n(\xi_1) > 0$.

ad a) From the recurrence relation (see Lemma 1) we have (by assumption)

$$\Delta_{N-n'+1}(\xi_1) = \xi\bar{\pi}\Delta^+_{N-n'}(1) + (1-\xi\bar{\pi})\Delta^+_{N-n'}(\xi_1(0)) + \tag{1}$$

$$+ \Delta^-_{N-n'}(\xi_1) > 0.$$

It has been proved by (Kelley, [5], p. 1058) that

$$\Delta_n(\xi_1) - \Delta_n(\xi_1') > 0 \quad \text{for } \xi_1' < \xi_1.$$

Therefore $\Delta_{N-n'}(1) \leqq 0$ is inconsistent with (1).
We can deduce $\Delta_{N-n'-n}(1) > 0$ by similar reasoning.

ad b) Suppose $\Delta_{n-1}(\xi_1(0)) \geqq \Delta_n(\xi_1) > 0$

Then we obtain the contradiction

$$\Delta_n(\xi_1) = \xi\bar{\pi}\,\Delta^+_{n-1}(1) + (1-\xi\bar{\pi})\Delta^+_{n-1}(\xi_1(0)) + \tag{1'}$$

$$+ \Delta^-_{n-1}(\xi_1) > \Delta_n(\xi_1)$$

as $\Delta^-_{n-1}(\xi_1) = \Delta^-_{n-1}(\xi_1(0)) = 0$

because of (2), and $\Delta_{n-1}(1) > \Delta_{n-1}(\xi_1(0))$.

Because of Lemma 2 the computation of the optimal strategy is simple: Before the game begins an integer (r^*-1) is chosen which represents the last possible instant of pulling arm 1 if one begins with arm 1 and experience with this arm is bad. Therefore it remains to show how the optimal r^* is computed:

Lemma 3: Given a strategy σ characterized by an integer r (see Lemma 2).

Define
$$M(1,N) = N \cdot \Pi_2$$
$$M(r,N) = \sum_{j=1}^{r-1} E\left(\sum_{n=1}^{N} X_{\sigma_n,n} | A_j\right) P(A_j) + E\left(\sum_{n=1}^{N} X_{\sigma_n,n} | B_{r-1}\right) P(B_{r-1})$$

for $r \in \{2,3,\ldots,N+1\}$

where
$$A_j = \bigcap_{n=1}^{j-1} \{X_{1,n} = 0\} \cap \{X_{1,j} = 1\}$$
$$B_j = \bigcap_{n=1}^{j} \{X_{1,n} = 0\}$$

then the optimal strategy σ^* is characterized by r^* which is defined as

$$r^* = \min \{r \mid M(r+1,N) - M(r,N) \leq 0\}.$$

Proof: Obviously one has to consider as candidates for optimal strategies only N strategies σ^r each characterized by an integer r. Then we have

$$E\left(\sum_{n=1}^{N} X_{\sigma_n^r,n}\right) = M(r,N),$$

as one can show by a straight forward generalization of De Groot's proof (De Groot, [4], p. 397/398)

$$M(r+1,N)-M(r,N) = \xi(1-\bar{\Pi})^{r-1}(\bar{\Pi}-\Pi_2)[(N-r+1)\bar{\Pi}+ \tag{1}$$

$$+1-\bar{\Pi}]-\Pi_2(1-\xi)$$

which is for $\bar{\Pi} > \Pi_2$ a decreasing function of r. Consequently the r which maximizes $E\left(\sum_{n=1}^{N} X_{\sigma_n^r,n}\right)$ is given by r* defined above.

In the following Corollary two conditions are given which characterize two extreme strategies: to play the same arm at every stage regardless of the experience gained during the game.

Corollary:

A player will always choose arm 2 iff

$$\xi \leqq 1/\{(\frac{\bar{\Pi}-\Pi_2}{\Pi_2})[\bar{\Pi}(N-1)+1]+1\} \tag{I.3}$$

a player will always choose arm 1 if

$$\xi > 1/\{(\frac{\bar{\Pi}-\Pi_2}{\Pi_2})(1-\bar{\Pi})^{N-1}+1\} \tag{I.4}$$

Proof: a) Condition (I.3) is easiliy derived by taking r = 1 in (1) (Lemma 3) and solving for ξ .
If the condition is violated at least one pull with arm 1 is made.
b) A sufficient condition for always pulling arm 1 is

$$M(r+1,N)-M(r,N) > 0 \tag{2}$$

for all $r \in \{1,\ldots,N\}$. As this difference is decreasing in r (see Lemma 3) (2) is equivalent to

$$M(N+1,N) - M(N,N) > 0. \qquad (2')$$

By solving (2') for ξ one obtains (I.4).
Even if this condition is violated it is still possible to pull arm 1 at every stage.

From an economic point of view it is interesting to note that the validity of the conditions (I.3) and (I.4) depends on the number of stages N and ξ. If one defines <u>myopic behaviour</u> of an economic agent by setting a "short" planning period for economic activities then it is implied by the Corollary that keeping ξ constant the choice of an arm (that is a particular economic activity) is determined by the degree of myopic behaviour of the agent. This remark will be important for the following economic applications.

II. Applications in Economics

(a) We think of the following problem which economic agents often have to solve: In an "uncertain environment" the success of a particular economic activity is known in a probabilistic sense. But for the success probability of a new activity only an a priori estimation is available. If there is a positive (a priori) probability that the expected success with the new activity is higher than the well known expected success with the old activity it may be profitable for the economic agent to change his behaviour.
We simplify this problem by assuming that the a priori estimation has the following simple structure: The expected success with the new activity will be either higher than that with the old one or will be zero. As

mentioned earlier this assumption is made mainly for keeping the mathematics as simple as possible. Beyond this technical point it is from an economic point of view not unreasonable to expect some economic agents to have a simple (two-valued) a priori estimation of a "strong"-form: Either one regards the new activity as perhaps much better resp. much worse than the old one or one does not anticipate any change at all. As an approximating value for the a priori estimation of the bad result of the experiment we take zero.

(b) To give concrete examples of the problem mentioned above we consider three related choice-problems under uncertainty:
Let $(P-C_i)$ $(i \in \{1,2\})$ denote the profit per sold unit produced by "activity" (i.e. production process) i, where "i = 1" denotes the new, "i = 2" denotes the old activity. Let Π_2 denote the (well-known) success probability of activity 2. Π_i will be interpreted as the probability that a customer will buy one unit of the product produced by activity i.
And the producer is confronted with the following simplified purchase process (See for example M. Rothschild, [ ]): In each period one customer enters his store and either buys one unit (with probability Π_i) or leaves the store without buying (with probability $(1 - \Pi_i)$).
To be more concrete we think of the following interpretations of an "activity":

- Suppose $C_1 < C_2$ and the new production process is distinguished from the old one by producing items with low quality. For instance the break down probability of the product may rise. Under these circumstances it is reasonable to assume that $\bar{\Pi} < \Pi_2$. Then a change of the production process would be considered only if

- 4. JUNI 1985 Weltwirtschaft Kiel

$$\bar{\Pi}(p - C_1) > \Pi_2(p - C_2). \qquad \text{(II.1)}$$

- Suppose the new activity is the production of a "new" product which may be made possible by technological improvements ("product innovation"). In this case it will also be reasonable to assume $C_1 < C_2$ and $\bar{\Pi} < \Pi_2$ at least for the first few periods. Then product innovation may be profitable if (II.1) is fulfilled.
- A more realistic variant of the latter interpretation is the following: Suppose the producer tries to compensate the expected reduction of the purchase probability to $\bar{\Pi}$ by a price reduction to p'. Then product innovation will be considered if

$$\bar{\Pi}(p' - C_1) > \Pi_2(p - C_2). \qquad \text{(II.1')}$$

(c) Clearly an appropriate formal model for the economic applications described above is the framework of the two-armed-bandit developed in chapter I. One has to take

$$x_{i,n} = \begin{cases} (p - C_i), \\ 0 \quad , \end{cases}$$

and together with the assumptions about the customer's purchasing process the problem of the optimal choice of a new activity is equivalent to the problem of a player pulling the arms of a two-armed-bandit where the success probability of one arm is known to him. Consequently the results of chapter I are applicable almost without modifications.

Let us denote the profit per sold unit on activity i by G_i. Then condition (I.3) can be transformed by straight forward manipulations into

$$\xi \leqq 1/\{\left(\frac{G_1\bar{\Pi} - G_2\Pi_2}{G_2\Pi_2}\right)[(N - 1)\ \bar{\Pi} + 1] + 1\}. \qquad \text{(II.2)}$$

Regarding all terms in (II.2) exept N as constant we can deduce the following interesting

Proposition 1: Suppose assumptions (II.1), (II.1') are valid. Then there exists a real number $N^* \geqq 1$ such that agents with $N < N^*$ stay with their old activity while agents with $N > N^*$ try out the new one for at least one period.

Proof: Let us denote the right hand side of (II.2) by F(N). As is easiliy seen F(N) is strictly decreasing in N. Defining N^* by $F(N^*) = \xi$ if $\xi \leqq F(1)$ and $N^* = 1$ otherwise the assertion follows.

As the set of agents with planning period $N < 1$ is empty the case $N^* = 1$ in Proposition 1 must be read as follows: All agents try out the new activity at least for one period.

The next question to be answered is the persistence of the period of experimentation.

For that condition (I.4) is transformed into

$$\xi > 1/\{\left(\frac{G_1\bar{\Pi} - G_2\Pi_2}{G_2\Pi_2}\right)(1 - \bar{\Pi})^{N-1} + 1\}. \qquad \text{(II.3)}$$

Keeping all variables except N constant we obtain

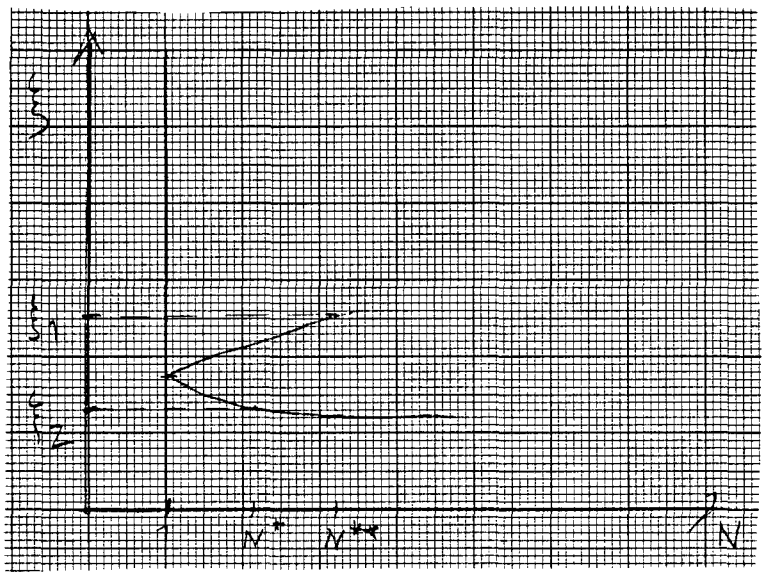
Proposition 2: Suppose (II.3) is fulfilled for at least one $N > 1$ and assumption (II.1) and (II.1') are valid then there exist a real number $N^{**} \geqq 1$ such that all agents with $N < N^{**}$ try out activity 1 for all periods.

Proof: Obvious.

The asymmetry in the assertions of proposition 1 and 2 stems from the fact that condition (II.2) is necessary

and sufficient while (II.3) is only sufficient (See Chapter 1, Corollary).
The following figure illustrates (for two values of ξ) the assertions expressed by the propositions

Fig. 1



(d) Now let us give a brief sketch of a "partial-partial" model of firm behaviour: Consider an economy with a finite number of firms which decide with the help of the two-armed-bandit paradigm which activity to pursue inasmuch as a new activity is available. The firms are assumed to be identical with respect to their a priori estimation ξ but differ from each other in the length of the planning horizon N. The latter could be motivated for instance by the assumption that firms use different discount rates.

Let us suppose for simplicity that before a fixed moment t* all firms pursue the same activity which they have pursued for a long time. At t* a new activity characterized by a possible expected return $G_1\bar{\Pi}$ (for instance a new technology) is available and all agents decide according to the calculations described above.

After $(t^* + N)$ a firm with planning period $(N - 1)$ stays with the activity it has selected at $t = (t^* + N)$[4].
The point in question now is the existence of activity dispersion at $t \geq t^*$.
The following proposition summarizes the main result of this paper that a dispersion of activities may be generated exclusively by an appropriate dispersion of myopic behaviour in the economy. To be more precise let us call a distribution of agents "sufficiently dispersed" if there exist agents with planning stages $N < N^*, N^{**}$ and $N > N^*, N^{**}$. That implies $N^*, N^{**} < \bar{N}$ (= longest planning period in the economy).
Finally let us assume that there exists a positive ("objective") purchase-probability $p \in (0,1)$ for the new activity. Consequently there exists a positive probability ($= (1 - p)^r$) that an "experimenting" agent with experience period r (for the definition see lemma 2) will switch over to the new activity.

Proposition 3:

a) Suppose $\xi < 1/\left(\frac{G_1\bar{\Pi} - G_2\Pi_2}{G_2\Pi_2} + 1 \right)$,

then there exists a non-degenerate dispersion of activities at t^* if the agents are "sufficiently dispersed". With positive probability a dispersion may be maintained for all $t > t^*$.

b) Suppose $\xi > 1/\left(\frac{G_1\bar{\Pi} - G_2\Pi_2}{G_2\Pi_2} + 1 \right)$,

then there exists a degenerate dispersion (concentrated at the new activity) at t^*. For all $t > t^*$ a nondegenerate dispersion may be maintained if the agents are "sufficiently dispersed".

In terms of our examples explained at the beginning of this chapter the latter proposition gives sufficient conditions for existence and persistence of quality and price dispersions generated exclusively by the agents myopic behaviour in face of uncertainty.

III. A generalization of the model

(a) As our knowledge about a priori probability distributions ξ is restricted a generalization of the results to arbitrary a priori distributions is required. Naturally the particular distribution ξ_1 (concentrated at $\Pi_1 = \bar{\Pi}$, $\Pi_1 = 0$) was not intended to represent a realistic estimation of firm's a priori beliefs. This distribution was introduced mainly for illustrative purposes. As sequential strategies in general are very difficult to describe we needed a simplification to illustrate the point relevant for economic applications.

On the other hand this distribution ξ_1 can be regarded as representative for "drastic" two-point a priori beliefs about the success probability of the "unknown arm" of a two-armed bandit ("either the new arm is better than the "old" one or its success probability is neglegible"). As it is easily understood the restriction to ξ_1 simplifies the problem considerably : A success on arm 1 (i.e. "$X_{1,n} = 1$") is inconsistent with the hypothesis $\Pi_1 = 0$.

(b) In the following the main results are derived which are needed to generalize the conclusions of the last chapter. Let ξ_1 denote a continous positive a priori density function of Π_1. To be more concrete let us assume henceforth that

$$\xi_1(\Pi_1) = \frac{\Pi_1^{\,r}(1-\Pi_1)^{r'}d\mu}{\int_0^1 \Pi_1(1-\Pi_1)^{r'}d\mu} \qquad \text{(III.1)}$$

where μ denotes a positive measure concentrated at the interior of $[0,1]$. Consequently ξ_1 can be characterized by two real numbers $r,\ r' \in \{(r,r') \in R^2) \mid \int_0^1 \Pi_1^r(1-\Pi_1)^{r'}d\mu < \infty\}$.

The optimal policy can now be characterized by a real number $Q(n,\xi_1)$ depending on the number n of trials remaining and ξ_1 at that time.

<u>Lemma 4: There exists a function $Q(n,\xi_1)$ such that arm 2 is required if and only if</u>

$$\Pi_2 \geqq Q(n,\xi_1).$$

<u>Proof</u>: By (I.2) arm 2 is required iff

$$E_{\xi_1}(\Pi_1) + E_{\xi_1}(\Pi_1)V^{n-1}(\xi_1(1)) + (1-E_{\xi_1}(\Pi_1))V^{n-1}(\xi_1(0)) \leqq \qquad (1)$$
$$\leqq W_2^n(\xi_1) = n\Pi_2$$

where the last equality follows from Lemma 1. Let us devide the left hand side of (1) by n and denote it by $F_n(\xi_1,\Pi_2)$. One can prove by induction that $F_n(.,.)$ is non-decreasing convex and continous in Π_2. Further $F_n(\xi_1,0) = E_{\xi_1}(\Pi_1) > 0$,

$$F_n(\xi_1,1) = (E_{\xi_1}(\Pi_1)+(n-1))/n = (1-(1-E_{\xi_1}(\Pi_1)))/n < 1.$$

Take for $Q(n,\xi_1)$ the (unique) fixed point of $F_n(\xi_1,\cdot)$.

The following property which is important for our applications has been proven by Bradt, Karlin, Johnson [3], p. 1069

$$Q(n,\xi_1) \geqq Q(n-1,\xi_1) \text{ for } n \geqq 2, \text{ and all } \xi_1. \qquad \text{(III.2)}$$

Now let us suppose that $\Pi_2 < Q(1,\xi_1)$. Then because of (III.2) arm 2 is never pulled at the first stage. In case $\Pi_2 \geqq Q(1,\xi_1)$ it is not clear whether there exists at least one planning horizon N such that $\Pi_2 < Q(N,\xi_1)$. In view of the economic applications it would be very undesirable if for all ξ_1 either $Q(N,\xi_1) \leqq \Pi_2$ or $Q(N,\xi_1) > \Pi_2$ for all N. In the following Lemma it is demonstrated that these inequalities will not be valid for all ξ_1 resp. N.

<u>Lemma 5</u>: <u>There always exists a ξ_1 and an integer $N \geqq 2$ such that</u>

$$Q(n,\xi_1) \leqq \Pi_2 \quad \underline{\text{for } n < N}, \qquad \text{(III.3)}$$

$$\underline{\text{and}} \quad Q(N,\xi_1) > \Pi_2.$$

<u>Proof</u>: We write Berry's recurrence relation in the following form (for an arbitrary $N < \infty, N \geqq 2$)

$$\begin{aligned} \Delta_N(r,r') = {} & E_{(r,r')}(\Pi_1)\Delta^+_{N-1}(r+1,r') + \\ & + (1-E_{(r,r')}(\Pi_1))\Delta^+_{N-1}(r,r'+1) + \\ & + \Delta^-_{N-1}(r,r'). \end{aligned} \qquad (1)$$

As Δ_{n-1} is continous in (r,r') there exists for any $\Pi_2 \in (0,1)$ a pair $(\bar{r},\bar{r}')$ such that

$$\Delta_{N-1}(\bar{r},\bar{r}') = 0.$$

As has been proven by Berry [2], p. 888, (5,17), we have $\Delta_{N-1}(\bar{r}+1,\bar{r}') > 0$ implying $\Delta_N(\bar{r},\bar{r}') > 0$ and $Q(n,\xi_1) > \Pi_2$ for all $n \geqq N$. From the result of lemma 1 we have $Q(n,\xi_1) \leqq \Pi_2$ for all $n < N$.

As is demonstrated in the proof: given Π_2 we can for any N find an appropriate ξ_1 which satisfies (III.3).

Before summarizing the results so far we will finally give a condition analogous to (I.4) for pulling arm 1 at all stages.

<u>Lemma 6</u>: (See Bradt, et. al. [3], p. 1074)
<u>Given N, ξ_1 (characterized by (r,r')) then arm 2 should never be used at all if</u>

$$\Pi_2 \geqq E(r,r'+N-1)(\Pi_1) \qquad \text{(III.4)}$$

<u>Proof</u>: Suppose arm 2 should be used if m pulls remain. By Lemma 1 arm 2 should be used at the last stage too implying

$$\Pi_2 \geqq E_{(r,r'+N-m)}(\Pi_1)$$

By the Cauchy-Schwartz inequality we have

$$E_{(r,r'+N-1)}(\Pi_1) = E_{(r,r'+N-m)}(\Pi_1)$$

in contrast to (III.4).

Summarizing the results relevant for economic applications we can conclude from the foregoing lemmas that with respect to the existence of activity dispersions the same propositions can be proven as in chapter II: Given the success probability Π_2 of the well known activity and ξ_1

- there exists a dispersion of activities at t* iff ξ_1 belongs to the type of distributions described by (III.3) and if the degree of myopic behaviour is "sufficient" in the sense that there are agents with planning periods larger than that required by ξ_1.
- there is a positive probability for a dispersion of activities at least at stage $\bar{N}$ if ξ_1 is such that (III.4) is not satisfied by all agent's planning horizons N, that means if the agents' myopic behaviour is "sufficiently dispersed". For the agents characterized by large N might turn to arm 2 again if experience with the new activity is disappointing.

For the results concerning the persistence of dispersions we can refer to propositions 1 and 2 (last chapter) whose generalization to arbitrary density functions is immediate.

FOOTNOTES

1. It is implicit in this assumption that the two arms are "independent".

2. $\text{proj}_n(\omega_i)$ denotes the projection of ω_i (= $(\omega_{i1},\ldots,\omega_{iN})$) on its n-th coordinate (ω_{in}).

3. If $\bar{\Pi} < \Pi_2$ it would trivially be optimal to pull arm two at all stages.

4. Obviously to be a reasonable assumption this presupposes a large number of small firms without (any significant) experience with the new activity.

References

[1] Bellman, R., A Problem in the Sequential Design of Experiments, Sankhya 16 (1956), 221 - 229.

[2] Berry, D., A Bernoulli Two-Armed-Bandit, Ann.Math.Stat., vol. 43 (1972), 871 - 897.

[3] Bradt, R./ Johnson, S./ Karlin, S., On Sequential Designs for Maximizing the Sum of n Observations, Ann.Math.Stat., vol. 27 (1956), 1068 - 1074.

[4] De Groot, M., Optimal Statistical Decisions, Mc. Graw Hill, New York (1970).

[5] Kelley T., A Note on the Bernoulli Two-Armed-Bandit Problem, Ann.Stat., vol. 2 (1974), 1056 - 1062.

[6] Robbins, H., Some Aspects of the Sequential Design of Experiments, Bull.Amer.Math.Soc., 68 (1952), 527 - 535.

[7] Rothschild, M., A Two-Armed-Bandit Theory of Market Pricing, J.E.T., vol. 9 (1974), 185 - 202.

[8] Yakowitz, S., Mathematics of Adaptive Control Processes, American Elsevier Publ. Comp., New York (1970).