

SEARCHING FOR THE LOWEST PRICE WHEN
THE DISTRIBUTION OF PRICES IS UNKNOWN

Michael Rothschild

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 150
August 1973

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

SEARCHING FOR THE LOWEST PRICE WHEN
THE DISTRIBUTION OF PRICES IS UNKNOWN

by

Michael Rothschild*

Princeton University

July 1973

* A preliminary version of this paper was presented at the NBER conference on stochastic control at the University of Chicago, June 1973. I have benefited from discussions with David Blackwell, Gregory Chow, Stanley Fischer, and Robert E. Lucas. This research was supported by the Industrial Relations Section of Princeton University, the Mathematical Social Science Board, and the National Science Foundation.

I. Introduction

Stories about search occupy a central position in the "new micro-economics of inflation and employment." This accounts for much of the recent work by economists^{1/} analyzing variants of the following problem: A man is considering purchase of some good which is sold at different stores at different prices. He can elicit price quotations from the various sellers by paying a fee.^{2/} What search strategy will the man follow?

Economists are interested in the rules which searchers follow because these rules determine the demand functions that sellers in such markets face, and thus, in part, the nature of the markets themselves. Although the only way to settle the question of what rules searchers follow is by observation, very little empirical work has been done on this problem.^{3/} Instead, researchers have proposed search rules that reasonable consumers might follow and examined their properties. Stigler, who is largely responsible for introducing this topic into economic theory, suggested that the individual should visit n stores, obtain price quotations from each one and then buy from the lowest price store. The expected price paid will be $M_n = \int_0^{\infty} (1 - F(p))^n dp$, where $F(p)$ is the distribution of prices. If such a procedure is followed, the only decision variable is n , the number of stores visited. Stigler pointed out that the familiar marginal calculations of micro-economic theory will suffice to determine n . Clearly M_n is a decreasing function of n , while the expected gain from searching, $G_n = M_{n-1} - M_n = \int_0^{\infty} (1 - F(p))^{n-1} \cdot F(p) dx$, also decreases with n . Therefore n should be chosen so that $G_n \geq c > G_{n+1}$.

This rule has several interesting properties:

- 1) If all its potential customers follow this rule then a firm faces

a well-behaved demand function; expected sales are a non-increasing function of the price it charges.

2) Customers' search behavior is a function of the cost of search c and the distribution of prices $F(\cdot)$. Thus, it is possible to do comparative statics so as to examine the consequences changes in $F(\cdot)$ and c . The most important results of these exercises are:

3) If costs of search increase, the amount of search decreases.

4) As prices become more dispersed, expected total costs decrease.

This follows from the fact that for all n , M_n decreases as the distribution becomes more variable.^{4/} Other things equal, customers prefer to draw from riskier distributions. This preference for risk or uncertainty on the part of those normally deemed risk averse, can explain apparently odd or perverse phenomena. For an application to the theory of migration see David (1973).

Interesting as these results are, they depend on people following a particular search rule -- and not a particularly attractive one at that. Although fixed sample size rules have a certain intuitive appeal, they are not the best search procedures and are in some circumstances simply silly. A person who rigidly follows a fixed sample size rule will, even if he gets a price quotation less than the cost of search, keep on sampling until his quota of price quotations is fulfilled. It is thus comforting to know that the optimal search rule has all the attractive properties of the fixed sample size rule. The optimal rule is sequential (after receiving each price quotation the searcher decides whether to continue searching or to accept the quoted price) and is characterized by a reservation price; there is a price R such that the searcher will accept any price less than or equal to R while he will reject a price higher than R . Once again,

familiar economic reasoning suffices to determine R . If the lowest price the customer has received to date is S then the expected gain from searching once more is

$$g(S) = \int_0^S (S - p) dF(p) = \int_0^S F(p) dp .$$

The optimal rule is to search whenever the cost of an additional search is less than the expected gain from that search so that R must satisfy

$$g(R) = c. \tag{1}$$

This reservation price rule has the four properties listed above. That properties 1) and 2) hold is obvious. Increasing c increases the reservation price and thus lowers the intensity of search so that 3) holds. Kohn and Shavell (1973) have shown that increasing dispersion lowers the reservation price and thus total expected costs^{5/} so that 4) holds for this rule as well. They do this by showing that if $F(p, t)$ is, as in fn. 4 above, a family of price distributions which becomes more dispersed as t increases, then

$$g(S, t) = \int_0^S f(p, t) dp$$

is an increasing function t . Therefore if

$$c = g(R, t) = g(R', t')$$

with $t' > t$, then $R' < R$. Thus the reservation price decreases as price variability increases or,

5) Increased price dispersion increases the intensity of search.

This last result -- which does not hold in general for fixed sample size rules^{6/} -- is a kind of stability property which could be used in a complete

model (one which explained price distributions exogenously) to show the existence of an equilibrium distribution of prices. It should turn out in most sensible models that increased search activity will decrease price dispersion.^{7/}

These results depend on the assumption that the searcher behaves as if he knows the distribution of prices. In any economic context, this is a very bad assumption. Little is known about the nature of price distributions and it seems absurd to suppose that consumers know them with any degree of accuracy.^{8/} Since the major reason for believing that searchers follow optimal sequential rules is that they are optimal, it is important that their cost minimizing properties not depend crucially on their being based on correct knowledge of the price distribution. This is unfortunately not true. Gastwirth (1971) explored the robustness of optimal reservation price rules. He found that modest specification errors could lead to dramatic increases in the expected number of searches and in the expected cost of buying. For example, someone who chose a reservation price on the assumption that prices were distributed uniformly on the unit interval ($F(x) = x$) when they really were distributed according to a right triangular distribution on the same interval ($F(x) = x^2$), would on the average incur roughly twice the total costs and search five times as much as he would if he were correctly informed. If the results 1) to 5) are to be salvaged, the problem of what the searcher should do if he does not know the price distribution must be attacked. This can be done in two ways: by exploring the properties of reasonable rules of thumb or by characterizing optimal rules. Telser (1973) took the first approach. He calculated (using Monte Carlo techniques) the expected costs of various search rules -- which had

properties 1) through 4) -- against several differently shaped price distributions and compared these costs to those which would ensue if the searcher followed the naive rule of taking the first price offered to him.

This paper takes the second tack. Optimal search rules from unknown distributions are derived and characterized.^{9/} The results of this exercise are as follows: In section II the problem of a man who knows prices belong to some finite set, but does not know how they are distributed is formally described. Section III discusses an important example -- the case where the prior distribution is a Dirichlet. In the next section dynamic programming is used to derive the optimal strategy. It is possible to parameterize the problem so that the optimal valuation functions are continuous -- a fact which is exploited in section V where it is shown that if a person follows the optimal strategy, search terminates after a finite number of searches. This is used to prove that property 3) holds for search from unknown distributions, that is that search decreases as its cost increases. In section VI it is shown if prior beliefs are Dirichlet the optimal search rule has a reservation price property. Searchers will accept a price if and only if it is less than some particular price p_R . The reservation price is a function of the searcher's beliefs; it changes as his beliefs are revised in the light of experience. Section VII is devoted to proving an analogue of properties 4) and 5) for the Dirichlet case.

The most important results of this paper are that for the example of section III, optimal search rules from unknown distributions have the same qualitative properties as optimal rules from known distributions. Since it is easy to construct examples for which this is not true, it is natural

to ask how general these results are. The final section considers this question. I believe them to be more general than it may appear, as they do not appear to depend on prices being confined to a finite set. However, they are still quite special as the proofs depend on the process of revising beliefs to accommodate new information having a particular -- and not terribly natural -- local property. Whether similar results hold if this assumption is abandoned is an open and difficult question.

II. Preliminaries

Consider a man trying to buy a good at the lowest total expected cost. At the beginning of each period this man pays c and receives a price quotation, which upon receiving he must decide whether to accept or to pay (at the beginning of next period) c again and receive another price quotation. Total costs include the price actually paid as well as the costs of search.^{10/} I assume that the man does not have the privilege of recall; that is, that an offer once spurned cannot be taken up again. As is well known, this assumption is innocent -- it has no effect on the optimal strategy -- when the price distribution is assumed known. It is not in the present case.

For simplicity, suppose that there are only a finite number of prices p_1, p_2, \dots, p_n and label them in ascending order so that

$$p_i < p_{i+1} \quad . \quad (2)$$

The probability distribution of prices is a multinomial distribution. It is completely characterized by the vector Π whose i^{th} element π_i is the probability that the i^{th} price is chosen. Since Π is a probability distribution,

$$\Pi \in \Delta = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \geq 0, \sum x_i = 1\} \quad .$$

Previous work has assumed that the searcher knows Π . Instead, I assume he has a prior distribution $F(\cdot)$ over Δ . As he continues to search he gathers more information about the distribution of prices which he assimilates by updating his prior according to Bayes Rule. All necessary information about his experience is contained in the statistics $N = (N_1, \dots, N_n)$, where N_i is the number of times price i has been observed.

It is convenient to parameterize his experience slightly differently. If $S(N) = \sum_i N_i$, then the vector whose i^{th} element is

$$\mu_i = \frac{N_i}{S(N)} \quad (3)$$

represents the average number of times that each price has been observed, while

$$\rho = S(N)^{-1} \quad (4)$$

represents the total number of prices he has observed; therefore, (μ, ρ) also contains all the information the searcher has accumulated. This parameterization permits a distinction between the content of this information, represented by μ , and its precision, represented by ρ . With these definitions, new information is assimilated as follows: With the observation of price i , (μ, ρ) becomes

$$h_i(\mu, \rho) = \left(\frac{\mu_1}{\rho + 1}, \dots, \frac{\mu_i + \rho}{\rho + 1}, \dots, \frac{\mu_n}{\rho + 1}, \frac{\rho}{\rho + 1} \right). \quad (5)$$

There is a slight technical difficulty with this convention. The updating rule (5) is not consistent with the definitions (3) and (4) when $N = 0$.

This problem may be finessed. Suppose $X = (x_1, \dots, x_n)$ and $t = (t_1, \dots, t_n)$; define

$$X^t = x_1^{t_1} x_2^{t_2} \dots x_n^{t_n}. \quad (6)$$

A man with initial beliefs $F(\Pi)$ and no experience, is no different from a man with initial beliefs $\tilde{F}(\Pi)$ and experience $N = (N_1, \dots, N_n)$ where $\tilde{F}(\Pi)$ is a probability distribution over Δ satisfying

$$\tilde{F}(\Pi) \Pi^N = K F(\Pi)$$

for some positive constant K , or, using the definitions (3) and (4),

$$\tilde{F}(\Pi) \Pi^{(\rho^{-1} \mu)} = K F(\Pi) .$$

In the sequel, we largely ignore the prior distribution $F(\cdot)$ and focus instead on the information

$$(\mu, \rho) \in \Gamma = \Delta \times [0, 1] \quad (7)$$

which is updated according to (5).

Corresponding to any (μ, ρ) is a vector defined by

$$\lambda_i(\mu, \rho) = \frac{\int_{\Delta} \pi_i \Pi^{\mu/\rho} dF(\Pi)}{\int_{\Delta} \Pi^{\mu/\rho} dF(\Pi)} . \quad (8)$$

$\lambda(\mu, \rho)$ is a probability distribution which represents the searchers' expected beliefs in that he would take a small bet on the proposition that the next price observed would be p_i at the odds $\lambda_i(\mu, \rho)$ to $(1 - \lambda_i(\mu, \rho))$. Since $\lambda(\mu, \rho)$ may be considered the index of a posterior distribution based on a sample size of ρ^{-1} which converges to a normal distribution with mean equal to the sample mean as sample size increases,

$$\lim_{\rho \rightarrow 0} \lambda_i(\mu, \rho) = \mu_i . \quad (9)$$

III. An example

A simple and important example will illustrate the nature of the searcher's problem and clarify the meaning of N , (μ, ρ) and $\lambda(\mu, \rho)$. Suppose that the searcher's prior is a Dirichlet distribution. Since the Dirichlet is the conjugate prior of the multinomial distribution, the posterior distribution will also be a Dirichlet. The Dirichlet is an n parameter distribution, completely characterized by the numbers $N = (N_1, \dots, N_n)$.^{11/} The properties of the Dirichlet are best illustrated by the following parable, which describes a problem exactly equivalent to that faced by the searcher.^{12/} Assume the N_i are integers. In an urn there are $S(N) = \sum N_i$ pieces of paper; p_1 is written on N_1 of these slips, p_2 on N_2 of them, and so on. A man draws from the urn at random. If he draws a slip with p_i on it he may either pay p_i or pay c , return the slip to the urn, place another slip with p_i on it in the urn, and draw from the urn again. That is, if he chooses to continue sampling, he faces the same problem as before except that the parameter describing the urn is

$$J_i(N) = N + e_i \quad (10)$$

where e_i is the vector with 1 in the i^{th} place and 0's elsewhere. The probability of getting price p_i from an urn with parameter N is just $\lambda_i(N) = N_i/S(N)$. $S(N)$ measures how fast these probabilities will change with successive drawings. If $S(N)$ is small new drawings will alter the composition of the urn considerably; if $S(N)$ is large they will hardly affect it. There is no reason why the N_i have to be integers. The above problem is equivalent to the searcher's problem for any positive N as long as the probability of drawing p_i from an urn with parameter N is $N_i/S(N)$

and N is updated by (10) when p_i is observed. Letting $\mu_i = N_i/S(N)$ and $\rho = S(N)^{-1}$ it is easy to check that the updating formula (10) is consistent with (5) and that $\lambda_i(\mu, \rho) = \lambda_i(N) = N_i/S(N) = \mu_i$.

It is important to realize that this story does not describe how price quotations are generated. In fact, there is a real price distribution which generates price quotation. However, the searcher does not know this distribution. His knowledge of the distribution is described by N , the composition of the urn. If price quotations are really generated by Π then the law of large numbers states that the proportion of prices in the urn will, with probability one, eventually be equal to Π .

The Dirichlet case is more general than it might appear. We have shown that if the prior is Dirichlet then it is possible to parameterize experience so that $\lambda_i(\mu, \rho) = \mu_i$. As we show below this property characterizes the Dirichlet completely. Thus, it follows from (9) that as experience accumulates, all searchers come to behave as if their priors were Dirichlet.

Proposition 1: *A searcher has a Dirichlet prior if and only if it is possible to parameterize his experience so that $\lambda_i(\mu, \rho) = \mu_i$ for all $(\mu, \rho) \in \Gamma$.*

Proof: The "only if" part of the proposition has already been demonstrated. Suppose $F(\cdot)$ is a probability measure on Δ and that when $\lambda_i(\mu, \rho)$ is defined by (8), $\lambda_i(\mu, \rho) = \mu_i$. There is a Dirichlet distribution $\tilde{F}(\cdot)$ so that, when $\tilde{\lambda}(\mu, \rho)$ is defined analogously, $\tilde{\lambda}(\mu, \rho) = \lambda(\mu, \rho)$. It follows that for all $N \geq 0$,

$$\frac{\int_{\Delta} \Pi^N + e_i dF(\Pi)}{\int_{\Delta} \Pi^N dF(\Pi)} = \frac{\int_{\Delta} \Pi^N + e_i d\tilde{F}(\Pi)}{\int_{\Delta} \Pi^N d\tilde{F}(\Pi)}$$

It is easy to see by induction that

$$\int_{\Delta} \Pi^N dF(\Pi) = \int_{\Delta} \Pi^N d\tilde{F}(\Pi)$$

for all non-negative integer N . The Stone-Weierstrass theorem implies that

$$\int_{\Delta} u(\Pi) dF(\Pi) = \int_{\Delta} u(\Pi) d\tilde{F}(\Pi)$$

for all bounded continuous functions, $u(\cdot)$, defined on Δ . Thus, by Theorem 1.3 of Billingsley (1968), $F(\cdot)$ and $\tilde{F}(\cdot)$ coincide.

IV. The optimal strategy

It is now possible to describe the optimal strategy of a man whose knowledge of prices is represented by the parameters (μ, ρ) which are updated according to (4). This is done in the standard way, by induction.

Let

$$V_0(\mu, \rho) = \sum_i \lambda_i(\mu, \rho) p_i \quad (12.a)$$

and

$$V_T(\mu, \rho) = \sum_i \lambda_i(\mu, \rho) [\text{Min}(p_i, V_{T-1}(h_i(\mu, \rho)) + c)]. \quad (12.b)$$

$V_{T-1}(\mu, \rho)$ is the minimum expected cost incurred by a man with prior experience (μ, ρ) who is allowed to search at most T times but must accept the T^{th} price offer made to him. It is easy to see by induction that, $V_T(\mu, \rho) \leq V_{T-1}(\mu, \rho)$ and $V_T(\mu, \rho) \geq p_1$ for all T , so that the $V_T(\mu, \rho)$ converge. Let

$$V(\mu, \rho) = \lim_{T \rightarrow \infty} V_T(\mu, \rho) .$$

Then $V(\mu, \rho)$ satisfies

$$V(\mu, \rho) = \sum_i \lambda_i(\mu, \rho) \text{Min}[p_i, V(h_i(\mu, \rho)) + c] , \quad (13)$$

an equation which defines the optimal policy: If p_i is drawn when beliefs are (μ, ρ) accept if

$$p_i \leq V(h_i(\mu, \rho)) + c \quad (14)$$

otherwise elicit another price offer.

Proposition 2: $V(\mu, \rho)$ is continuous.

This fact, although of little interest itself, is the basis of the

proof of the next section and justifies the use of extreme examples in the sections that follow. It is important to realize that all the proofs given hold for any $(\mu, \rho) \in \Gamma$ including such strange boundary values as $(\mu = e_i, \rho = 0)$ or $(\mu = e_i, \rho = 1)$ where e_i is the vector whose i^{th} component is 1 while all other components are equal to 0.

Proof: Since $\lambda_i(\mu, \rho)$ is continuous so is $V_0(\mu, \rho)$. A simple induction establishes that $V_T(\mu, \rho)$ is continuous for all T . It remains to show that the V_T converge uniformly.

Let $\bar{V}_T(\mu, \rho)$ be the expected total cost to a man who for T periods follows the policy of accepting the elicited price only if (14) is satisfied; if after T periods no price has been accepted, the man must pay p_n . $\underline{V}_T(\mu, \rho)$ is defined similarly except that if he has not chosen after T periods he receives p_1 . From their definitions it follows that

$$\bar{V}_T(\mu, \rho) \geq V_T(\mu, \rho) \geq V(\mu, \rho) \geq \underline{V}_T(\mu, \rho) \quad .$$

It will suffice to show that $|\bar{V}_T(\mu, \rho) - \underline{V}_T(\mu, \rho)|$ converges to 0 uniformly.

$$|\bar{V}_T(\mu, \rho) - \underline{V}_T(\mu, \rho)| \leq \beta_T(\mu, \rho) p_n$$

where $\beta_T(\mu, \rho)$ is the probability -- calculated according to the prior (μ, ρ) -- that sampling will not terminate after T periods. However, from the definition of $V_T(\mu, \rho)$ it must be that

$$\beta_T(\mu, \rho) T c \leq V_T(\mu, \rho) \leq p_n$$

so

$$\beta_T(\mu, \rho) \leq \frac{p_n}{Tc}$$

and $\beta_T(\mu, \rho)$ converges to 0 uniformly. This completes the proof.

V. Finiteness of search

In this section it is shown that search will cease after a finite time. The idea of the proof (for which I am indebted to Robert Lucas) is a straightforward one. Regardless of his beliefs a person will always accept p_1 . If p_1 does not appear for a long time then he will come to believe that p_2 is the lowest possible price and will of course accept it. Thus, there is a t_2 such that after t_2 trials either the person has stopped sampling or he will be willing to accept p_2 . Similar arguments establish the existence of t_3, t_4, \dots, t_n with the same property. After t_n trials, the searcher has either stopped or will accept any price that occurs.

Theorem I: *There is a number t such that a person following the optimal strategy will have stopped searching after t trials.*

Proof: The theorem is not a probability statement. Let Ω be the set of all infinite sequences of the prices p_1, \dots, p_n and by ω denote an element of Ω . We shall show that there is a t such that for any ω , sampling will have stopped by the time the t^{th} price is quoted. Before we do so some notation is necessary. For any ω , let ω_s denote its s^{th} coordinate. Since (μ, ρ) are updated as functions of ω , we may define $(\mu, \rho)(\omega(q))$ as the values of (μ, ρ) obtained after observation of $\omega_1, \dots, \omega_q$.

Define the acceptance set A_i by $A_i = \{(\mu, \rho) \in \Gamma \mid p_i < V(h_i(\mu, \rho)) + c\}$. It follows from (9) that if $K_2 = \{(\mu, \rho) \in \Gamma \mid \mu_1 = \rho = 0\}$ then $K_2 \subset A_2$. Since A_2 is open and D compact there is an $\varepsilon > 0$ such that if $(\mu_1, \dots, \mu_n, \rho) \in \Gamma$ and $\mu_1 + \rho < \varepsilon$, then $(\mu_1, \dots, \mu_n, \rho) \in A_2$. We may choose t_2 such that if

$$\omega_s \neq p_1 \text{ for } s < t_2 \quad (15)$$

then $\mu_1(\omega(t_2)) + \rho(\omega(t_2)) < \epsilon$; thus if ω satisfies (15), $(\mu, \rho)(\omega(t_2)) \in A_2$. Furthermore, if ω satisfies (15) and $\omega_{s_1} = p_2$ for $s_1 > t_2$, then sampling will have terminated by s_1 since either $\omega_s = p_1$ for $t_2 < s < s_1$ or $(\mu, \rho)(\omega(s_1)) \in A_2$.

Now let $K_3 = \{(\mu, \rho) \in \Gamma \mid \mu_1 = \mu_2 = \rho = 0\}$. Clearly $K_3 \subset A_3$ and, by arguments analagous to those used above, there exists a t_3 such that if ω satisfies (15) and

$$\omega_s \neq p_1; \quad \omega_s \neq p_2 \quad \text{for } t_2 \leq s < t_3$$

then

$$(\mu, \rho)\omega(t_3) \in A_2 \cup A_3 .$$

Continuing in this we establish the existence of a t_n such that for all ω either sampling has stopped by t_n or $(\mu, \rho)(\omega(t_n)) \in \bigcup_{i=1}^n A_i$. Letting $t = t_n$ completes the proof.

This theorem has two important implications. First, it suggest that computing optimal search rules and the expected costs of following them is a finite problem; thus that it should be possible to compute the loss from following such ad hoc rules as those discussed by Telser (1973) rather than following optimal rules. Secondly, it shows that to prove propositions about $V(\mu, \rho)$ it is sufficient to prove them about all $V_T(\mu, \rho)$. This technique is used repeatedly in the sequel. For example, it may be used to prove that Property (3) holds in general.

Proposition 3: *As the costs of search increase, search decreases.*

Proof: Since the optimal search rule is to keep on searching whenever

$p_i \geq V(h_i(\mu, \rho)) + c$, it will suffice to show that if $V(\mu, \rho)$ is the value of having information (μ, ρ) when costs are c and $\hat{V}(\mu, \rho)$ is the value of having the same information when costs are $\hat{c} \geq c$, then $\hat{V}(\mu, \rho) \geq V(\mu, \rho)$. Theorem I implies that this will be true if, for all T ,

$$\hat{V}_T(\mu, \rho) \geq V_T(\mu, \rho)$$

But this is clearly true if $T = 0$ and that it is true for all T follows by induction from (12.b), the definition of V_T .

VI. Reservation prices

When the distribution of prices is known, optimal search rules are characterized by a reservation price; there is a price p_R such that the searcher will accept all prices less than or equal to p_R and reject all prices greater than p_R . When the distribution of prices is unknown, acceptable prices change as information changes, so that optimal Bayesian search procedures cannot be characterized by a single reservation price. It is however worthwhile to ask whether or not they have a reservation price property, that is, whether for every state of information (μ, ρ) there is a $p_{R(\mu, \rho)}$ such that prices below $p_{R(\mu, \rho)}$ are accepted and those above it are rejected; if the acceptance sets of all customers in a market are characterized by reservation prices, then the demand function of each seller in that market is well behaved -- if he raises prices expected sales will not increase. Raising prices may increase sales if customers acceptance sets are not characterized by reservation prices.

In general, optimal Bayesian search rules do not have the reservation price property. Counter examples are easy to construct. Suppose there are three prices, \$1.00, \$2.00 and \$3.00 and that the cost of search is \$.01. Prior beliefs admit the possibility of only two distributions of prices. Either all prices are \$3.00 or they are distributed between \$1.00 and \$2.00 in the proportions 99 to 1. A man with these beliefs should accept a price of \$3.00 (as this is a signal that no lower prices are to be had) and reject a quote of \$2.00 (which indicates that the likelihood that a much better price will be obtained on another draw is high).

It is easy to see what makes this counter example work. Price quotations have value as information. If my beliefs are (μ, ρ) then, the informa-

tional value of getting price i is just $V(h_i(\mu, \rho))$. In the counter example above, the differences in the information value of prices far exceeded the differences in the prices themselves. If differences in the value of price information are less than differences in prices, then optimal search rules have the reservation price property.

Proposition 3: If for all $(\mu, \rho) \in \Delta$,

$$|p_i - p_k| \geq |V(h_i(\mu, \rho)) - V(h_k(\mu, \rho))|, \quad (16)$$

then

$$V(h_i(\mu, \rho)) + c \geq p_i \geq p_k \quad (17)$$

implies

$$V(h_k(\mu, \rho)) + c \geq p_k. \quad (18)$$

Note that (17) implies (18) is precisely the reservation price property.

Proof: Suppose that (16) holds. Then, using (17),

$$|p_i - p_k| = p_i - p_k \geq |V(h_i(\mu, \rho)) - V(h_k(\mu, \rho))| \geq V(h_i(\mu, \rho)) - V(h_k(\mu, \rho)).$$

Thus $p_k - V(h_k(\mu, \rho)) + c \leq p_i - V(h_i(\mu, \rho)) + c \leq 0$.

This is a useful criterion. It is used to prove

Theorem II. Optimal rules for searchers with Dirichlet priors have the reservation price property.

In view of Proposition 1 and Equation (9) this is about the most general result which could be hoped for in this context. The proof which follows is purely formal and conveys little insight. It is not, however, hard to see why the theorem should be true. Proposition 3 implies that search rules will have the reservation price property whenever observing

a price does not convey information about the relative likelihood -- as measured by $\lambda(\mu, \rho)$ -- of observing other prices in the future. The rule for updating the Dirichlet is completely neutral in this sense. For all distinct i, j , and k observing p_i has no effect on λ_j/λ_k .

Proof: In view of Propositions 1 and 3 it will suffice to show that

$$\lambda_i(\mu, \rho) = \mu_i \quad (19)$$

implies

$$|p_i - p_k| \geq |V(h_i^s(\mu, \rho)) - V(h_k^s(\mu, \rho))| \quad (20)$$

where

$$h_i^s(\mu, \rho) = \left(\frac{\mu_1}{1 + s\rho}, \dots, \frac{\mu_i + s\rho}{1 + s\rho}, \dots, \frac{\mu_n}{1 + s\rho}, \frac{\rho}{1 + s} \right) \quad (21)$$

is (μ, ρ) updated according to (5) after s observations of p_i . Note that (16) is just (21) with $s = 1$. Theorem I implies that it will suffice to show, for all t , 13/

$$|p_i - p_k| \geq |V_t(h_i^s) - V_t(h_k^s)| \quad (22)$$

which can be done by induction. Suppose, as we shall throughout the proof,

$$p_i > p_k \quad (23)$$

Then from (12) and (21),

$$V_0(h_i^s) = \frac{1}{1 + s\rho} \sum_j \mu_j p_j + \frac{s\rho}{1 + s} p_i \quad .$$

Thus,

$$V_0(h_i^S) - V_0(h_k^S) = \frac{s\rho}{1+s\rho} (p_i - p_k) \leq (p_i - p_k)$$

so that (22) holds for $t = 0$. The inductive step, that if (22) holds for $t = T - 1$, then it holds for $t = T$, is proven in a series of lemmas below.

Lemma 1: $p_i \geq p_k$ implies $V_T(h_i^S) \geq V_T(h_k^S)$.

Proof: We have already shown that

$$V_0(h_i^S) - V_0(h_k^S) = \frac{s\rho}{1-\rho} (p_i - p_k) > 0.$$

Suppose the lemma true for $t = T - 1$. Then letting

$$h_j h_i^S = h_j (h_i^S(\mu, \rho)) = h_i^S h_j,$$

$$\begin{aligned} V_T(h_i^S) - V_T(h_k^S) &= \frac{1}{1+\rho} \sum_j \mu_j [\text{Min}(p_j, V_{T-1}(h_i^S h_j) + c) \\ &\quad - \text{Min}(p_j, V_{T-1}(h_k^S h_j) + c)] \\ &\quad + \frac{\rho}{1+\rho} [\text{Min}(p_i, V_{T-1}(h_i^{S+1}) + c) \\ &\quad - \text{Min}(p_k, V_{T-1}(h_k^{S+1}) + c)] \end{aligned} \quad (24)$$

The terms in square brackets which multiply $\frac{1}{1+\rho}$ are of the form $J(A,B,C) = \text{Min}(A,B) - \text{Min}(A,C)$, where $B > C$ by the induction hypotheses. Thus they are all non-negative. Consider the term multiplying $\frac{\rho}{1+\rho}$. It is of the form

$$J(A,B,C,D) = \text{Min}(A,B) - \text{Min}(C,D) \quad (25)$$

where $A = p_i$, $B = V_{T-1}(h_i^{S+1}) + c$, $C = p_k$, and $D = V_{T-1}(h_k^{S+1}) + c$. Thus $A \geq C$ and $B \geq D$ by the induction hypotheses. These two inequalities imply (25) is non-negative. Since (24) is a weighted average of non-negative

quantities, it is non-negative.

Lemma 2: If (22) holds for $t = T - 1$ then

$$\begin{aligned} \text{Min}(p_j, V_{T-1}(h_j, h_k^s) + c) - \text{Min}(p_j, V_{T-1}(h_j, h_k^s) + c) \\ \leq p_i - p_k \end{aligned} \quad (26)$$

Proof: The L.H.S. of (26) has the form

$$J(A, B, C) = \text{Min}(A, B) - \text{Min}(A, C)$$

where $A = p_j$, $B = V_{T-1}(h_j, h_i^s) + c$ and $C = V_{T-1}(h_j, h_k^s) + c$. Together (22) and Lemma 1 imply

$$0 \leq B - C \leq p_i - p_k \quad (27)$$

If $B \leq A$, then $C \leq A$ and $J = B - C$. If $B > A$, either $C \leq A$ and $J = A - C \leq B - C$ or $C > A$ and $J = A - A = 0 \leq B - C$. Thus $J \leq (B - C) \leq p_i - p_k$.

Lemma 3: If (22) holds for $t = T - 1$,

$$\begin{aligned} \text{Min}(p_i, V_{T-1}(h_i^{s+1}) + c) - \text{Min}(p_i, V_{T-1}(h_k^{s+1}) + c) \\ \leq p_i - p_k \end{aligned} \quad (28)$$

Proof: The R.H.S. of (28) is of the same general form as (25) with $A = p_i$, $B = V_{T-1}(h_i^{s+1}) + c$, $C = p_k$ and $D = V_{T-1}(h_k^{s+1}) + c$. Also by hypothesis, $A \geq C$ and Lemma 1 implies $B \geq D$. We want to show that $J \leq A - C$. There are four cases to consider:

- (i) $A \leq B$ and $C \leq D$; so that $J = A - C$.
- (ii) $A \leq B$ and $C > D$. This case may be disregarded since if it obtains $B - D \geq A - D > A - C$ or $V_{T-1}(h_i^{s+1}) - V_{T-1}(h_k^{s+1}) > p_i - p_k$ which contradicts

(22).

(iii) $A > B$ and $C > D$ so that $J = (B - D) \leq A - C$ by (22).

(iv) $A > B$ and $C \leq D$ so that $J = (B - C) < A - C$ since $A > B$.

Lemma 4: If (22) holds for $t = T - 1$ it holds for $t = T$.

Proof: Use Lemmas 2 and 3 to calculate

$$\begin{aligned}
 & |V_T(h_i^S) - V_T(h_k^S)| = \\
 & V_T(h_i^S) - V_T(h_k^S) = \\
 & \frac{1}{1 + sp} \sum_j \mu_j [\text{Min}(p_j, V_{T-1}(h_j h_i^S) + c) - \text{Min}(p_j, V_{T-1}(h_j h_k^S) + c)] \\
 & + \frac{sp}{1 + sp} [\text{Min}(p_i, V_{T-1}(h_i^{S+1}) + c) - \text{Min}(p_k, V_{T-1}(h_k^{S+1}) + c)] \\
 & \leq \frac{1}{1 + sp} \sum_j \mu_j (p_i - p_k) + \frac{sp}{1 + sp} (p_i - p_k) = \\
 & (p_i - p_k) = |p_i - p_k| .
 \end{aligned}$$

This completes the proof of Theorem II.

Theorem II implies that the reservation price of those searchers who remain in the market must eventually increase to p_n . I conjecture (but have not been able to prove) that if the prior is Dirichlet, the reservation price will increase monotonically -- as is occasionally postulated in models of the market behavior of searchers (e.g., Diamond 1971).

VII. Effects of increasing uncertainty

The introduction listed some results on the effects of increased price dispersion when the price distribution is known. It is natural to ask whether these results -- that increased price dispersion lowers total expected costs and increases search activity -- hold when the distribution is unknown. Since the searcher's knowledge of the price distribution is defined in the parameters (μ, ρ) , the question is what effect changes in μ and ρ which represent increased uncertainty have on $V(\mu, \rho)$. To answer this it is necessary to decide what changes in (μ, ρ) represent increased uncertainty. There are two obvious candidates: $\lambda(\mu, \rho)$ is the expected price distribution in the sense that the man is just willing to bet that the next price will be p_i at odds $\lambda_i(\mu, \rho)$ to $(1 - \lambda_i(\mu, \rho))$. However, ρ represents the precision of the searcher's knowledge of the price distribution he faces ($\rho = 0$ is subjective certainty). Thus $\lambda(\mu, \rho)$ represents what the searcher believes while ρ represents how firmly he believes it. Increased uncertainty could correspond either to an increase in the dispersion of $\lambda(\mu, \rho)$, holding ρ constant, or to an increase in ρ , holding $\lambda(\mu, \rho)$ constant. With what I hope is a pardonable abuse of language, I shall call the first case increasing objective uncertainty, and the second case increasing subjective uncertainty. This terminology is justified by the fact that increasing objective uncertainty is the natural analogue of the increases in uncertainty studied when price distributions are assumed known. Both cases are discussed below.

The effects of increased subjective uncertainty are ambiguous. As seems reasonable, it appears that whether increasing the subjective certainty with which a searcher holds his beliefs increases or decreases expected costs

depends on what those beliefs are. If my expectations are of the best, $\lambda_i(\mu, \rho) = (1, 0, \dots, 0; \rho)^{14/}$ then increased confidence will decrease expected costs. That is,

$$V(1, \dots, 0; \rho) > V(1, \dots, 0; 0) = p_1$$

for any $\rho > 0$, so that $V(\mu, \rho)$ is increasing in ρ for at least some ρ . Similarly,

$$V(0, \dots, 1; \rho) < V(0, \dots, 1; 0) = p_n$$

so that $V(\mu, \rho)$ is decreasing in ρ for some ρ . It follows that nothing general can be said about the effect of increased subjective uncertainty on the willingness of searchers to accept or reject prices.

The effects of increased objective uncertainty, at least for the Dirichlet case, are not ambiguous.

Theorem III: *If the searcher's prior is a Dirichlet ($\lambda(\mu, \rho) = \mu$), then if $\tilde{\mu}$ is riskier than μ , $V(\tilde{\mu}, \rho) \leq V(\mu, \rho)$.*

Proof: Suppose $\tilde{\mu}$ is riskier than μ . It will suffice to prove the theorem if μ and $\tilde{\mu}$ differ by a single mean preserving spread (Rothschild and Stiglitz, 1970) so we shall consider four prices

$$p_1 < p_2 < p_3 < p_4$$

such that $\mu_i = \tilde{\mu}_i$ for $i \notin (1, 2, 3, 4)$, and

$$\tilde{\mu}_1 - \mu_1 = \mu_2 - \tilde{\mu}_2 = A > 0 \quad (29)$$

$$\mu_3 - \tilde{\mu}_3 = \tilde{\mu}_4 - \mu_4 = B > 0 \quad (30)$$

where

$$A(p_2 - p_1) + B(p_3 - p_4) = 0 \quad (31)$$

We shall prove the theorem by showing

$$V_T(\tilde{\mu}, \rho) \geq V_T(\mu, \rho) \quad (32)$$

for all T by induction. That (32) holds for $T = 0$ follows from (31) and (12). In an obvious notation, let $\tilde{h}_j = h_j(\mu, \rho)$.

Let R satisfy

$$V_T(\mu, \rho) = \sum_1^R \mu_j p_j + \sum_{R+1}^N \mu_j (V_{T-1}(h_j) + c).$$

Theorem II guarantees the existence of such an R . Note that

$$\begin{aligned} V_T(\tilde{\mu}, \rho) &= \sum \tilde{\mu}_j \text{Min} [p_j, V_{T-1}(\tilde{h}_j) + c] \\ &\leq \sum_1^R \tilde{\mu}_j p_j + \sum_{R+1}^n \tilde{\mu}_j (V_{T-1}(\tilde{h}_j) + c) \\ &\leq \sum_1^R \tilde{\mu}_j p_j + \sum_{R+1}^n \tilde{\mu}_j (V_{T-1}(h_j) + c) \end{aligned}$$

(the last inequality follows from the induction hypothesis). Let

$$H(\mu, \tilde{\mu}) = \sum_1^R (\mu_j - \tilde{\mu}_j) p_j + \sum_{R+1}^n (\mu_j - \tilde{\mu}_j) (V_{T-1}(h_j) + c) \quad (33)$$

Then, it will suffice to show that $H(\mu, \tilde{\mu}) \geq 0$.

There are five cases to consider, depending on the relationship of p_R to p_1, p_2, p_3, p_4 .

1. $p_R \geq p_4$. In this case, $H(\mu, \tilde{\mu}) = (\mu_1 - \tilde{\mu}_1) p_1 + (\mu_2 - \tilde{\mu}_2) p_2 + (\mu_3 - \tilde{\mu}_3) p_3 + (\mu_4 - \tilde{\mu}_4) p_4 = A(p_2 - p_1) + B(p_3 - p_4) = 0$.

2. $p_4 > p_R \geq p_3$. In this case, $H(\mu, \tilde{\mu}) = A(p_2 - p_1) +$

$$B(p_3 - (V_{T-1}(h_4) + c)) \geq A(p_2 - p_1) + B(p_3 - p_4) = 0.$$

3. $p_3 > p_R > p_2$. In this case, $H(\mu, \tilde{\mu}) = A(p_2 - p_1) +$

$$B(V_{T-1}(h_3) - V_{T-1}(h_4)) \geq A(p_2 - p_1) + B(p_3 - p_4) = 0. \text{ The second inequality}$$

follows from (22).

4. $p_2 > p_R \geq p_1$. In this case,

$$\begin{aligned} H(\mu, \tilde{\mu}) = & A(V_{T-1}(h_2) + c - p_1) + B(V_{T-1}(h_3) - V_{T-1}(h_4)) \\ & A(V_{T-1}(h_2) - V_{T-1}(h_1)) + B(V_{T-1}(h_3) - V_{T-1}(h_4)). \end{aligned} \quad (34)$$

It is shown below in Lemma 5 that

$$\frac{V_{T-1}(h_2) - V_{T-1}(h_1)}{p_2 - p_1} \geq \frac{V_{T-1}(h_4) - V_{T-1}(h_3)}{p_4 - p_3} \quad (35)$$

so that

$$A(V_{T-1}(h_2) - V_{T-1}(h_1)) \geq \frac{A(p_2 - p_1)}{p_4 - p_3} (V_{T-1}(h_4) - V_{T-1}(h_3)).$$

Combining (34) and (35) we obtain

$$H(\mu, \tilde{\mu}) \geq \frac{V_{T-1}(h_4) - V_{T-1}(h_3)}{p_4 - p_3} (A(p_2 - p_1) + B(p_4 - p_3)) = 0.$$

5. $p_1 \geq p_R$. In this case, $H(\mu, \tilde{\mu})$ is just equal to the last expression in (34) and (35) can be used to show $H(\mu, \tilde{\mu}) \geq 0$.

To complete the proof we only need prove

Lemma 5: (35) holds for all T .

Proof: If $T = 0$, both sides of (35) are equal to $(\rho/(1 + \rho))$. Suppose (35) holds for $t = T - 1$. Let R_k satisfy

$$V_T(h_k) = \sum_1^{R_k} \mu_j p_j + \sum_{R_k+1}^n \mu_j (V_{T-1}(h_{jk}) + c) .$$

It follows from Lemma 1 that $R_1 \leq R_2 \leq R_3 \leq R_4$, so that

$$V_T(h_2) - V_T(h_1) \leq \sum_{R_3+1}^n \mu_j (V_{T-1}(h_{2j}) - V_{T-1}(h_{1j})) \quad (36)$$

while

$$V_T(h_4) - V_T(h_3) \leq \sum_{R_3+1}^n \mu_j (V_{T-1}(h_{4j}) - V_{T-1}(h_{3j})) . \quad (37)$$

The induction hypothesis implies

$$\begin{aligned} 0 &\leq \sum_{R_3+1}^n \mu_j \left[\frac{V_{T-1}(h_{2j}) - V_{T-1}(h_{1j})}{p_2 - p_1} - \frac{V_{T-1}(h_{4j}) - V_{T-1}(h_{3j})}{p_4 - p_3} \right] \\ &\leq \frac{V_{T-1}(h_2) - V_{T-1}(h_1)}{p_2 - p_1} - \frac{V_{T-1}(h_4) - V_{T-1}(h_3)}{p_4 - p_3} . \end{aligned}$$

The last inequality follows from (36) and (37).

An immediate consequence of Theorem III is an analogue of the property 5) that increased price dispersion leads to increased search.

Corollary: *If the searcher's prior is Dirichlet, then increasing objective uncertainty lowers the reservation price.*

The proof which is obvious is omitted.

VIII. Conclusion

The most important results of this paper, Theorems II and III, state that for what Proposition 1 suggests is a quite significant example, the qualitative behavior of persons searching optimally from unknown distributions is the same as that of persons searching optimally from known distributions. Since it is easy to construct examples for which this is not true, it is natural to ask just how general these results are. This problem is best stated by focusing on a very general and abstract formulation of the problem of optimal search. Let K be a compact ^{15/} set of the real line and let $\mathcal{M}(K)$ be the set of all probability measures on K . $\mathcal{M}(K)$ is a compact separable metric space so that we can, in the standard way, define probability measures on $\mathcal{M}(K)$.^{16/} Such a probability measure, say \mathcal{P} , represents a searcher's a priori beliefs about the price distribution he faces (which is some probability distribution in $\mathcal{M}(K)$). If a price p is observed then as long as the conditional distributions (on $\mathcal{M}(K)$ given p) can be calculated, Bayes rule may be used to update his beliefs from \mathcal{P} to, say, \mathcal{P}' . As in sections II and IV above, the optimal Bayesian strategy can be devised. \mathcal{P} completely determines the optimal search rule. The question at hand is: For what class of \mathcal{P} do analogues of Theorems II and III hold?

Each \mathcal{P} has two rather distinct aspects. The first is $S(\mathcal{P})$, the support of \mathcal{P} , which is roughly the set of price distributions which the man whose beliefs are described by \mathcal{P} judges to be conceivable. This paper has focused on the case where $S(\mathcal{P})$ contains only multinomial distributions. The second aspect of \mathcal{P} is the updating rule. Theorems II and III required that the updating rule be of a very special kind -- if p_i is observed then the (subjective) likelihood of observing p_i in the future is increased while

the likelihood of observing all other prices is decreased (and by the same proportion).

Some recent work of Ferguson (1973)^{17/} suggests that it is only the restrictions on the updating rule and not those on $S(\mathcal{P})$ which are required for Theorems II and III. Ferguson introduces a class of distributions whose support is essentially all of $\mathcal{M}(K)$; I conjecture that analogues of Theorems I, II, and III hold for these distributions. However, these probability measures imply very special updating rules. Ignoring many technicalities, Ferguson's work may be summarized as follows: Let α be a finite measure on K and let $B = (B_1, B_2, \dots, B_n)$ be a measurable partition of K (that is, the B_i are measurable and $\cup_i B_i = K$ and $B_i \cap B_j = \emptyset$ if $i \neq j$). Consider the probability that the next observed price belongs to $B_1, B_2,$ or B_n . This is a multinomial distribution described by the numbers $(P(B_1), \dots, P(B_n)) \in \Delta$. Each such multinomial distribution belongs to $\mathcal{M}(K)$ and is thus assigned probability by \mathcal{P} . \mathcal{P}_α is a Ferguson distribution with parameter α if the distribution of $(P(B_1), P(B_2), \dots, P(B_n))$ is Dirichlet with parameter $(\alpha(B_1), \alpha(B_2), \dots, \alpha(B_n))$. Ferguson shows (1973, pp. 215-216) that the support of \mathcal{P}_α is the set of all measures $\mu \in \mathcal{M}(K)$ which are absolutely continuous with respect to α . By choosing α appropriately $S(\mathcal{P}_\alpha)$ can be made very large.

The rule for updating \mathcal{P}_α is very simple. If \mathcal{P}_α is a Ferguson distribution with parameter α , after price p is observed it is updated to a Ferguson distribution with parameter

$$h_p(\alpha) = \alpha + \delta_p \quad (39)$$

where δ_p is the probability measure concentrated on the point p . This formula is an exact analogue of the rule (10) given above for updating

Dirichlet priors.

Proving analogues of Theorems I, II, and III for Ferguson distributions whose parameters have compact support should involve only technical problems.^{18/} Therefore the results of this paper are quite general in the sense that they apply to searchers whose beliefs are not restricted in any way beyond boundedness. However, they are quite special in that they apply only to people who revise their beliefs in a very special way. Formula (39) implies that information is strictly local, or to put the matter more colloquially, that a miss is as good as a mile. My future beliefs about the likelihood of observing p are affected in exactly the same way by observing $p + 10^{-10}$ as by observing $p + 10^{10}$. I would like to know whether the conclusions of theorems II and III hold when information is not so completely local. At present there seems no way to answer this question with generality. I know of no class of $\varphi \in \mathcal{M}(K)$ which both has large support and tractable updating rules other than the Ferguson distributions.

It is possible to examine the property of φ 's with rather restricted support by examining parametric families of distributions with tractable updating rules. For example, it is a simple consequence of DeGroot's (1968) work that optimal search rules from normal distributions with unknown means and known variance have the reservation price property when the prior distribution is also normal.

Perhaps the strongest possible result along the lines of Theorem II^{19/} which could be hoped for is the following

Conjecture: *Optimal search rules from exponential families of distributions have the reservation price property if the unknown parameters have*

conjugate prior distributions.

Establishing this would significantly expand the applicability of the results of this paper. Still I think enough has been done to establish that economists can without great loss assume that the qualitative properties of demand functions which arise from optimal search from unknown distributions are the same as those which arise from optimal search from known distributions.

FOOTNOTES

1/ See Stigler [1961, 1962], McCall [1965, 1970], Nelson [1970], Gastwirth [1971], Kohn and Shavell [1973] and Telser [1973].

2/ This fee, the cost of search, is generally interpreted as the cost of visiting a store and obtaining a price quotation. It could also include the cost of doing without the item in question while search continues, in which case the fact that search takes time would have to be formally incorporated in the model, see fn. 10 below. Other rather trivial, from the formal point of view, generalizations are possible. Price and cost may be measured in utility rather than money. However, the utility function must be linear so this is a small generalization. The same framework may be used to analyze search for jobs when wages (or the utility of jobs) are random and unknown. I shall largely ignore these matters of interpretation.

3/ A cynic might suspect that the lack of empirical work reflects a lack of interest in the real world on the part of those concerned with search rules. A glance at the authors in fn. 1, above, suggests that this is not the only reason. It seems to this author -- whose interest in such matters is admittedly casual -- that the problem of determining characteristics of search rules empirically is a very difficult one.

4/ This is generally demonstrated by calculating $\frac{dM}{dn} < 0$ for particular parametric families of distributions. However, the result is much more general. Let $F(p, t)$ be a family of distribution functions with support in $[0, B]$ indexed by t . Then, as Diamond and Stiglitz [1973] have shown, increases in t correspond to increases in risk in the sense of Rothschild and Stiglitz [1970] if

$$\int_0^B dF(p, t) = 1 \quad (i)$$

$$\int_0^Y F_t(p, t) dp \geq 0 \quad 0 \leq Y \leq B \quad (ii)$$

and

$$\int_0^B F_t(p, t) dp = 0 \quad (iii)$$

If $h(p)$ is any increasing function, then

$$\int_0^B h(p) F_t(p, t) dp \leq 0$$

Note that $M_n(t) = \int_0^B (1 - F(p, t))^n dp$ and

$$\frac{dM_n(t)}{dt} = \int_0^B -n(1 - F(p, t))^{n-1} \cdot F_t(p, t) dp \leq 0$$

since $-n(1 - F(p, t))$ is increasing in p .

5/ Total expected costs to a searcher following a reservation price rule are just equal to the reservation price. This follows from the fact that expected costs do not change as the searcher continues to sample. The logic of the rule (1) might be stated as: Accept any price less than the expected costs of continuing to search, otherwise continue.

6/ If in the notation of fn. 3, $G_n(t) = M_{n-1}(t) - M_n(t)$, then it is not true that $G_n(t)$ increases with t whenever $F(p, t)$ satisfies (i), (ii), and (iii). I suspect that most increases in price dispersion which do not lead to piling up of probability mass at the extreme points of the distribution will increase $G_n(t)$ but I have not yet been able to formulate this precisely.

7/ For discussion of the need for such a complete model see Rothschild [1973].

8/ I suspect they vary a great deal. Characteristics of empirical wage

distributions -- which are related to -- but distinct from distributions of wage offers -- are unstable over occupations and across cities, Buckley [1968].

9/ The statistical literature on optimal search rules is extensive. Breiman [1964], DeGroot [1970, Chapter 13], and Chow, Robbins and Siegmund [1971] are surveys of increasing sophistication. Some of the results of this paper appear to be new, particularly Theorems I, II and III. The work most closely related to that presented here is that of Yahav [1966] and DeGroot [1968]. Yahav considered the general problem of searching with recall (the searcher is allowed to return to stores previously sampled) from an unknown distribution, showed that optimal rules existed under quite general circumstances, and discussed how they could be approximated. De Groot completely solved the problem faced by a man searching (with or without recall) from a normal distribution of known variance and unknown precision.

10/ Given that the process takes place in time, perhaps it would be more logical to minimize discounted expected costs. This would introduce no complications. See Kohn and Shavell [1973] for a discussion of the effect of varying the discount rate on search rules.

11/ See, for example, DeGroot [1970, pp. 49-51] for a discussion of the properties of the Dirichlet distribution.

12/ I am indebted to David Blackwell for pointing out to me this way of looking at Dirichlet distributions.

13/ The arguments (μ , ρ) will be suppressed where this may be done without

confusion.

14/ Since $V(u, \rho)$ is continuous on Γ this is a meaningful example.

15/ I continue to insist that price distributions be bounded. No meaningful generality and much complication ensues if this assumption is abandoned.

16/ See Parthasarthy [1967, pp. 39 ff.] for details.

17/ Blackwell and MacQueen [1973] show how Ferguson distributions may be understood in terms of the urn model of section III. In terms of that model, we can phrase our question as follows: Do our results hold when the composition of the urn is altered (upon observation of price p_i) by some other way than adding another slip with p_i written on it?

18/ It should be possible to push through proofs very similar to those in the text, once a topology is found in which Proposition 2 holds.

19/ I am not certain how propositions about the effects of increasing dispersion are best stated in this context.

- J. L. Gastwirth, "On Probabilistic Models of Consumer Search for Information." Unpublished paper, 1971.
- M. G. Kohn and S. Shavell, Try it, you'll like it: A theory of rational search. To appear in Journal of Economic Theory (1973).
- J. J. McCall, The economics of information and optimal stopping rules. Journal of Business 38 (1965). 300-317.
- _____, Economics of information and job search. Quarterly Journal of Economics 84 (1970). 113-126.
- P. Nelson, Information and consumer behavior. Journal of Political Economy 78 (1970). 311-329.
- K. R. Partasarthy, Probability Measures on Metric Spaces. Academic Press, New York, 1967.
- M. Rothschild, Models of market organization with imperfect information. Journal of Political Economy 81 (1973).
- _____ and J. E. Stiglitz, Increasing risk: I. A definition. Journal of Economic Theory 2 (1970). 225-243.
- G. J. Stigler, The economics of information. Journal of Political Economy 69 (1961). 213-225.
- _____, Information in the labor market. Journal of Political Economy 70 No. 5 Part 2 (1962). 94-105.
- L. G. Telser, Searching for the lowest price. American Economic Review, Proceedings 63 (1973). 41-49.
- J. Yahav, On optimal stopping. Ann. Math. Stat. 37 (1966). 30-35.

REFERENCES

- P. Billingsley, Convergence of Probability Measures. Wiley, New York, 1968.
- D. Blackwell and J. B. MacQueen, Ferguson distributions via Polya urn schemes. Annals of Statistics 1 (1973): 353-355.
- L. Breiman, "Stopping-rule Problems," in Beckenbuch (ed.), Applied Combinatorial Mathematics. University of California Press, Berkeley, 1964.
- J. E. Buckley, Wage dispersion in metropolitan areas. Monthly Labor Review 92, No. 9 (September 1969). 24-29.
- Y. S. Chow, H. Robbins and D. Siegmund, Great Expectations: The Theory of Optimal Stopping. Houghton Mifflin, Boston, 1971.
- P. A. David, "Fortune, Risk and the Micro-economics of Migration." Harvard Institute of Economic Research Discussion Paper No. 274, 1973.
- M. H. DeGroot, Some problems of optimal stopping. J. Roy. Stat. Soc. B. 30 (1968). 108-122.
- _____, Optimal Statistical Decisions. McGraw Hill, New York, 1970.
- P. A. Diamond, A model of price adjustment. Journal of Economic Theory 3 (1971).
- _____ and J. E. Stiglitz, "Increases in Risk and in Risk Aversion." M.I.T. Department of Economics Working Paper No. 97, 1973.
- T. S. Ferguson, A Bayesian analysis of some nonparametric problems. Annals of Statistics 1 (1973). 209-230.