LONG-TERM COMMITMENTS, DYNAMIC OPTIMIZATION, AND THE BUSINESS CYCLE

by

BEN SHALOM BERNANKE

A.B., Harvard University
(1975)

Submitted in Partial Fulfillment
of the Requirements for the
Degree of

DOCTOR OF PHILOSOPHY

At the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1979

Signature of Author.......................................
Department of Economics, May 1979

Certified by.. ....
Thesis Supervisor

Accepted by...... .....
Chairman, Department Committee

Long-term Commitments, Dynamic Optimization,
and the Business Cycle

by

Ben Shalom Bernanke

Submitted to the Department of Economics on May 14, 1979, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

The thesis consists of three loosely connected essays. Each paper is a theoretical study of some form of long-term commitment made by economic agents. The goal is to relate the derived micro-level decision models to macroeconomic phenomena, especially the business cycle.

Chapter 1 analyzes the problem of making irreversible investment decisions when there is uncertainty about the true parameters of the stochastic economy. It is shown that increased uncertainty provides an incentive to defer such investments in order to wait for new information. Uncertainty and the volatility of investment demand are connected at the aggregate level.

In Chapter 2 we look at the commitment of resources to specific sectors of the economy. It is assumed that relative sectoral productivities vary over time, and that it is costly to transfer resources between sectors. In both planning and market economy contexts, we show that dynamic considerations can make periods of unemployment and excess capacity part of an efficient growth path.

Chapter 3 studies labor contracting in an environment with capital and a quasi-fixed labor force. We argue that for exogenous reasons real labor contracts may be incomplete; i.e., unable to contain certain types of provisions. The resulting second-best contracts may lead to situations of apparent (but only apparent) labor market disequilibrium. The contracting model provides a framework for analyzing numerous sources of unemployment.

Thesis Supervisor: Stanley Fischer
Professor of Economics

Acknowledgements

Many people provided me with help and support in the completion of this thesis. Deserving of special mention is my advisor, Stanley Fischer. Professor Fischer carefully read numerous drafts of my papers, annotating them with many excellent suggestions. His advice and encouragement throughout my entire graduate education is gratefully acknowledged.

Other faculty members also gave generously of their time, reading and discussing my work. These include Professors Rudiger Dornbusch and Robert Solow (my second and third readers), as well as Professors Peter Diamond and Dale Jorgenson (Harvard).

Throughout my graduate career my fellow M.I.T. students were an invaluable source of information and ideas. I cannot list all with whom I had fruitful discussions, but I would like to mention Ralph Braid, Jeremy Bulow, and my office-mates, Lex Kelso and Maurice Obstfeld.

I would like to acknowledge the financial assistance of the National Science Foundation and the M.I.T. Department of Economics. My graduate studies could not have been undertaken without that help.

Finally, thanks are due Peter Derksen for his excellent editorial assistance and typing of this manuscript.

Dedication

This is for my parents, Philip and Edna, who sacrificed so that I could obtain the best possible education; and for my wife, Anna, who created an atmosphere of love in which hard work became easy.

Long-term Commitments, Dynamic Optimization,

and the Business Cycle

Ben Shalom Bernanke

<u>Table</u> <u>of</u> <u>Contents</u>

CHAPTER ONE

ON THE TIMING OF IRREVERSIBLE INVESTMENTS

Introduction

Economic theorists are usually willing to assume the existence of a great deal of flexibility in the economy. Factors are mobile, prices shift readily, techniques of production are changed, the capital stock is as easily decreases as increased. Observation suggests that, however reasonable the assumption of flexibility may be in the long run, it is an increasingly bad approximation as the horizon under consideration shortens. Analysis of the business cycle -- a short-to-medium term phenomenon -- needs to recognize the difficulty the economy may have in adjusting to new events. Slow or incomplete adjustment is not the antithesis of rational economic behavior, but a result of the economic necessity of making long-term commitments under incomplete information. When new information arrives, agents who have made commitments cannot react flexibly; those who have not committed themselves may wait to find out the long-term implications before they act.

With the ultimate goal of analyzing the economy's short-run response to new information, this paper studies a particular form of commitment under uncertainty: the making of durable, irreversible investments. In this class we include almost any purchase of producers' or consumers' durables, structures, or investment in human capital. Indeed, given that a real investment is durable, the qualifier "irreversible" is hardly necessary (viz., the Second Law of Thermodynamics). Once a machine tool is made, for instance, it cannot be transformed into anything very unlike a machine tool without prohibitive loss of economic value -- this is what we mean by irreversibility. An individual can sell his machine tool, but society as a whole is still committed to it; this fact is reflected in the price the seller can get. Moreover, some investments -- for example, in human capital -- are irreversible even for the individual.

The addition of the assumption of irreversibility, in combination with the assumptions of uncertainty and investible resource scarcity, has interesting implications for investment theory. Irreversibility creates an asymmetry, not usually accounted for in the theory, between the acts of investing and not-investing. If an agent invests, and new information reveals that he should not have, then he cannot undo his mistake; his loss accrues over the life of the investment. If an agent fails to invest, when he should have, he can still make up most of the loss by investing in the next period. Willingness to invest in a given period depends not only upon risk-discounted returns but on the rate of arrival of new information. When there is a high "information potential" (usually, when the environment is in a state of flux or uncertainty), a wait-and-see approach is most profitable and investment is low. When certainty about the economy is high, and there does not seem to be much to be learned by waiting, investors are relatively more willing. The timing of investment is seen to be an important part of the decision problem. It is argued that the fact of irreversibility helps explain the volatility of durable purchases over the cycle.

The organization of the paper is as follows: Section I sets up the irreversible investment problem and interprets the solution conditions. The concepts of the asymmetry in the investment decision and information potential are introduced and motivated.

Section II develops the model for the case where agents have Dirichlet priors and the underlying stochastic structure is stationary. In this example there is a natural exact measure of information potential, and we verify that it belongs in the desired capital stock equation, along with return.

Section III is a heuristic look at the case where the stochastic structure is nonstationary. It is argued that in that situation information potential may increase or decrease, potentially leading to volatility of investment demand.

An application is presented in Section IV. We consider investment in an energy-importing economy faced with an energy cartel of uncertain duration.

Section V concludes.

I. Irreversible Investment: Statement of the Problem

This section studies the T-period, stochastic decision problem of an agent who must distribute his wealth between liquid and illiquid (irreversible) assets. We employ a simple model that reduces the problem to a choice of optimal stock levels. Our goals are to motivate the idea of an asymmetry in the investment decision caused by irreversibility and to develop tools used in the later sections.

A basic assumption to be used throughout is that the agent's stock of wealth, W, is an exogenously-given, nondecreasing function of time. This assumption plays two roles: 1) it permits separation of the asset choice problem and the life-cycle savings problem; and 2) it ensures that the agent faces a less-than-perfectly elastic supply of investible resources in each period. The first of these allows great simplification but has no essential bearing on the argument. Some form of (2), however, must be assumed. If the supply of investible resources is perfectly elastic, then the fact of irreversibility does not affect the investment decision.

The model is as follows. An agent holds a given quantity W_t of liquid wealth at the beginning of period t. The agent observes the state of nature in time t, which determines the current returns to holdings (and possibly also revises the agent's priors on future states of nature). After observing the state, the agent has the option to convert all or part of his liquid holdings into one or more of k available illiquid assets. He does this with the knowledge that an illiquid asset cannot be reconverted to liquid form or to an alternate nonliquid asset. We assume a fixed rate of transformation between liquid and nonliquid assets; one unit of liquid wealth always exchanges for one unit of illiquid. All assets are perfectly divisible, durable, and available in any quantity.

Once the agent has made his portfolio choice, he receives his current return. The return takes the form of a quantity of a homogeneous, perishable consumption good. The level of return is a function of the agent's holdings of the k+1 assets and the state of nature prevailing in t. We write the aggregate return function $R(\cdot)$ as

$$(1.1) \qquad R(K_{1t},K_{2t},\ldots,K_{kt}) = r_0(W_t - \Sigma K_{it}) + r_1(K_{1t},s_t) + \ldots + r_k(K_{kt},s_t)$$

where

K_{it} = holdings of the i-th illiquid asset in t

$W_t - \sum_i K_{it}$ = holdings of liquid wealth in t

s_t = the observed state of nature in t

and the individual return functions $r_i(\cdot)$ are increasing and concave in holdings. Note that there is a return on liquid holdings, which is assumed not to depend on s_t. For simplicity of exposition, and for this section only, we make the assumption that $r_i'(0)$ equals infinity, i = 0, 1,...,k, so that the agent always wishes to hold some of each asset. After the return is received, a new period begins, with a new state of nature. The agent must make the new portfolio choice; this he does subject to the constraint that he cannot reduce his holdings of an illiquid asset and that his total holdings cannot exceed W_{t+1}. This process is repeated until the terminal period T is reached.

The agent's objective is to maximize the expected utility of his returns over the horizon. In the usual way, expected utility will be taken to be separable and concave in the quantity of the consumption good consumed in each period and state. With these assumptions, and given both the wealth and return functions, we can conveniently write the decision problem in period t as

$$(1.2) \qquad \max_{\{K_{it}\}} \sum_{\tau=t}^{T} E_t \, \beta^{\tau-t} \, U(K_{1\tau}, K_{2\tau}, \ldots, K_{k\tau}, s_\tau, \tau)$$

$$\text{subject to} \quad K_{i\tau} \geqq K_{i\tau-1} \qquad \begin{array}{l} i = 1,2,\ldots,k \\ \tau = t,t+1,\ldots,T \end{array}$$

where β is the discount factor, $K_{i\tau}$ is the holding of the i-th illiquid asset in period τ, and expectations are with respect to period t. Here we have used our knowledge of the return functions and the τ-th period wealth constraint to write the utility functions directly in terms of the holdings of illiquid assets. For values of $K_{i\tau}$ and $K_{j\tau}$ not subject to inequality constraints, we can use our previous curvature assumptions to write, for a given state of nature

$$\begin{array}{ll} (1.3a) & \partial U/\partial K_{i\tau} \geqq 0 \\ (1.3b) & \partial^2 U/\partial K_{i\tau}\partial K_{j\tau} \leqq 0 \end{array} \qquad \text{all } i,j,\tau$$

Condition (1.3b) is interpreted as follows. To increase his holdings of K_i, the agent must run down his stock of liquid wealth. If he wishes to increase K_j as well, it must be done out of liquid wealth that has been reduced and therefore has a higher marginal opportunity cost (due to the concavity of $r_0(\cdot)$). Thus an increase in one illiquid holding reduces the <u>net</u> marginal consumption return, and hence the marginal utility, of an increase in an alternate holding.

We have not said anything about how the agent forms his expectations. In the sequel we will employ some specific models. For the present let us assume that there are thought to be a finite number of possible states in each period (the set of states possibly differing from period to period); and that the agent has set up a subjective "probability tree" giving transition probabilities at each stage as a function of the history of states.

The solution technique for this type of problem is stochastic dynamic programming. We define the sequence of value functions

$$(1.4)\qquad V_t(0,s_t) = \max_{\overline{K}_t} \{U(\overline{K}_t,s_t) + \beta\sum_j p(s_{t+1}^j|s_t)\ V_{t+1}(\overline{K}_t,(s_t,s_{t+1}^j))\}$$

$$V_{t+1}(\overline{K}_t,(s_t,s_{t+1})) = \max_{\overline{K}_{t+1}\geq\overline{K}_t} \{U(\overline{K}_{t+1},s_{t+1}) + \beta\sum_j p(s_{t+2}^j|(s_t,s_{t+1}))\ V_{t+2}(\overline{K}_{t+1},(s_t,\ldots,s_{t+2}^j))\}$$

$$\vdots$$

$$V_T(\overline{K}_{T-1},(s_t,\ldots,s_T)) = \max_{\overline{K}_T\geq\overline{K}_{T-1}} U(\overline{K}_T,s_T)$$

where $\overline{K}$ is the k-vector of illiquid asset holdings in period τ.

Thus, $V_\tau(\overline{K}_{\tau-1},(s_t,\ldots,s_\tau))$ represents the maximum expected utility attainable from periods τ to T, given inherited illiquid stocks $\overline{K}_{\tau-1}$ and history $(s_t,\ldots,s_\tau)$. For a fully specified problem the V_τ can be evaluated by backward recursion over all possible sequences $(s_\tau,s_{\tau+1},\ldots,s_T)$. Optimal asset holdings for a given time-state node and inherited holdings are found by solving for $\overline{K}^*$ satisfying the k conditions

$$(1.5)\qquad \frac{\partial U}{\partial K_{i\tau}}(\overline{K}^*_\tau,s_\tau)+\beta\sum_j p(s^j_{\tau+1}|s_\tau)\cdot$$

$$\frac{\partial V}{\partial K_{i\tau}}(\overline{K}^*_\tau,s^j_{\tau+1})-\lambda_{i\tau}=0$$

$$i=1,2,\ldots,k$$

where $\lambda_{i\tau}$ is the (positive) Lagrange multiplier corresponding to the constraint $K_i \geqq K_{i\tau-1}$. For this to be a maximum it is sufficient to demonstrate that V(·) is concave in $\overline{K}$. This is shown in an appendix.

In principle, at least, a computer could use the above approach to produce numerical solutions of fully-specified problems. In practice, the "curse of dimensionality" would prohibit the solution of large problems, especially if there were many possible states of nature in each period. Our problem is still too general for explicit solution of either the numerical or analytic variety. However we can, at this stage, use the dynamic programming concept to find some characterizations of the solution.

Let us return to the decision problem in period t. There are no inherited illiquid stocks at this point; we have already assumed that the non-negativity constraints on desired stocks are not binding. Then the agent's optimal holdings of the k illiquid assets after observing s_t are given by the simultaneous solution of the k equations

$$(1.6) \qquad \frac{\partial U}{\partial K_{it}}(\overline{K}_t^*, s_t) + \beta \sum_j p(s_{t+1}^j | s_t) \cdot \frac{\partial V}{\partial K_{it}}(\overline{K}_t^*, s_{t+1}^j) = 0$$

$$i = 1, 2, \ldots, k$$

Let us interpret these conditions. The first term of the left-hand side sum in the i-th equation is the agent's net marginal utility, in the current period, from expanding his holdings of the i-th asset by an additional unit. The second term is the marginal effect of increased current holdings of K_i on the value function. Note that this term is always less than or equal to zero. This follows from the fact that the higher the inherited levels of illiquid holdings, the more restricted is the scope of portfolio choice in subsequent periods. Since the maximum over a set is always at least as large as the maximum over a subset of that set, we have the implication that $\partial V/\partial K_i \leqq 0$, always.

$\partial V/\partial K_i$ can be interpreted as (the negative of) "expected marginal regret" induced by a current-period increase in holdings of K_i. Thus the optimal holdings, as given by (1.6), equate the current marginal return of each asset and the expected marginal losses arising from the restriction of the future choice set. Note also that $\partial V/\partial K_i(\overline{K}_t^*, s_{t+1}^j)$ is strictly equal to zero for values of s_{t+1}^j in which desired holdings of K_i will exceed those planned in period t. For values of s_{t+1}^j in which K_i appears less attractive than it did in period t, the agent experiences regret ($\partial V/\partial K_i$ is strictly less than zero).

Examination of (1.6) allows us to verify some well-known conclusions about illiquid investment. First, in the complementary problem in which the k assets are perfectly liquid, optimization requires only the maximization of current-period utility; i.e., in (1.6) the second term is always equal to zero in the liquid case. Since in the illiquid case the second term is generally less than zero, we have $\partial U/\partial K_i(\overline{K}_t^*, \text{illiquid}) \geqq \partial U/\partial K_i(\overline{K}_t^*, \text{liquid})$. That is, everything else being equal, agents will hold less of an asset if it is illiquid. (This is the basic point of recent papers on environmental preservation and irreversibility. See Arrow-Fisher (1974)). Second, all else equal, the more agents discount the future, the more willing they are to hold illiquid assets. (This does not follow directly from (1.6), since the discount factor implicitly appears also in $V(\cdot)$; however, it can be shown by induction.) Finally, the higher their prior probability on the occurrence of future states in which they will regret their illiquid investments (i.e., those future states in which desired illiquid stocks are less than those currently planned), the less the agents will invest in illiquid stocks.

We can also use (1.6) to demonstrate an asymmetry between the acts of investing and not-investing which occurs because of irreversibility. The regret terms in that first-order condition, $\partial V/\partial K_i(\bar{K}_t^*, s_{t+1}^j)$, are nonzero only for states s_{t+1}^j which are bad for holding K_i, relative to the decision period. (In 'bad' states s_{t+1}^j the optimal unrestricted holding of K_i is strictly less than in the decision period.) 'Good' states s_{t+1}^j, no matter how good, exert no counterbalancing effect in the current investment decision. This reflects the fact that underinvestment is remediable under uncertainty;

errors in this direction can be made up as soon as new information is received. Overinvestment is not remediable and induces permanent regret.

Consider this thought experiment. Suppose that, from an initial equilibrium position, the agent suddenly decided that those period t+1 states which he had thought were good for investment K_i are really (much) 'better'; but that those period t+1 states which were previously thought to be bad are actually (a little) 'worse' (i.e., they induce more regret at the original level of investment in K_i). This change of beliefs could be done in such a way that the expected value of the holding of K_i ('q', we might call it) is the same or even higher than originally. Nevertheless, due to (1.6) and the asymmetry, with the new set of beliefs the optimizing agent unambiguously reduces planned investment in K_i, relative to the original holding.

This thought experiment goes through even when we drop the assumption that investments are immediately realized as capital. The case of a nonzero gestation period is treated in Appendix 2.

In this example there has been an increase in what we shall call the system's "information potential", which we shall define heuristically as the average expected impact of new observations on the agent's beliefs.* The example suggests that when information potential is high, there is an incentive for investors to wait for the new information. This leads to a decrease in current investment. This idea is developed in the subsequent sections, under the assumption that beliefs can be summarized by a particularly convenient Bayesian distribution.

*This heuristic definition cannot always be made precise. For the example in Section II we find a natural exact measure of information potential. In Sections III and IV this concept's role is basically expositional.

II. A Dirichlet Example: Stationary Case

In this and the next section we develop an extended example that illustrates our model of investment. This section considers the case where the underlying structure which generates observed returns, though not perfectly known, is thought to be stationary over time. Thus the investor's information about his environment can increase but never decrease.

Suppose that there are a finite number of discrete states of nature possible, and that the probability of a state occurring in a given period is constant and independent of the history of states. To have perfect knowledge of the (stationary) underlying structure in this case is to know the parameters of the multinomial distribution from which the state-outcomes are drawn in each period. We shall assume that the true distribution is not known, but that the agent has a prior distribution over the multinomial parameters.

We shall take the agent's priors to be in the form of a Dirichlet distribution.* The Dirichlet is an n-parameter distribution defined over the (n-1)-simplex,

---

*For a derivation of the Dirichlet's properties, the reader should consult DeGroot (1970) or Murphy (1965). For an interesting application of this family in the theory of search, see Rothschild (1974).

where n is the number of variables in the joint distribution; i.e., it is defined only for sets of n random variables that are positive and sum to one. Thus it is an appropriate prior over the parameters of an n-nomial density.

The Dirichlet has a number of useful properties, notably that is its own posterior density and that it is statistically consistent as an estimator for the true density. We employ it here because its use for inference implies a very simple belief-updating rule. The beliefs about the environment of an agent with a Dirichlet prior can be described at time t by an n-vector, $(a_{1t}, a_{2t}, \ldots, a_{nt})$ corresponding to the parameters of his prior. Define $r_t = \sum_{i=1}^{n} a_{it}$. Then 1) the agent's prior probability (at t) on the occurrence of state j is given by a_{jt}/r_t. 2) The posterior probability is given by $\dfrac{a_{jt} + h_j(t,t+d)}{r_t + d}$, where $h_j(t,t+d)$ is the number of times state j is observed in the interval (t,t+d). To restate the belief-updating rule simply: when a new state is observed, increase the parameter corresponding to that state by one. Leave the other parameters unchanged. The updated probability of a given state is just the ratio of the updated parameter corresponding to that state to the sum of parameters.

Notice that r_t, the sum of Dirichlet parameters in time t, is a natural (inverse) measure of the information potential of the environment. When r_t is small, the effect of a new observation on the agent's priors is large. When r_t is large (infinity, in the limit), the effect of a new observation on priors is relatively small (zero, in the limit). Thus we will be especially interested to see how changes in r, return probabilities held constant, affect investment behavior.

The example we develop is the simplest possible. We consider a fixed-wealth investor who can choose between only two assets, one liquid and one illiquid (irreversible). As before, there is a T-period horizon, but now there are only two states. In state 1, the marginal returns to the illiquid asset, given holdings, are "high"; in state 2 the returns to the illiquid asset are "low". Given holdings, the return to the liquid asset is the same in both states.

The agent has Dirichlet priors on the underlying binomial distribution. His beliefs at any time t are completely characterized by the pair (a_t, r_t). The agent's prior probability that state 1 will occur in t+1 is given by a_t/r_t. If state 1 does occur in t+1, the revised priors are $(a_{t+1} = a_t + 1;\ r_{t+1} = r_t + 1)$. Similarly, the agent's probability for state 2 occurring in t+1 is $1 - a_t/r_t =$

$(r_t - a_t)/r_t$. If state 2 does occur in t+1, priors are revised by $(a_{t+1} = a_t;\ r_{t+1} = r_t + 1)$.

With this setup we can show the following proposition:

Proposition: Consider the problem of choosing an (unrestricted) portfolio $(K_t, W-K_t)$ to maximize expected utility $\sum_{\tau=t}^{T} \beta^{\tau-t}\, EU(K_\tau, s_\tau)$, $K_\tau \geqq K_{\tau-1}$, where $\partial U/\partial K(K,s=1) > \partial U/\partial K(K,s=2)$, and the agent's prior at time t is Dirichlet with parameters $(a_t, r_t - a_t)$. Define $x_t = a_t/r_t$. Then, letting $T \to \infty$, there exists a rule for desired (unrestricted) illiquid holdings of the form $K^* = K^*(x,r)$, such that $\partial K^*/\partial x > 0$, and $\partial K^*/\partial r > 0$.

Proof: (The proof is expository. The reader not interested in details may still wish to read part 1).

The existence of the rule is tantamount to the existence of a solution to the dynamic programming problem. This existence must usually be assumed, and we do so here. Conditional on existence we prove the two derivative properties. This is done by induction.

1) We show the derivative properties of the rule for period T-2. (In T the decision is trivial; in T-1 the second property holds with equality rather than inequality.) Begin by defining two quantities, K_{max} and K_{min}.

(2.1) $K_{max} = \max(W, \bar{K})$, where $\bar{K}$ is such that

$$\frac{\partial U}{\partial K}(\overline{K}, s=1) = 0$$

(2.2) $$K_{min} = \min(0, \underline{K}), \text{ where } \underline{K} \text{ is such that}$$

$$\frac{\partial U}{\partial K}(\underline{K}, s=2) = 0$$

K_{max} and K_{min} are respectively the largest and smallest quantities of the illiquid asset the agent could feasibly and rationally choose. K_{max} will be held when state 1 is thought to prevail forever. K_{min} will be held, or at least desired, whenever state 2 occurs.

We want a defining expression for K^*_{T-2}, the optimal <u>unrestricted</u> holding in T-2. (The decision maker will then actually hold $\max(K^*_{T-2}, K_{T-3})$ in T-2). If $s_{T-2} = 2$ (the "bad" state), $K^*_{T-2} = K_{min}$. So we only consider $s_{T-2} = 1$. From the point of view of T-2, the future then looks like Figure 1, where (a,r) are the Dirichlet parameters in T-2.

The level parts of the figure describe the two possible future states in T-1 (denumerated 1 and 4), and the four possible states in T (2,3,5,6), plus the corresponding desired (unrestricted) holdings. Note that in any situation where s = 2 (states 3,4,6), the desired holding is K_{min}; there is no desire to hold more, since $\partial U/\partial K(K_{min}, s=2) = 0$, and no desire to hold less, since a

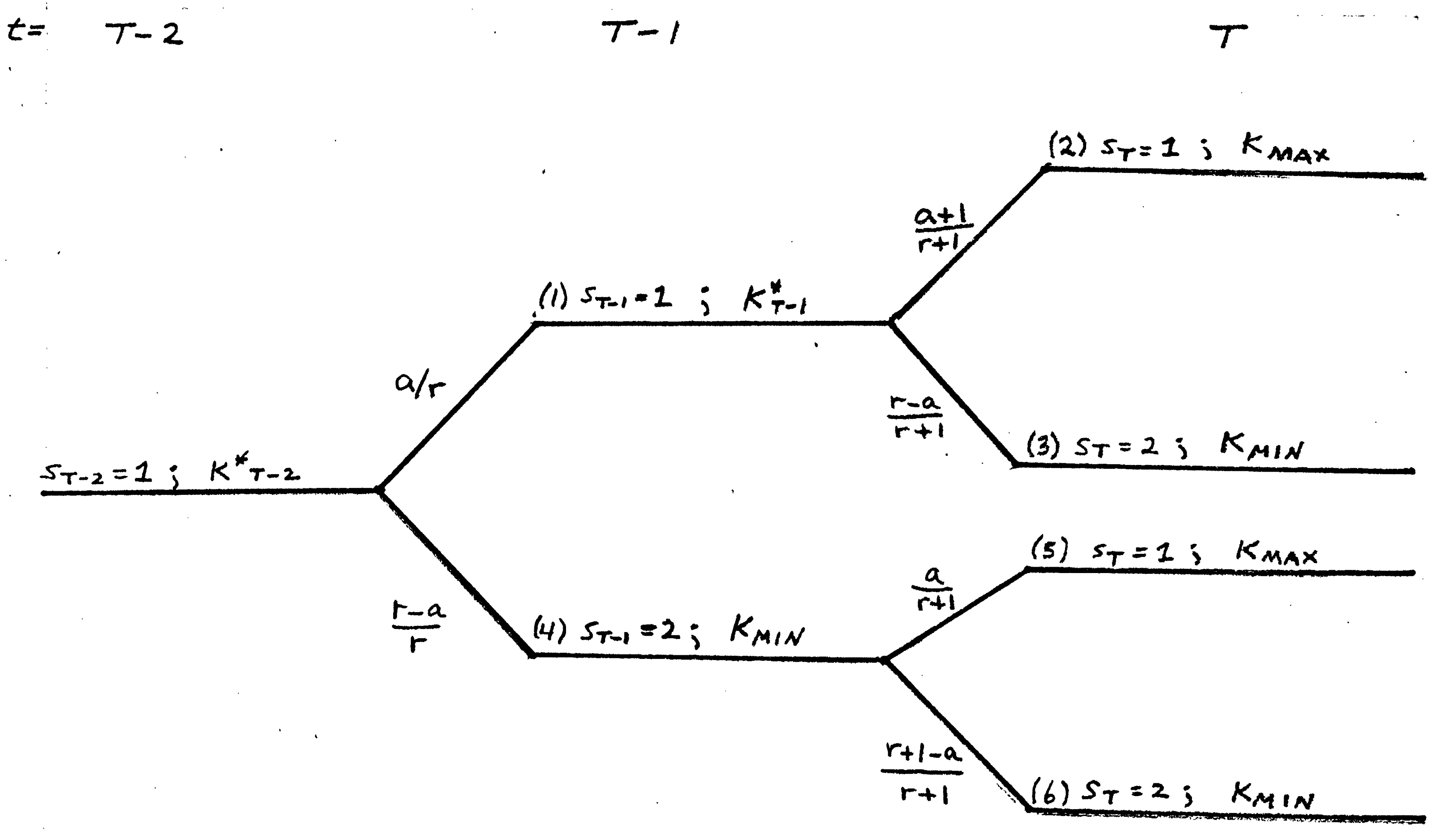

t= T-2 T-1 T
(2) $s_T = 1$; K_{MAX}
$\frac{a+1}{r+1}$
(1) $s_{T-1} = 1$; K^*_{T-1}
a/r
$\frac{r-a}{r+1}$
(3) $s_T = 2$; K_{MIN}
$s_{T-2} = 1$; K^*_{T-2}
(5) $s_T = 1$; K_{MAX}
$\frac{a}{r+1}$
$\frac{r-a}{r}$
(4) $s_{T-1} = 2$; K_{MIN}
$\frac{r+1-a}{r+1}$
(6) $s_T = 2$; K_{MIN}

Figure 1

holding of K_{min} will never restrict future decisions. When $s_T = 1$, (states 2 and 5) the desired holding is K_{max}; again, $\partial U/\partial K(K_{max}, s=1) = 0$, and, as T is the last period, there are no future decisions to worry about. Finally, K^*_{T-1} in state 1 is given by the equation

$$(2.3) \qquad \frac{\partial U}{\partial K}\,(K^*_{T-1}, s=1) + \beta\,\frac{r - a}{r + 1}\,\frac{\partial U}{\partial K}\,(K^*_{T-1}, s=2) = 0 \;.$$

Written along the sloped parts of the figures are the subjective transition probabilities, in terms of the Dirichlet parameters from T-2.

K^*_{T-2} is to be found at the point where the marginal current gains from increasing K_{T-2} are just offset by expected losses due to the restriction of future choice sets. We must determine the future states in which a small change in K_{T-2} around its optimum will constrain choice. We show first that $K^*_{T-1} \geqq K^*_{T-2}$, always. Both have the same current return schedule, and they impose identical restrictions when an unfavorable state (s = 2) occurs in the subsequent period. But 1) as the figure shows, the probability of a subsequent unfavorable state is greater when picking K_{T-2} than when picking $(K_{T-1} | s_{T-1}=1)$, and 2) K_{T-2} may impose restrictions in other subsequent states, while $(K_{T-1} | s_{T-1}=1)$ obviously does not. We conclude $K^*_{T-1} \geqq K^*_{T-2}$. Since $K^*_{T-1} \geqq K^*_{T-2}$, small changes in K_{T-2} around its optimum can cause no restriction in any state in the upper branch of the figure -- states 1, 2, or 3. K_{T-2} obviously also causes no restriction in state 5, where the desired holding is K_{max}. Increases in K_{T-2} are subsequently costly only in states 4 and 6, which have desired holdings equal to K_{min}. The probability of state 4 is (r-a)/r; the probability of state 6 is $(\frac{r-a}{r})(\frac{r+1-a}{r+1})$. K^*_{T-2} is thus given by:

$$(2.4) \qquad \frac{\partial U}{\partial K}(K^*_{T-2}, s+1) + \beta\left(\frac{r-a}{r}\right)\left(1 + \beta\left(\frac{r+1-a}{r+1}\right)\right) \cdot$$

$$\frac{\partial U}{\partial K}(K^*_{T-2}, s=2) = 0$$

(So we have been able to write down an explicit rule for choice of K^* in period T-2).

Set x = a/r and use (2.4) to define $K^*_{T-2}(x,r)$. Implicit differentiation yields

$$(2.5) \qquad \frac{\partial K^*_{T-2}}{\partial x} = \frac{d_1 \frac{\partial U}{\partial K}(K^*_{T-2}, s=2)}{\frac{\partial^2 U}{\partial K^2}(K^*_{T-2}, s=1) + d_2 \frac{\partial^2 U}{\partial K^2}(K^*_{T-2}, s=2)}$$

which is greater than zero because

$$d_1 = \beta\left(1+\beta\left(1-x\frac{r}{r+1}\right)\right) + \beta^2(1-x)\frac{r}{r+1} > 0$$

$$d_2 = \beta(1-x)\left(1+\beta\left(1-x\frac{r}{r+1}\right)\right) > 0 .$$

$\partial U/\partial K(K^*_{T-2}, s=2) < 0$ for $K^*_{T-2} > K_{min}$, and $\partial^2 U/\partial K^2 < 0$. We also get

$$(2.6) \qquad \frac{\partial K^*_{T-2}}{\partial r} = \frac{d_3 \frac{\partial U}{\partial K} (K^*_{T-2}, s=2)}{\frac{\partial^2 U}{\partial K^2} (K^*_{T-2}, s=1) + d_2 \frac{\partial^2 U}{\partial K^2} (K^*_{T-2}, s=2)}$$

which is also greater than zero because

$$d_3 = \beta^2 (1-x) \; x \; (r+1)^{-2} > 0 \; .$$

This shows the rule for T-2. Note that the property $\frac{\partial K^*_{T-2}}{\partial r} > 0$ is true because 1) the asymmetry between investing and not-investing means agents worry only about how bad things can get, not how good things can get, and 2) when r is low, things can get bad faster than when r is high. (They can also get good faster, but this is irrelevant.)

2) We now make the inductive step. Assume the existence of the rule, with properties $\partial K^*/\partial x > 0$, $\partial K^*/\partial r > 0$, for periods t+1, t+2,...,T.

Define the value function in period τ as $V(K, x_\tau, r_\tau)$, where K is inherited irreversible capital and (x_τ, r_τ) summarizes current beliefs.

In period t we suppose $s_t = 1$, as usual. Then

$$(2.7) \qquad \frac{\partial U}{\partial K} (K^*(x_t, r_t), s_t=1) + \beta \frac{\partial V}{\partial K} (K^*(x_t, r_t), x_t, r_t)$$

$$= 0 \; .$$

Differentiating by x_t and then by r_t yields

$$(2.8)\qquad \frac{\partial K^*}{\partial r} = \frac{-\beta \dfrac{\partial^2 V}{\partial K \partial x}}{\dfrac{\partial^2 U}{\partial K^2} + \beta \dfrac{\partial^2 V}{\partial K^2}}$$

$$(2.9)\qquad \frac{\partial K^*}{\partial r} = \frac{-\beta \dfrac{\partial^2 V}{\partial K \partial r}}{\dfrac{\partial^2 U}{\partial K^2} + \beta \dfrac{\partial^2 V}{\partial K^2}}$$

Our problem is reduced to showing $\partial^2 V/\partial K \partial x > 0$ and $\partial^2 V/\partial K \partial r > 0$ at the optimum points. We offer a heuristic demonstration (which may be formalized), and then an algebraic one.

Recall that $-\partial V/\partial K(K,x,r)$ measures the marginal regret induced by an increase in current irreversible holdings K. To show $\partial^2 V/\partial K \partial x > 0$ and $\partial^2 V/\partial K \partial r > 0$ we need to show that a small increase in either x or r in the agent's priors reduces the expected regret associated with a given holding K. For a given (x,r), consider the set of future consequences S(K) in which the agent will be constrained away from his optimal holdings because of the irreversibility of K. Let $\bar{s}$ be a sequence $(s_{t+1}, s_{t+2}, \ldots, s_\tau)$. By definition, $K^*_\tau(s_\tau) < K^*_t$, all $\bar{s}_\tau \varepsilon S(K)$. Also, for any $\bar{s}_\tau \varepsilon\ S(K)$, $h_1(t+1,\tau)/(\tau-t) < x$. (This follows from the inductive hypothesis $\partial K^*/\partial x > 0$. Since $K^*_\tau(\bar{s}_\tau) < K^*_t$, the fraction of good states observed between t+1 and τ must be less than x.)

A small increase in x or r could affect the regret level in three ways: by changing the set S(K), by changing the current loss due to constraint experienced in each sequence in S(K), and by changing the prior probabilities on those unfavorable sequences. 1) By the inductive hypotheses and the definition of S(K), a small increase in x or r can remove sequences from S(K) but cannot add any. 2) The loss due to constraint in each sequence in S(K) depends only on K and the states in the sequence, and therefore is unaffected by changes in x and r. 3) Define $x_\tau(\bar{s}_\tau)$ to be the agent's prior probability of a good state occurring, given history $\bar{s}_\tau \in S(K)$. Then $x_\tau(\bar{s}_\tau) = (xr + h_1(t+1,\tau))/(r + \tau - t)$. It is easy to show that $\partial x_\tau/\partial x > 0$, and $\partial x_\tau/\partial r > 0$. Thus a small increase in x or r always increases the probability of a more favorable sequence relative to a less favorable one. We conclude that increases in x or r reduce marginal regret; that is, $\partial^2 V/\partial K \partial x > 0$, $\partial^2 V/\partial K \partial r > 0$.

Algebraically we can in fact show that $\partial^2 V/\partial K \partial x$ and $\partial^2 V/\partial K \partial r$ are greater than zero not just for K around the optimum, but for any K such that $K_{min} < K < K_{max}$. Consider the quantity $\partial^2 V/\partial K \partial x$ in period T-2. Our example in part 1 of this proof already has shown $\partial^2 V/\partial K \partial x > 0$ for K such that $K_{min} < K < K^*_{T-1}$. If K is such that

$K^*_{T-1} < K < K_{max}$, we can calculate

$$(2.10) \qquad \frac{\partial V}{\partial K} = x \frac{\partial U}{\partial K}(K, s=1) + 2(1-x) \frac{\partial U}{\partial K}(K, s=2)$$

so that

$$(2.11) \qquad \frac{\partial^2 V}{\partial K \partial x} = \frac{\partial U}{\partial K}(K, s=1) - 2 \frac{\partial U}{\partial K}(K, s=2).$$

Since $K > K_{min}$, $\partial U/\partial K(K, s=2)$ is less than zero and the expression for $\partial^2 V/\partial K \partial x$ is greater than zero.

We proceed inductively. In period t, if K is such that $K_{min} < K < K^*(s_{t+1}=1)$, then

$$(2.12) \qquad \frac{\partial V}{\partial K} = (1-x) \frac{\partial U}{\partial K}(K, s_{t+1}=2) + \beta \frac{\partial V}{\partial K}(K, \frac{x\ r}{r+1}, r+1)$$

$$\text{and} \quad \frac{\partial^2 V}{\partial K \partial x} > 0.$$

If $K^*(s_{t+1}=1) < K < K_{max}$, we have

$$(2.13) \qquad \frac{\partial V}{\partial K}(K, x, r) = x\Big(\frac{\partial U}{\partial K}(K, s_{t+1}=1) + \beta \frac{\partial V}{\partial K}(K, \frac{xr+1}{r+1}, r+1)\Big) + (1-x)\Big(\frac{\partial U}{\partial K}(K, s_{t+1}=2) + \beta \frac{\partial V}{\partial K}(K, \frac{x\ r}{r+1}, r+1)\Big)$$

and again $\frac{\partial^2 V}{\partial K \partial x} > 0$. We conclude $\frac{\partial^2 V}{\partial K \partial x} > 0$ for $K_{min} < K < K_{max}$.

A proof of similar form works for $\frac{\partial^2 V}{\partial K \partial r}$. In the interests of space we only describe the calculation. Consider T-2. Part 1 showed $\frac{\partial^2 V}{\partial K \partial r} > 0$ for $K_{min} < K < K^*_{T-1}$.

Differentiating (2.10) with respect to r shows $\frac{\partial^2 V}{\partial K \partial r} = 0$ for $K^*_{T-1} < K < K_{max}$. Proceeding inductively, differentiating (2.12) shows $\frac{\partial^2 V}{\partial K \partial r} > 0$ for $K_{min} < K < K^*(s_{t+1}=1)$. (Note that the appearance of $\frac{\partial^2 V}{\partial K \partial x}$ in the expression for $\frac{\partial^2 V}{\partial K \partial r}$ makes it strictly positive.) For K such that $K^*(s_{t+1}=1) < K < K_{max}$, differentiate (2.13) with respect to r. The derivative of the second term, corresponding to $s_{t+1}=2$, is positive, but the derivative of the first term, corresponding to $s_{t+1}=1$, has ambiguous sign. Expand the first term into terms corresponding to $s_{t+2}=1$ and $s_{t+2}=2$. Again the derivative of the second term is positive, the first term ambiguous. Proceeding in this manner, we always find the ambiguous term is of the form $\frac{\partial}{\partial r}\left(\frac{\partial V}{\partial K}\left(K, \frac{xr + \tau - t}{r + \tau - t}, r + \tau - t\right)\right)$, corresponding to a perfect string of good states from t+1 to τ. But as $\tau \to \infty$, this term goes to zero for any $K < K_{max}$. We conclude $\frac{\partial^2 V}{\partial K \partial r} > 0$. q.e.d.

We have worked out an example in which the optimal holding of an illiquid asset is positively related not only to the subjective probability that it will bring a good return (x_t), but also to the certainty which the investor attaches to that probability (r_t). This seems rather realistic. Note that in this example, the parameter r_t can equally well be interpreted as the degree of investor certainty or as the inverse of the expected impact of new information on priors -- an "information potential" measure. r_t can <u>not</u> be interpreted, however, as a measure of riskiness. Risk has already been accounted for by the expected utility function. An investor could still consider a project to be very risky even if his r_t equalled infinity. On the other hand, even a risk-neutral investor may defer investment if his r_t is too low.

Another realistic aspect of this example is the importance of timing of investments. Timing does not enter most theories of investment; usually, agents are theorized either 1) to make an investment, or 2) not to make an investment, according to some set of criteria. The present analysis adds another option for the agent: 3) wait and get new information. Thus there is a decision about <u>when</u> as well as <u>whether</u> to invest. Section IV of this paper presents an application in which investment timing is of paramount importance.

III. A Dirichlet Example: Nonstationary Case

Our purpose in introducing irreversibility into investment theory was to help explain the volatility of durable investment over the cycle. So far, however, we have not seen much reason for investment to be unstable over time. This is remedied in this section as we drop the assumption that the distribution generating the returns to investment is stationary. We consider the behavior of an agent who, even as he learns about the true contemporaneous distribution of returns, is aware that that distribution itself may make discrete, random shifts at random intervals.

The agent's statistical decision problem varies according to whether or not he knows, independently of the observed states, when a distribution shift has occurred. It is more realistic and more interesting to assume that the agent does not know directly when a shift occurs, but must infer it. His problem -- detecting a change in the distribution of a random variable from realizations of the random variable alone -- is studied in statistics under the heading of dynamic inference.* Our plan for this section is 1) briefly to

*For a good description of dynamic inference, see Howard (1964). Our exposition of the subject relies heavily on that paper.

illustrate the method of dynamic inference, using the Dirichlet distribution as an example; 2) to contrast the nonstationary information structure to our previous case; and 3) to reconsider the irreversible investment problem in a particularly simple example.

Suppose there are n possible states of nature, $s = 1,2,\ldots,n$. One state is drawn independently each period from an underlying, imperfectly known multinomial distribution. The underlying distribution itself may change. The probability of the distribution changing in a given period is equal to a fixed number q, independent of the history of changes. When a change occurs, the old distribution is replaced by a new drawing from a Dirichlet meta-distribution with parameters $(a_1,a_2,\ldots,a_n;r=\sum_1^n a_i)$. Successive distributions are independent.*
The decision problem in time t is to infer the probability of each state in t+1, given the history of observations $s_1,s_2,\ldots,s_t$.

We define some notation. Let

$\bar{s}_t = (s_1,s_2,\ldots,s_t)$ be the vector of observations up to t

$\bar{c}_t^i = (c_1^i,c_2^i,\ldots,c_{m(i)}^i)$ be the i-th possible "change vector".

*A Markov process assumption would probably be more realistic here.

Each change vector represents a possible history of dates in which the underlying distribution changed. c_1^i is the period in which the last distribution change occurred, c_2^i the next-to-last change date, and so on. $m(i)$ is the total number of changes in change vector $\bar{c}_t^i$. All possible change vectors have to be considered because the agent cannot observe the true history of changes. We will adopt the convention that a change occurring in period τ occurs after the realization of s_τ. As before, let $h_j(t_1,t_2)$ equal the number of times state j is observed between periods t_1 and t_2, inclusive.

The agent's problem is to find

$$(3.1) \qquad P(s_{t+1}^j \mid \bar{s}_t).$$

Using the laws of conditional probability, (3.1) can be expanded as

$$(3.2) \qquad \sum_c P(s_{t+1}^j \mid \bar{s}_t, \bar{c}_t^i)\ P_j(\bar{c}_t^i \mid \bar{s}_t).$$

$P(s_{t+1}^j \mid \bar{s}_t, \bar{c}_t^i)$ is just the probability of s_{t+1}, given the history of observations and the date of the distribution changes. But if we know when the distribution last changed, we are back in the old Dirichlet situation and can write

$$(3.3) \qquad P(s_{t+1}^j \mid \bar{s}_t, \bar{c}_t^i) = \frac{a_j + h_j(c_1^i+1,t)}{r + t - c_1^i} \quad .$$

The agent must also consider $P(\overline{c}_t^i | s_t)$, the probability of a given change vector given the observations. (We will subsequently think of the $P(\overline{c}_t^i | \overline{s}_t)$'s as weights on the Dirichlet distributions defined by (3.3).) Using Bayes's Law, we write

$$(3.4) \qquad P(\overline{c}_t^i | \overline{s}_t) = \frac{P(\overline{s}_t | \overline{c}_t^i)\ P(\overline{c}_t^i)}{\sum_{c} P(\overline{s}_t | \overline{c}_t^j)\ P(\overline{c}_t^j)} \quad .$$

$P(\overline{s}_t | \overline{c}_t^i)$ is the probability of observing the actual history of states, given a particular change vector. This is found as follows: 1) Divide history into $m(i) + 1$ regimes, whose boundaries are the change dates in $\overline{c}_t^i$. 2) For each regime, use the Dirichlet priors and the states observed during that regime to calculate the posterior multinomial distribution. 3) Find $P(\overline{s}_t | \overline{c}_t^i)$ as the product over the regimes of the probability of the states actually observed in each regime, given the posterior distributions.

We need only find $P(\overline{c}_t^i)$ and we will have completely specified the appropriate way of inferring $P(s_{t+1}^j | \overline{s}_t)$. $P(\overline{c}_t^i)$ is the unconditional probability of a given change vector. This depends on $m(i)$, the number of changes. Since changes in successive periods are independent and occur with fixed probability q, $P(\overline{c}_t^i)$ is given by a binomial density where the number of "successes" equals m(i), with a probability of success equalling q. That is,

$$(3.5) \qquad P(\overline{c}_t^i) = \binom{t}{m(i)} m(i)^q \; (t-m(i))^{1-q} \; .$$

This completes the evaluation of the constituent parts of $P(\overline{s}_{t+1}^j | \overline{s}_t)$.

We are ready to contrast heuristically the evolution of information potential in this environment with that of the stationary environment of the last section.*
Information potential has been defined as the expected impact of a new observation on an agent's subjective probabilities. In the stationary Dirichlet case, information potential decreases monotonically over time, with $1/r_t$. A long-enough history of observations reduces the value of a new observation effectively to zero.

The behavior of information potential is different in the non-stationary case. As we see from (3.2) and (3.3), priors in this model are not described by a single Dirichlet distribution, but rather by a weighted sum (over all possible change vectors) of Dirichlets.

---

*It is to be understood that the rest of this section contains no formal results, other than the simple example. This discussion should be taken as a form of "intellectual venture capital" or as a description of future research.

A new observation in this model changes beliefs in two ways. First, it updates each of the Dirichlets in the weighted sum, just as new observations update the simple Dirichlet in the stationary case. Second, it changes the weights with which the individual Dirichlets count in the prior, increasing the weights of distributions that tend to predict the new observation, decreasing the weights of others.

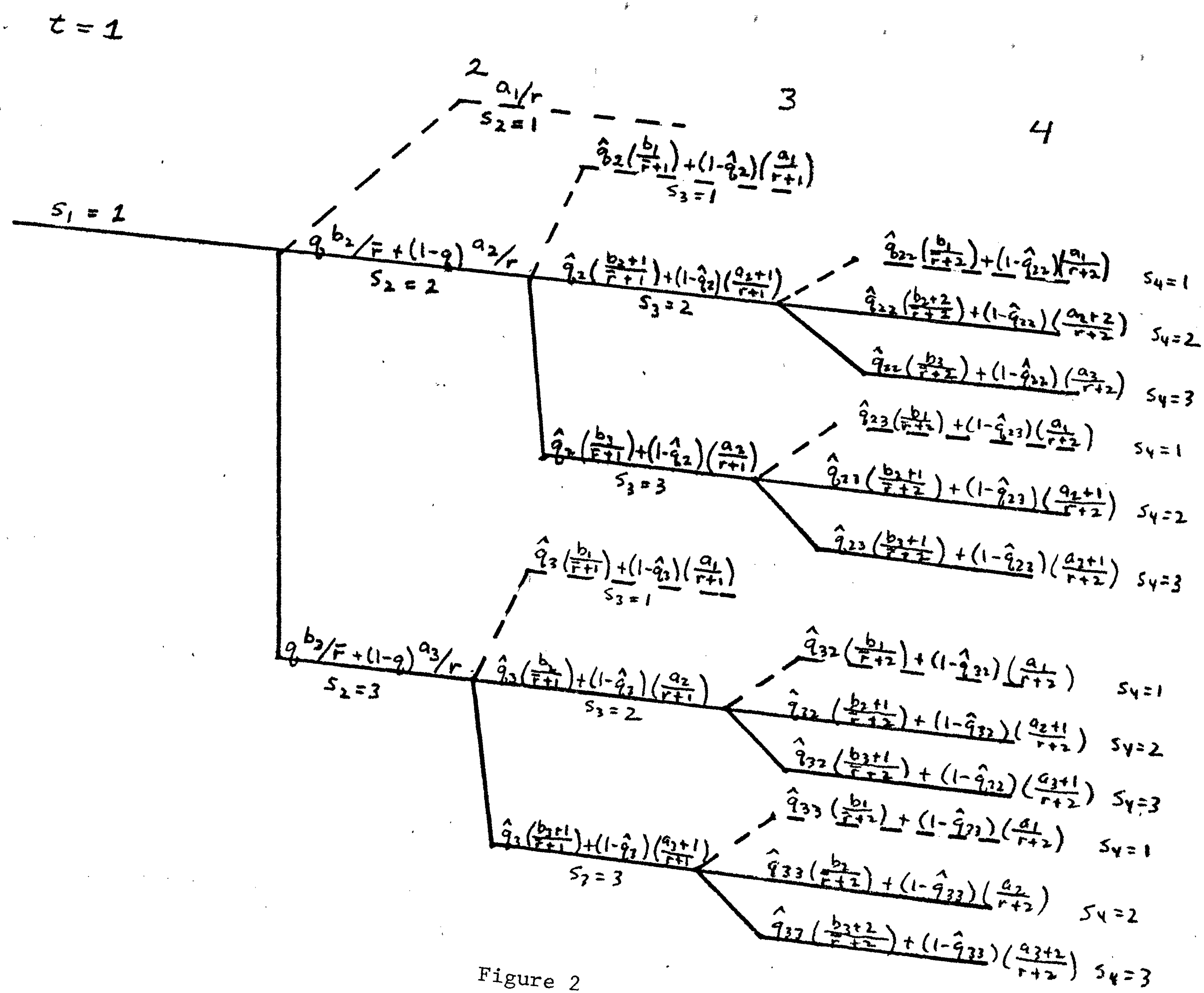
Because of the second way that new observations affect priors, the information potential at a given time in the nonstationary case may either increase or decrease. Consider this example. Suppose that q, the probability of a change, is small, and that for many periods the possible states have appeared in relatively stable proportions. Then the probabilities of change-vectors that include a recent change are low, and the highest weights in the prior are given to "old" Dirichlets with correspondingly high values of r. At this point a new observation can have little effect on beliefs. But now suppose there follow a number of "unusual" (relative to the prior) observations. This makes change-vectors which include a recent change relatively more likely, so that more weight is given to "new" low-r Dirichlets. Because, in some sense, average certainty has decreased, the information value of a new observation is larger than before.

More generally: In a nonstationary environment, new observations carry information not relevant in the stationary case -- i.e. information bearing on the probability that there has been a recent change in the underlying distribution. Unusual observations may tend to suggest that there has been a recent shift; expected observations may suggest that none has recently occurred. When the probability of a recent shift is large, new observations are important; they are given a lot of weight in the agent's attempt to tell "where he is". When the probability of a recent shift is small, less a priori value is attached to making a new observation.

The combination of irreversibility, as analyzed in the last section, and this characterization of nonstationary environments can be used as a descriptive theory of investment volatility. Willingness to undertake irreversible investment in a given period varies inversely with the amount of relevant information that can be gained by waiting. If the pattern of returns in the economy tends to be relatively stable over time, there will not be a large average premium associated with irreversible investment. If there is introduced into this usually stable environment the possibility of structural change, though the change may as likely be for good as for bad, investors will cut back to await new information. This suggests that, when capital is irreversible, planned investment in a given period can change radically, though long-run returns on average change little or not at all. It is worth noting also that the more "invested up" agents are, the more willing they are to sacrifice current returns for new information. This may explain the increasing vulnerability of an economy in a long recovery to collapses in investment demand.

It would be nice if we could present here an intuitive investment rule for the nonstationary case like the one in the last section. Unfortunately, while we can still characterize the solution in the manner of Section I, this gives us no additional insights. Unlike the stationary Dirichlet case, the nonstationary model has no sufficient statistics other than the complete history of observations. Thus we can pick no summary measure of belief corresponding to some notion like "certainty" to prove theorems about. As a second-best, we briefly present a simple example that illustrates some of the points of this section.

Figure 2 illustrates a four-period model, with three possible states in each period. In each period,



$t = 1$
2
3
4
$s_1 = 1$
a_1/r
$s_2 = 1$
$q\, b_2/\bar{r} + (1-q)\, a_2/r$
$s_2 = 2$
$q\, b_3/\bar{r} + (1-q)\, a_3/r$
$s_2 = 3$
$\hat{q}_2\left(\frac{b_1}{\bar{r}+1}\right) + (1-\hat{q}_2)\left(\frac{a_1}{r+1}\right)$
$s_3 = 1$
$\hat{q}_2\left(\frac{b_2+1}{\bar{r}+1}\right) + (1-\hat{q}_2)\left(\frac{a_2+1}{r+1}\right)$
$s_3 = 2$
$\hat{q}_2\left(\frac{b_3}{\bar{r}+1}\right) + (1-\hat{q}_2)\left(\frac{a_3}{r+1}\right)$
$s_3 = 3$
$\hat{q}_3\left(\frac{b_1}{\bar{r}+1}\right) + (1-\hat{q}_3)\left(\frac{a_1}{r+1}\right)$
$s_3 = 1$
$\hat{q}_3\left(\frac{b_2}{\bar{r}+1}\right) + (1-\hat{q}_3)\left(\frac{a_2}{r+1}\right)$
$s_3 = 2$
$\hat{q}_3\left(\frac{b_3+1}{\bar{r}+1}\right) + (1-\hat{q}_3)\left(\frac{a_3+1}{r+1}\right)$
$s_3 = 3$
$\hat{q}_{22}\left(\frac{b_1}{\bar{r}+2}\right) + (1-\hat{q}_{22})\left(\frac{a_1}{r+2}\right)$ $s_4 = 1$
$\hat{q}_{22}\left(\frac{b_2+2}{\bar{r}+2}\right) + (1-\hat{q}_{22})\left(\frac{a_2+2}{r+2}\right)$ $s_4 = 2$
$\hat{q}_{22}\left(\frac{b_3}{\bar{r}+2}\right) + (1-\hat{q}_{22})\left(\frac{a_3}{r+2}\right)$ $s_4 = 3$
$\hat{q}_{23}\left(\frac{b_1}{\bar{r}+2}\right) + (1-\hat{q}_{23})\left(\frac{a_1}{r+2}\right)$ $s_4 = 1$
$\hat{q}_{23}\left(\frac{b_2+1}{\bar{r}+2}\right) + (1-\hat{q}_{23})\left(\frac{a_2+1}{r+2}\right)$ $s_4 = 2$
$\hat{q}_{23}\left(\frac{b_3+1}{\bar{r}+2}\right) + (1-\hat{q}_{23})\left(\frac{a_3+1}{r+2}\right)$ $s_4 = 3$
$\hat{q}_{32}\left(\frac{b_1}{\bar{r}+2}\right) + (1-\hat{q}_{32})\left(\frac{a_1}{r+2}\right)$ $s_4 = 1$
$\hat{q}_{32}\left(\frac{b_2+1}{\bar{r}+2}\right) + (1-\hat{q}_{32})\left(\frac{a_2+1}{r+2}\right)$ $s_4 = 2$
$\hat{q}_{32}\left(\frac{b_3+1}{\bar{r}+2}\right) + (1-\hat{q}_{22})\left(\frac{a_3+1}{r+2}\right)$ $s_4 = 3$
$\hat{q}_{33}\left(\frac{b_1}{\bar{r}+2}\right) + (1-\hat{q}_{33})\left(\frac{a_1}{r+2}\right)$ $s_4 = 1$
$\hat{q}_{33}\left(\frac{b_2}{\bar{r}+2}\right) + (1-\hat{q}_{33})\left(\frac{a_2}{r+2}\right)$ $s_4 = 2$
$\hat{q}_{33}\left(\frac{b_3+2}{\bar{r}+2}\right) + (1-\hat{q}_{33})\left(\frac{a_3+2}{r+2}\right)$ $s_4 = 3$

Figure 2

state 1 is the "high-return" state for the irreversible capital good; states 2 and 3 have identical "low returns". State 1 prevails in period 1, the decision period. The agent puts probability q on the event that a distribution change has occurred in period 1. He puts probability zero on a distribution change in any other period. (This last is the key simplifying assumption in this example.) If the agent knew for sure that there had been no change of distribution, we assume he would have Dirichlet priors with parameters $(a_1, a_2, a_3; r = \sum_i a_i)$. If he knew for sure that a change <u>had</u> occurred, he would have a prior with parameters $(b_1, b_2, b_3; \bar{r} = \sum_i b_i)$. Let us assume that a change does not affect expected returns but does reduce certainty; i.e. $a_1/r = b_1/\bar{r}$ and $r > \bar{r}$. We also simplify our problem by assuming $a_1/r \leqq 1/2$. We would like to know the relation of the parameters of the problem to the optimal investment decision.

First let us see how the agent's priors evolve. Transition properties <u>into</u> each state are shown in Figure 2. These transition probabilities are indeed in the form of weighted sums of the Dirichlet formula. The weights correspond to his posterior probabilities on the change having occurred, given the new observations. q_i is the agent's probability that the change has occurred, given an observation of state i in period 2. q_{ij} is this probability after observation of state i in period 2 and state j in period 3. These weights are defined by

$$(3.5) \qquad q_i = \frac{q(b_i+1)/(\bar{r}+1)}{q(b_i+1)/(\bar{r}+1) + (1-q)(a_i+1)/(r+1)}$$

$$(3.6) \qquad q_{ij} = \frac{q(b_i+h_i(2,3))(b_j+h_j(2,3))/(\bar{r}+2)^2}{\text{DENOM}}$$

where $\text{DENOM} = q(b_i+h_i(2,3))(b_j+h_j(2,3))/(\bar{r}+2)^2 + (1-q)(a_i+h_i(2,3))(a_j+h_j(2,3))/(\bar{r}+2)^2$.

Let us suppose, for example, that $b_2/\bar{r} > a_2/r$, $a_3/r > b_3/\bar{r}$, so that the relative likelihood of state 2 versus state 3 increases if a distribution change occurs. Then one can verify, using (3.5) and (3.6), that observations of state 2 increase the probability that a change has occurred, while observations of state 3 decrease this probability. Thus the agent is not indifferent between observing state 2 and state 3; even though they both imply the same current return, they have different information content. Contrast this to the stationary case, in which there would be no point in distinguishing state 2 from state 3.

The determination of the optimal holding of K in period 1 is along the same lines as our example in Section II. We begin by finding the future states in which a small increase in K_1 around its optimum imposes effective restrictions on choice. These are the states marked with a solid line in Figure 2; in each of these states the desired holding is K_{min}. In states marked with a dashed line, K_1 does not restrict holdings; this is seen by symmetry arguments like the one in Section II and requires our simplifying assumption that $a_1/r \leqq 1/2$. The first-order condition for K_1 is of the form

$$(3.7) \qquad \frac{\partial U}{\partial K}(K_1^*, s_1=1) + d\,\frac{\partial U}{\partial K}(K_1^*, s_1=2,3) = 0 \ .$$

where d has fourteen terms, corresponding to the probabilities of the fourteen future states in which the desired holding is K_{min}. Inspection or tedious differentiation gives the following results about K_1^*:

1) An increase in a_1 or b_1, holding other parameters constant, increases K_1^*. An increase in r (a_1/r constant) or $\bar{r}$ ($b_1/\bar{r}$ constant) also increases K_1^*.

2) An increase in the prior change probability q unambiguously reduces K_1^*, despite the fact that the probability of a good state is unaffected by a distribution change.

3) Given that $a_1/r = b_1/\bar{r}$, K_1^* is at its maximum when we also have $a_2/r = b_2/\bar{r}$ and $a_3/r = b_3/\bar{r}$. When

these equalities hold, the information which an observation of state 2 or 3 contains about whether a change has occurred is minimal. Since the information potential of the environment is relatively lower, K_1^* is relatively higher.

IV. An Application: Investment When There is an Energy Cartel

As an illustration and application of the ideas of this paper, we introduce a simple model of investment and output in an energy-importing economy after the unanticipated formation of an energy-exporter's cartel. It will be shown in this model that uncertainty can make investment collapse, even if capital dominates the alternative asset in every period.

We consider the behavior of risk-neutral agents in the energy-importing economy. At time t this economy is assumed to have two possible domestic factors of production: a stock of energy-intensive capital K_t^e, and a stock of energy-saving capital, K_t^s. Both stocks are durable and irreversible. These factors are used to produce a homogeneous good y_t according to the relation

$$(4.1) \qquad y_t = x_t^e K_t^e + x_t^s K_t^s \qquad \begin{array}{l} 0 \leqq x_t^e \leqq 1 \\ 0 \leqq x_t^s \leqq 1 \end{array}$$

where x_t^e and x_t^s are the utilization rates of K_t^e and K_t^s, respectively. Utilization rates are chosen by the agents in each period in a manner to be specified shortly. These rates are introduced basically because we want to assume that energy-using capital is used less intensively when the cartel is in existence.

Over time the agents may augment K^e and K^s from their stock of investible resources. This stock, W_t, is assumed to be an exogenously-given, increasing function of time.* The investible resources may be converted costlessly and at a one-for-one basis into units of K^e or K^s. We assume that these resources pay no return in liquid form and have no alternative uses. Conversion of investible resources to a specific form of capital is irreversible. The constraints on the choice of K_t^e and K_t^s are thus given by:

$$(4.2) \qquad K_t^e + K_t^s \leqq W_t$$

$$K_t^e \geqq K_{t-1}^e$$

$$K_t^s \geqq K_{t-1}^s \; .$$

The state of nature in each period in this model depends on the status of the energy-exporter's cartel. We define the state of nature s_t by

$$(4.3) \qquad s_t = \begin{cases} 1, & \text{if the cartel exists in period } t \\ 0, & \text{otherwise.} \end{cases}$$

*The interpretation of W_t varies with the choice of agent. For a small firm, W_t is the available line of credit. For an industry (e.g., electric power), W_t is potential plant sites, or demand markets. For a national economy, W_t is aggregate investible resources: labor, land, raw materials.

The cartel is assumed to have formed in period t_0. Agent's beliefs about its continued existence are given by

$$(4.4) \qquad \Pr(\, s_t = 1 \mid s_{t-1} = 1 \,) = p_t$$

$$\Pr(\, s_t = 1 \mid s_{t-1} = 0 \,) = 0$$

where

$$dp_t/dt > 0, \; \lim_{t\to\infty} p_t = 1 \, .$$

Thus if the cartel fails, it is assumed to be gone forever. The longer the cartel lasts, the greater is the agents' common subjective probability of its survival through the next period. If the cartel survives long enough, it is assumed to be permanent.

The agents are risk-neutral, so their goal is to maximize $E_t \sum_{\tau=t}^{\infty} \beta^{\tau-t} y_\tau$. We shall chart their optimal production and investment paths as the cartel stubbornly continues to exist.

First we specify how utilization rates (and current profitability) are determined. We assume that there are three per-unit-capital cost functions:

$$(4.5) \qquad C^s(x_t^s) = \text{cost per unit } K_t^s \text{ associated with rate } x_t^s \, .$$

$$C^{e,0}(x_t^e) = \text{cost per unit } K_t^e \text{ for rate } x_t^e$$

$$\text{and } s_t = 0$$

$$C^{e,1}(x_t^e) = \text{cost per unit } K_t^e \text{ for rate } x_t^e$$

$$\text{and } s_t = 1 \text{ .}$$

Costs are in terms of the output good y_t. Assume $C(0) = 0$, $dC/dx > 0$, $d^2C/dx^2 > 0$ for all cost functions; for a given x, take $dC^{e,0}/dx < dC^s/dx < dC^{e,1}/dx$. Thus energy-saving capital has higher marginal costs than energy-using capital when there is no cartel, but has lower costs when the cartel exists.

Net output maximization now yields optimal utilization rates $\bar{x}^s$, $\bar{x}^{e,0}$, and $\bar{x}^{e,1}$ as solutions to

$$(4.6) \qquad \begin{aligned} dC^s(\bar{x}^s)/dx^s &= 1 \\ dC^{e,0}(\bar{x}^{e,0})/dx^{e,0} &= 1 \\ dC^{e,1}(\bar{x}^{e,1})/dx^{e,1} &= 1 \text{ .} \end{aligned}$$

These utilization rates are independent of time and the sizes of the capital stocks, and satisfy the relation $0 < \bar{x}^{e,1} < \bar{x}^s < \bar{x}^{e,0}$. With them we can define the per-period gains from a unit of capital:

$$(4.7) \qquad \pi^s = \bar{x}^s - C^s(\bar{x}^s) = \text{profit per period from each unit of } K^s$$

$$\pi^{e,0} = \bar{x}^{e,0} - C^{e,0}(\bar{x}^{e,0}) = \text{profit per}$$

period per unit of K^e when $s = 0$

$\pi^{e,1} = \bar{x}^{e,1} - C^{e,1}(\bar{x}^{e,1})$ = profit per period per unit of K^e when $s = 1$.

Note that the relation $0 < \pi^{e,1} < \pi^s < \pi^{e,0}$ holds.

We now look at the evolution of the capital stock when the cartel refuses to disappear. We note that, since 1) investors are risk-neutral, 2) investment in either K^e or K^s is always guaranteed a positive return, 3) uninvested resources pay no return, and 4) investment is free up to the resource constraint W_t, capital appears to dominate the alternative in the traditional risk-return sense. Thus it may appear that investment will never be below its maximum, equal to $W_t - (K^e_{t-1} + K^s_{t-1})$. This turns out not to be true: it is possible in this model to have an investment 'pause', during which even risk-neutral investors are content to cumulate barren liquid resources and wait for new information.

The analysis that follows is aimed at finding sufficient conditions for this pause to occur. We begin by defining $V_t(K^e_{t-1}, K^s_{t-1}, s_t)$ (in the manner of Section I) as the maximum expected discounted consumption available from period t to the horizon, given inherited stocks K^e_{t-1}, K^s_{t-1}, and current state of nature s_t.

If $s_t = 0$, (i.e., the cartel has failed) then, by assumption, it is known with certainty that $s_\tau = 0$, all $\tau \geqq t$. The investors' best plan is to invest all available resources in K^e, so that for all $\tau \geqq t$ we have $K^e_\tau = W_\tau - K^s_{t-1}$. Thus we can write $V_t(K^e_{t-1}, K^s_{t-1}, s_t=0)$ explicitly as

$$(4.8) \qquad V_t(K^e_{t-1}, K^s_{t-1}, 0) = \sum_{\tau=t}^{\infty} \beta^{\tau-t}\{\pi^s K^s_{t-1} + \pi^{e,0}(W_\tau - K^s_{t-1})\} .$$

Abbreviating the expression in (4.8) by V^0_t, we can write

$$(4.9) \qquad \partial V^0_t/\partial K^e_{t-1} = 0$$

$$\partial V^0_t/\partial K^s_{t-1} = \sum_{\tau=t}^{\infty} \beta^{\tau-t}(\pi^s - \pi^{e,0})$$

$$= (\pi^s - \pi^{e,0})/(1-\beta) < 0 .$$

As predicted in Section I, "marginal regret" $(-\partial V/\partial K)$ for a given investment is either zero or positive, depending on whether the subsequent state is "good" or "bad" for that investment.

When $s_t = 1$ (the cartel is still in existence), we have

$$(4.10) \qquad V_t(K^e_{t-1}, K^s_{t-1}, s_t=1) = \max_{K^e_t, K^s_t} \pi^{e,1} K^e_t + \pi^s K^s_t +$$

$$\beta\{p_t V_{t+1}(K_t^e, K_t^s, s_{t+1}=1) +$$

$$(1-p_t)V_{t+1}(K_t^e, K_t^s, s_{t+1}=0)\}$$

subject to

$$K_t^e \geqq K_{t-1}^e$$

$$K_t^s \geqq K_{t-1}^s$$

$$W_t \geqq K_t^e + K_t^s$$

so that

$$\partial V_t^1 / \partial K_{t-1}^e = -\lambda_t^e \tag{4.11}$$

$$\partial V_t^1 / \partial K_{t-1}^s = -\lambda_t^s$$

where $\lambda_t^e \geqq 0$, $\lambda_t^s \geqq 0$ are the (constant for given t and s_t) Lagrange multipliers associated with the constraints $K_t^e \geqq K_{t-1}^e$ and $K_t^s \geqq K_{t-1}^s$, respectively.

The investor's problem in period t with $s_t = 1$ is

$$\max_{K_t^e, K_t^s} \pi^{e,1} K_t^e + \pi^s K_t^s + \tag{4.12}$$

$$\beta\{p_t V_{t+1}(K_t^e, K_t^s, s_{t+1}=1) +$$

$$(1-p_t)V_{t+1}(K_t^e, K_t^s, s_{t+1}=0)\}$$

subject to the three constraints. Using (4.9) and

(4.11), the first-order conditions are

$$(4.13a)\qquad (K^e)\qquad \pi^{e,1} - \beta p_t \lambda^e_{t+1} + \lambda^e_t - \lambda^w_t = 0$$

$$(4.13b)\qquad (K^s)\qquad \pi^s + \frac{\beta}{1-\beta}(1-p_t)(\pi^s - \pi^{e,0}) -$$

$$\beta p_t \lambda^s_{t+1} + \lambda^s_t - \lambda^w_t = 0$$

where $\lambda^w_t \geqq 0$ is the multiplier associated with the constraint $W_t \geqq K^e_t + K^s_t$. λ^w_t is strictly greater than zero when the resource constraint is binding. Since the risk-neutral investor always picks a corner solution, we can identify an investment 'pause' with periods t such that $\lambda^w_t = 0$ (or, equivalently, with periods t such that $\lambda^e_t > 0$ and $\lambda^s_t > 0$). The following proposition gives sufficient conditions for $\lambda^w_t = 0$.

Proposition. Sufficient conditions for $\lambda^w_t = 0$ in this problem are

1) $\pi^s < \alpha_1 \pi^{e,0}$ $\qquad \alpha_1 = \frac{\beta(1-p_t)}{1-\beta+\beta(1-p_t)}$

and

2) $\pi^{e,1} < \alpha_2 \pi^s$ $\qquad \alpha_2 = \frac{\beta p_t}{1-\beta+p_t}$.

Proof. Let t go to infinity, holding $s_t = 1$. By assumption $\lim_{t\to\infty} p_t = 1$. If $p_t = 1$ all investment is in K^s, so that $\lim_{t\to\infty} \lambda^w_t = \pi^s$. This implies $\lim_{t\to\infty} \lambda^s_t = 0$,

$\lim_{t\to\infty} \lambda_t^e = (\pi^s - \pi^{e,1})/(1-\beta)$. The limits of λ_t^s and λ_t^e represent their lower and upper bounds, respectively.

We want sufficient conditions for $\lambda_t^w = 0$. $\lambda_t^w = 0$ is equivalent to ($\lambda_t^e > 0$, $\lambda_t^s > 0$). By (4.13a), $\lambda_t^e > 0$ if $\pi^{e,1} - \beta p_t \lambda_{t+1}^e - \lambda_t^w < 0$. Since $\lambda_t^w \geq 0$ and λ_{t+1}^e is bounded above by $(\pi^s - \pi^{e,1})/(1-\beta)$, a sufficient condition is $\pi^{e,1} < \frac{\beta p_t}{1-\beta} (\pi^s - \pi^{e,1})$ which is equivalent to 2).
Similarly, by (4.13b), $\lambda_t^s = \lambda_t^w - (\pi^s + \frac{\beta}{1-\beta} (1-p_t)$ $(\pi^s - \pi^{e,0})) + \beta p_t \lambda_{t+1}^s$, which is unambiguously positive if 1) holds. q.e.d.

Since p_t is monotonic for s = 1, if either of these conditions is true it will be true over a continuous interval of time. The continuous interval in which the two conditions intersect will have $\lambda^w = 0$, i.e., there is no investment for any value of W_t. It is numerically plausible that this intersection will exist. Suppose $\beta = .9$, $p_t = .5$. Then $\lambda_t^w = 0$ if $\pi^{e,1} < .82\ \pi^s$ and $\pi^s < .82\ \pi^{e,0}$. Other things equal, the more disparate are the profitabilities $\pi^{e,1}$, π^s, and $\pi^{e,0}$, the more likely it is that a pause will occur, and the longer it will be if it does occur. An example: Say that we have, at time t, a set of profitabilities ($\pi^{e,1}$,π^s,

$\pi^{e,0}$) such that agents invest all available resources. Now suppose that $\pi^{e,0}$ were to be multiplied by a thousand, π^{s} by a hundred, and $\pi^{e,1}$ by ten. This huge increase in the value of capital will likely drive current investment to zero! This is because the increased value of waiting for new information more than offsets the improvement in current returns.

Under the assumption that a pause occurs and taking W_t as linear, Figure 3 traces the path of the energy-importing economy over time.

In the figure the pause runs from t_1 to t_2. (We show $t_1 > t_0$, the period of the cartel's formation; an alternative possibility is $t_1 = t_0$.) The history of the economy is as follows. From t_0 to t_1 investors give insufficient credence to the cartel to desist from energy-intensive investment. By t_1 the future has become sufficiently ambiguous that investors prefer to remain liquid and wait for new information. Finally, at t_2, the continued existence of the cartel seems sufficiently likely that investors commit themselves with a bang to energy-saving capital. There is an investment spurt as cumulated liquidity is transformed to a stock of K^s.

To put the development of the economy in terms of our heuristic "information potential" measure: In the

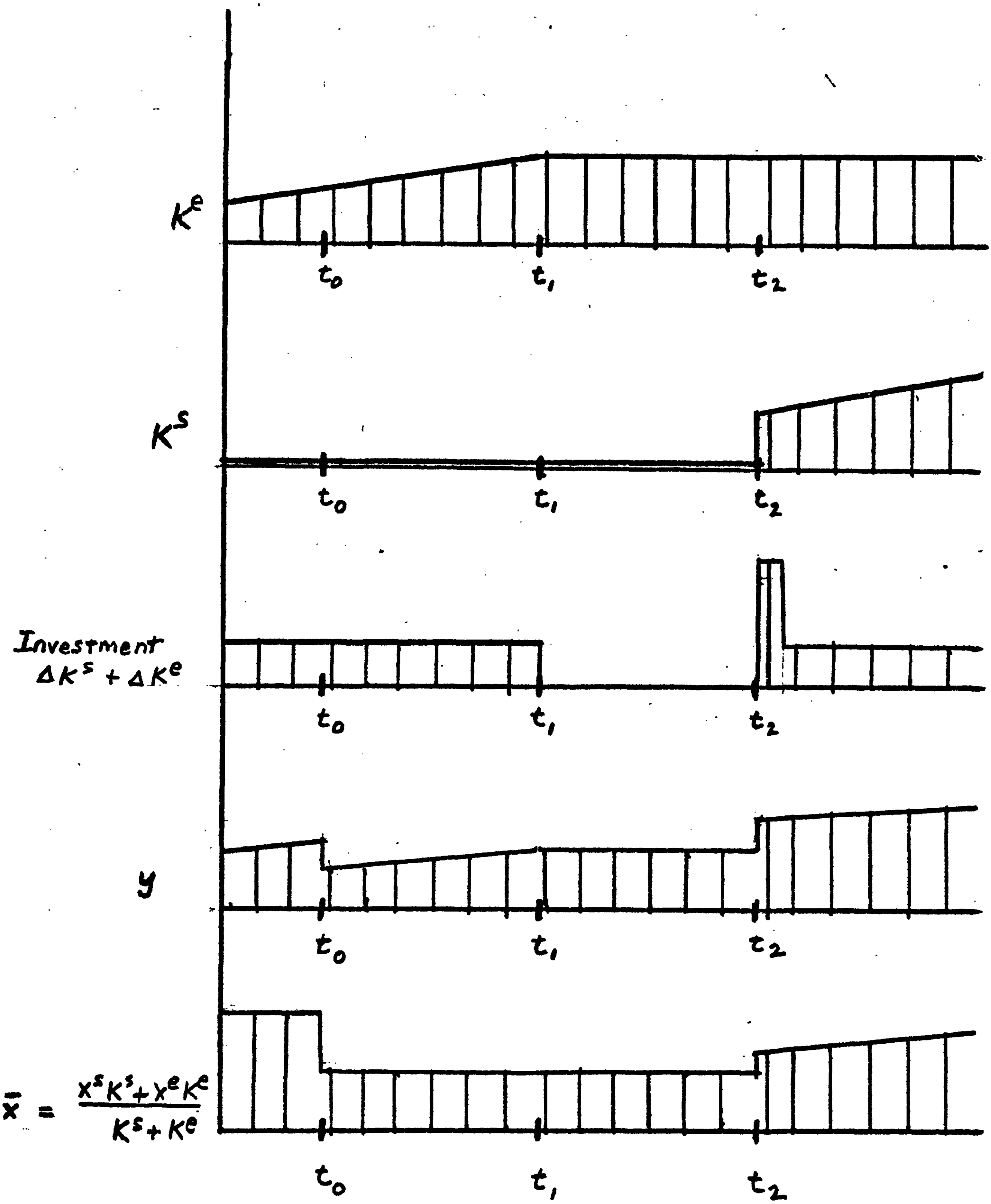

K^e
t_0
t_1
t_2
K^s
Investment
$\Delta K^s + \Delta K^e$
y
$\bar{x} = \frac{x^s K^s + x^e K^e}{K^s + K^e}$

Figure 3

intervals (t_0, t_1) and (t_2, ∞), investors feel that they know the true long-run situation with a relatively high probability. Information potential is low and investors are relatively willing. In the period (t_1, t_2), the future is more uncertain. The information value of waiting is high, and investors hold back.

We can see that, although nothing observable changes in the investors' environment after t_0, the development of the economy is not smooth. Investment is quite volatile. As the figure shows, output also dips and then rises. This output movement reflects only changes in utilization rates and the composition of capital; the variability of output would be increased if we explicitly included the production of capital goods. Aggregate capacity utilization is cyclical, reflecting the existence of the cartel and the changing composition of capital.

We did not include a labor input in the model, but it would be easy to do so. If we postulated that 1) labor supply is inelastic, 2) there are fixed costs in training a worker for a specific job, or in moving a worker from one job to another, and 3) current labor costs vary with the rate of utilization of the labor force, then the allocation of labor would parallel that of capital. There would be a pause, followed by a spurt, of new hires as entrepreneurs wait to see to which technology new workers should be committed. Meanwhile, while the cartel lasts, workers already employed in the energy-intensive sector would face low utilization rates (layoffs and short hours). If the cartel failed, these workers would experience higher rates (callbacks and overtime).

Two comments conclude our discussion of this example.

First, as an explanation of the recent recession, our model is obviously oversimplified. It does seem, however, that uncertainty has been a major reason for the weakness of investment since 1973 -- uncertainty not only about the effectiveness of OPEC, but about the long-run nature of domestic policy, the prospects of new technologies, the future of worldwide economic conditions. Caution is the order of the day for investors.

Second, we note that the "cycle" generated by this model represents a completely efficient use of resources, given technology and beliefs. The output changes are supply-induced and do not depend on demand shortfall. Thus, government interference for efficiency reasons would not be warranted in this economy, given that markets work well in accommodating output variations. This last proviso is an important one, however. It is considered in the third chapter of this thesis.

Conclusion

This paper has argued that when investment is irreversible, it will sometimes pay agents to defer commitment of scarce investible resources in order to wait for new information. Uncertainty about the long-run environment which is potentially resolvable over time thus exerts a depressing effect on current levels of investment. We have conjectured that changes in the general level of uncertainty may explain some of the volatility of investment demand associated with cyclical fluctuations.

There are numerous avenues for future research suggested by this topic.

First, the basic model should be generalized to a more realistic description of the investment decision. Some interesting extensions are:

1) The incorporation of information flows that are not purely exogenous. For example, the possibility of "learning-by-doing" induced by the investment process may create a <u>positive</u> incentive investment in some uncertain situations.

2) The removal of the "zero-one" character of irreversibility in our model. If we allow for partial convertibility of capital stock, we can analyze the

decision to commit to, say, flexible (but higher-cost) technologies versus more restrictive options.

3) The addition of flow constraints. If there are high costs associated with converting investible resources into capital at a rapid rate, the results of the model are modified. Keeping one's portfolio completely in investible resources is clearly no longer the most cautious option in this case, since the penalty for underinvestment will be greater than one period's foregone output.

Second, the relation of this model to business cycle theory must be taken beyond heuristics and put into a general equilibrium structure. An important task is to show how the central planner's solution of this paper (as in Section IV) is duplicated by a competitive economy. It should be possible to show that the aggregate decision of competitive investors looks like that of the planner, even when the investors believe that capital markets are perfect (so that the "scarcity of investible resources" assumption is violated on the micro-level). The mechanism that enforces this is speculation in the investible resources market. This speculation adds a premium to the price of investible resources analogous to the "user cost" added to the price of exhaustible resources. When uncertainty is high, a high premium in the price of investible resources depresses competitive investment.

Finally, this work has many potential microeconomic applications. An example is the problem of choosing a technique in a field where the technology is changing rapidly. Should a firm buy the current-generation computer system or speculate by waiting for a system that is better and cheaper? The decision to wait or commit in a given period depends not only on expected system improvement (return) but also on how much one can expect to learn in the short run about long-run technical possibilities (information potential).

Appendix 1

To show $V_t(K_{1t},K_{2t},\ldots,K_{kt},\overline{s}_t)$ is concave in $(K_{1t},K_{2t},\ldots,K_{kt})$.

Lemma: Let f be a concave function $\mathbb{R}^k \to \mathbb{R}^1$. Let x and y be k-vectors. Define $g(y) = \max_{x>>y} f(x)$. Then g(y) is concave in y.

Proof will show g not concave implies f not concave. g not concave implies $g(ty_1 + (1-t)y_2) < tg(y_1) + (1-t)g(y_2)$. Let $f(x_1) = \max_{x>y_1} f(x)$, $f(x_2) = \max_{x>y_2} f(x)$. Now $tx_1 + (1-t)x_2 > ty_1 + (1-t)y_2$, so $f(tx_1 + (1-t)x_2) \leqq g(ty_1 + (1-t)y_2)$, by definition of g. But $g(ty_1 + (1-t)y_2) < tg(y_1) + (1-t)g(y_2) = tf(x_1) + (1-t)f(x_2)$. By transitivity, this implies $f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2)$, which implies that f is not concave. So f concave implies g concave. //

Main result is by induction. $V_T(\overline{K},\overline{s}) = \max_{K>>\overline{K}} U(K,s_T)$. U is concave so V_T is concave. Now suppose V_{t+1} is concave. $V_t(\overline{K},\overline{s}) = \max_{K>>\overline{K}} \left(U(K,s_t) + \beta\Sigma p_i\, V_{t+1}(K,\overline{s}_{t+1}) \right)$. The sum of concave functions is concave, so by the same lemma V_t is concave. q.e.d.

Appendix 2

To show the existence of an asymmetry in the investment decision when the gestation period is nonzero.

Let the gestation period be of length g. Thus decisions about $K_{\tau+g}$ are made in period τ. Assume that investments in the pipeline are irreversible.

Analogously to (1.4), define

$$V_t(\overline{K}_{t+g-1}, s_t) = \max_{\overline{K}_{t+g} \geq \overline{K}_{t+g-1}} \beta^g \sum_h p(s_{t+g}^h | s_t) \cdot U(\overline{K}_{t+g}, s_{t+g}^h) + \beta \sum_j p(s_{t+1}^j | s_t) \cdot V_{t+1}(\overline{K}_{t+g}, s_{t+1}^j)$$

where now V_t is the maximum expected utility for (t+g,T). The first-order condition analogous to (1.6) is

$$\beta^g \sum_h p(s_{t+g}^h | s_t) \frac{\partial U}{\partial K_i} (\overline{K}_{t+g}, s_{t+g}^h) + \beta \sum_j p(s_{t+1}^j | s_t) \frac{\partial V_{t+1}}{\partial K_i} (\overline{K}_{t+g}, s_{t+1}^j) = 0 .$$

As when the gestation period was zero, $\partial V_{t+1}/\partial K_i$ is less than zero for states s_{t+1} in which the investment decision of period t is 'regretted'; i.e., $K_{i,t+g} >$

$K^{*}_{i,t+g+1}(s^{j}_{t+1})$. When the constraint is not effective, $\partial V_{t+1}/\partial K_i = 0$. The thought experiment of Section I (pp. 21-22) goes through with no difficulty. Make those states s_{t+1} for which $\partial V_{t+1}/\partial K_i < 0$ "a little worse" (increase $-\partial V_{t+1}/\partial K_i$ slightly). No improvement of the prospects of K_i in states for which $\partial V_{t+1}/\partial K_i = 0$ can prevent a reduction of investment in period t.

CHAPTER TWO

EFFICIENT EXCESS CAPACITY AND UNEMPLOYMENT IN A TWO-SECTOR ECONOMY WITH FIXED INPUT PROPORTIONS

Introduction

The macroeconomic policy goal of having the economy reach "potential GNP" in each period is based, at least in part, on the assumption that it is economically inefficient to allow capital and labor stocks to stand idle. This assumption may seem reasonable, especially if depreciation is independent of rates of production and variable costs are small. In fact, this assumption is correct only if input stocks have no "specificity"; i.e., if they are costlessly transferable between sectors or uses. If transfer is costly, then under some circumstances dynamic efficiency will require assigning input stocks to sectors where they will be temporarily idle, rather than to sectors where they could be currently productive.

In order to develop this and related results, this paper studies the class of economies with the following characteristics:

1) There is more than one productive sector.
2) Capital is durable; purchases of plant and equipment are made with the knowledge that their useful lives extend into the uncertain future. Capital is also "bolted down", i.e., sector-specific.
3) Future "investment opportunities", in a broad

sense, are perceived as uncertain.

4) There are worker mobility costs. This imparts a quality of durability to labor analogous to that of capital, in the sense that the use of labor requires initial sunk costs.

5) There is limited factor substitutability *ex ante*; at least in the short- and medium-runs, the set of possible capital-labor ratios has fixed positive bounds.

To analyze this class of economies, we will employ a simple model that embodies these assumptions in an extreme form. For example, rather than discussing economies where factor substitutability is merely bounded, we shall take the limiting case and speak mainly about fixed-coefficients technologies (in which there is no substitution at all). These restrictions, however, are largely for the purposes of exposition. Intuition should suggest that our results will hold approximately in the most general case.

This paper has two sections. Part I sets up one simple two-sector model which is to be used throughout. We consider the problem of a central planner for this economy searching for the optimal allocation of resources over time. A key result is that best allocation will sometimes require the temporary unemployment of capital, labor, or both.

In Part II we get rid of the planner and introduce a monetary market version of the model. The behavior of consumers, firms, and workers is characterized in a unified way by the introduction of a time-state discount factor derived from expected utility theory. This economy is compared with the planning version. Here also we see that there are efficiency reasons for the existence of idle resources, even resources that are not "used up" when employed in production.

I. Throughout the paper we will be considering the following economic model:

The economy is assumed to function over a horizon of T discrete periods, indexed $t = 1,\ldots,T$ with given initial conditions.

There are two sectors, consumption goods and capital goods, each of which produces a homogeneous product. Capital and labor, the only inputs, are used in fixed proportions - an extreme form of the assumption of limited factor substitutability. With normalization we can write

$$(1.1)\qquad C_t = \min(K_{ct}, L_{ct})$$

$$I_t = \min(K_{kt}/a, L_{kt}/b)$$

where C_t is consumption, I_t is investment (the output of the capital goods sector), inputs are indexed by sector and period, and a and b are constants.

We want somehow to convey the idea that future investment opportunities are uncertain. There are many ways to model this. Let us assume that the future effectiveness of investment goods in creating new capacity (or, alternatively, the relative productivity of the investment goods sector) is a random variable. Rewrite the investment goods production function as

$$(1.2)\qquad I_t = \min(K_{kt}/a, L_{kt}/b) \cdot \theta_t$$

where $\{\theta_t\}$ is a stochastic sequence. (Note that this is equivalent to retaining a nonstochastic capital goods production function and writing

$$(1.2a) \qquad K_{ct} = K_{c,t-1} + I_{c,t-1} \cdot \theta_{t-1}$$

$$K_{kt} = K_{k,t-1} + I_{k,t-1} \cdot \theta_{t-1}$$

where I_{ct} and I_{kt} are investment in each of the two sectors.)

The specification (1.2) means that some periods (when θ_t is large) are "good" for investment - i.e., a fixed amount of input in the investment goods sector produces a large investment to capacity - and other periods (small θ_t) are "bad". (The current realization of θ_t is assumed to be known when investment decisions are made.)

In this model, investment goods are homogeneous when produced, but once they are added to the capacity of either sector we shall assume that they are bolted down and cannot be transferred. Capacity is also perfectly durable. These conditions, plus the requirement that total investment may not exceed the output of the capital goods sector, may be summarized as

$$(1.3) \qquad K_{ct} = K_{c,t-1} + I_{c,t-1} \qquad\qquad I_{c,t-1} \geqq 0$$

$$K_{kt} = K_{k,t-1} + I_{k,t-1} \qquad I_{k,t-1} \geqq 0$$

$$I_{ct} + I_{kt} \leqq I_t$$

where, again, K_{ct} and K_{kt} are sectoral capital stocks in period t; I_{ct} and I_{kt} are investment in each of the two sectors in period t; and I_t is the output of the capital goods sector.

Labor can be assigned either to the consumption or capital goods sector, up to the current total labor pool:

$$(1.4) \qquad L_{ct} + L_{kt} \leqq L_t \ .$$

However, there are real mobility costs for labor; transferring a worker from the unassigned pool to one of the sectors, or from one sector to the other, costs v_t units of consumption goods. We will want to assume that the future effective labor force (which may include labor-augmenting technical change) is known. To remove some difficulties not of direct relevance here, we will also assume that the labor force increases monotonically.

$$(1.5) \qquad L_t \geqq L_{t-1} \ .$$

Now let us imagine that this economy is centrally planned. The planner has perfect information about the state of the economy in the current period, t, and has complete control of allocations. He also knows, or thinks he knows, the stochastic process that is generating Θ. His problem in period t is to assign the labor force and allocate investment so as to maximize a social welfare function over the horizon. The SWF is a discounted sum of expected aggregate utilities, utility being a function of consumption net of labor mobility costs. Formally, he must solve

$$(1.6) \qquad \max_{\substack{I_{ct}, I_{kt},\\ L_{ct}, L_{kt}}} \sum_{i=t}^{T} \beta^{i-t} E(U(\bar{c}_i(\Theta)))$$

subject to:

1) given initial values of K_{ct}, K_{kt}, $L_{c,t-1}$, $L_{k,t-1}$ and Θ_t

and 2) constraints (1.1) to (1.5), replicated for each period $i = t, t+1, \ldots, T$,

where β = a constant discount factor,

$$\bar{c}_i(\Theta) = c_i(\Theta) - v_i(\max(L_{ci} - L_{c,i-1}, 0) + \max(L_{ki} - L_{k,i-1}, 0))$$

is consumption net of mobility costs, and expectations are taken with respect to and as of period t.

This is, of course, a problem in dynamic optimization;

we may think of it as a sequence of single period problems, with the decisions of each period determining the initial conditions of subsequent periods. There is a well-developed methodology, due to Bellman and others, for solving this - at least in specific cases. We consider this methodology briefly. The first step is to define a new function, V_t, equal to the maximum attainable value of the objective function as of period t:

$$(1.7) \qquad V_t = V_t(K_{ct}, K_{kt}, L_{c,t-1}, L_{k,t-1}, \theta_t)$$

$$= \max_{\substack{I_{ct}, I_{kt}, \\ L_{ct}, L_{kt}}} \sum_{i=t}^{T} \beta^{i-t} E(\ U\ (\bar{c}_i(\theta)))$$

subject to the constraints and to the initial conditions that form the arguments of V_t. The planner's problem can now be rewritten

$$(1.8) \qquad \max_{\substack{I_{ct}, I_{kt}, \\ L_{ct}, L_{kt}}} U(\min(K_{ct}, L_{ct}) - v_t(\max(L_{ct} - L_{c,t-1}, 0) + \max(L_{kt} - L_{k,t-1}, 0))) + E(V_{t+1}(K_{ct} + I_{ct}, K_{kt} + I_{kt}, L_{ct}, L_{kt}, \theta_{t+1}))$$

subject only to current-period constraints and initial conditions. This is Bellman's Optimality Principle, that any optimal path can be broken up into subpaths that are optimal with respect to their initial conditions.

If we knew the form of V_{t+1}, then the problem would be reduced to the one-period type and could be easily solved. V_{t+1} can be obtained, in principle at least, by working recursively backwards from period T. Begin by noting that we do know the form of V_T:

(1.9) $$V_T(K_{cT},K_{kT},L_{c,T-1},L_{k,T-1},\Theta_T) =$$

$$\max_{L_{cT}} \; U(\min(K_{cT},L_{cT}) - v(\max(L_{cT}-L_{c,T-1},0))) \; .$$

V_{T-1} can now be obtained as a function of initial conditions, by solving (1.8) for $t = T-1$. Proceeding recursively, work back to the decision period. The result of this exercise is a nonstochastic optimal first period allocation. As later values of the stochastic sequence $\{\Theta_i\}$ become known, nonstochastic values for the allocations in periods t+1 to T can be calculated from the V_i.

We know, then, how the planner can solve his problem for any specific set of initial values, parameters, and functional forms. Unfortunately, there is no simple way to write down this solution in the general case; the expressions for the optimal allocations grow more complicated with each stage of recursion, and the side conditions multiply rapidly.

While we cannot find an explicit solution to the general planner's problem, we can at least hope to characterize that solution in an interesting way. One approach is suggested by the dynamic stochastic programming procedure just discussed. With that procedure, current decisions are made under the assumption that all future allocations will be optimal, given the current decisions (the Optimality Principle). Let us think of future optimal allocations explicitly as functions of current (period t) allocations and of realizations of Θ_i . (We take initial conditions in period t as fixed.) Denoting an optimal allocation with an asterisk, we can write

$$
\begin{aligned}
(1.10) \qquad I^*_{ci} &= I^*_{ci}(I_{ct}, I_{kt}, L_{ct}, L_{kt}, \{\Theta_{t+1}, \ldots, \Theta_i\}) \\
I^*_{ki} &= I^*_{ki}(I_{ct}, I_{kt}, L_{ct}, L_{kt}, \{\Theta_{t+1}, \ldots, \Theta_i\}) \\
L^*_{ci} &= L^*_{ci}(I_{ct}, I_{kt}, L_{ct}, L_{kt}, \{\Theta_{t+1}, \ldots, \Theta_i\}) \\
L^*_{ki} &= L^*_{ki}(I_{ct}, I_{kt}, L_{ct}, L_{kt}, \{\Theta_{t+1}, \ldots, \Theta_i\})
\end{aligned}
$$

for $i = t+1, \ldots, T$.

In the obvious way we would now write down the Lagrangian of the general planner's problem, including the constraints for each period, in a form depending only on the expectations (as of period t) of $(I^*_{ci}, I^*_{ki}, L^*_{ci}, L^*_{ki})$ $i = t+1, \ldots, T$ and on the current decision variables: I_{ct}, I_{kt}, L_{ct}, and L_{kt}. This Lagrangian can be viewed as a function of only the current decision variables, since expected future optimal allocations depend only on these variables. We maximize by differentiating with respect to the current variables, noting that the envelope theorem permits us to ignore changes in $(I^*_{ci}, I^*_{ki}, L^*_{ci}, L^*_{ki})$. Assume that in the current period it is optimal to invest in both sectors, so that the nonnegativity constraints on I_{ct} and I_{kt} are not binding. Then this procedure yields the following necessary conditions:

(1.11) (Consumption sector investment)

$$\lambda_{it} = \sum_{i=t+1}^{T} \beta^{i-t} E(U' (\bar{c}_i(\theta))\ Z_{1i}(\theta))$$

$$\text{where} \quad Z_{1i} = \begin{cases} 1 & \text{if } K^*_{ci} < L^*_{ci} \\ 1-\lambda_{2i} & \text{if } K^*_{ci} = L^*_{ci} \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{and} \quad K^*_{ci} = \begin{cases} K_{ct} + I_{ct} + \sum_{j=t+1}^{i} I^*_{cj} & , \ i > t \\ K_{ct} & , \ i = t\ . \end{cases}$$

(1.12) (Capital sector investment)

$$\lambda_{1t} = \sum_{i=t+1}^{T} \beta^{i-t} E(\lambda_{1i} \cdot (\frac{\theta i}{a}) \cdot Z_{2i}(\theta))$$

$$\text{where} \quad Z_{2i} = \begin{cases} 1 & \text{if } K^*_{ki} \quad L^*_{ki} \\ 1-\lambda_{2i} & \text{if } K^*_{ki} = L^*_{ki} \\ 0 & \text{otherwise.} \end{cases}$$

(1.13) (Consumption sector labor)

$$\lambda_{2t} = U'(\bar{c}_t)((1-Z_{1t}) - v_t(Z_{3t})) + \sum_{i=t+1}^{T} \beta^{i-t} E(U'(\bar{c}_i(\Theta))(1-Z_{1i}(\Theta))-\lambda_{2i}(\Theta)).$$

$$\text{where} \quad Z_{3i} = \begin{cases} 1 & \text{if } L^*_{ci} > L^*_{c,i-1} \\ 0 & \text{otherwise.} \end{cases}$$

(1.14) (Capital sector labor)

$$\lambda_{2t} = \lambda_{1t}(\frac{\Theta t}{b})(1-Z_{2t}) - U'(\bar{c}_t)v_t(Z_{4t}) + \sum_{i=t+1}^{T} \beta^{i-t} E(\lambda_{1i}(\frac{\Theta i}{b})(1-Z_{2i}))-\lambda_{2i}(\Theta)) .$$

$$\text{where} \quad Z_{4i} = \begin{cases} 1 & \text{if } L^*_{ki} > L^*_{k,i-1} \\ 0 & \text{otherwise.} \end{cases}$$

Expectations are understood to be conditional on information in period t and to be with respect to t.

Despite the notational difficulties, these conditions have obvious economic interpretations. First, note that the Lagrange multipliers λ_{1i} and λ_{2i} represent the marginal values in period t of a unit of uncommitted capital or labor, respectively. As the necessary conditions imply, at an optimum these marginal values must be the same in each of the two alternate uses.

The marginal value of a unit of consumption sector capital, given by (1.11), is the discounted consumption value of its marginal product in all future periods. In periods when consumption sector capital is expected to be greater than consumption sector labor (the indicator $Z_{1i}=0$), this marginal product is zero; when consumption capital is expected to be the binding constraint ($Z_{1i}=1$), the expected marginal product is a unit of consumption goods.

Similarly, the expected marginal product of a unit of investment sector capital (1.12) is zero for periods when excess capacity in that sector is expected ($Z_{2i}=0$). In periods of expected insufficient capacity, the marginal product is Θ_i/a units of investment goods, each of which has a discounted value of $\beta^{i-t}\lambda_{it}$. The marginal value of an unassigned labor unit is expressed either by (1.13) or (1.14). The expressions are similar to those for capital; i.e., the expected marginal product of a unit of labor is positive only in those periods when expected optimal allocations make sectoral labor, rather than sectoral capital, the binding constraint. A difference between labor and capital is that labor has mobility costs, which must be deducted from the expression for value. Also, λ_{2i} must be deducted from product to correct for the possibility that labor might have been brought in later at lower mobility cost. If we assumed the existence of variable installation costs for capital, the two sets of equations would be exactly analogous.

Note that $\lambda_{1t}=0$ implies that the economy is saturated with capital, both in the current and future periods. $\lambda_{2t}=0$ means that the discounted product of a worker for periods t to T is not sufficient to overcome current mobility costs.

From these conditions we can draw several conclusions about the nature of the planner's optimal solution:

1) When capital is durable and there are labor mobility costs, current optimal allocation depends not only on the current state of the economy but, in a complicated way, on all future states. Any econometric model, say, that looks at investment or employment as dependent only on current variables is implicitly assuming a very naive set of economic agents. This, of course, is a major point of Lucas's well-known critique.

2) An optimal plan for this economy may include (intentional) periods of excess capital capacity. To see this, we note that necessary conditions (1.11) and (1.12) ascribe positive value to capital additions as long as there is some future period in which the sector is expected not to have excess capacity. Thus it is conceivable that an optimal plan might call for, say, capital additions to a sector which currently has idle capital, at the expense of the other sector which may currently be capital-short. If, for example, future values of Θ are expected to be high, it may well be efficient to hoard capital in the capital goods sector, despite current shortages of consumption sector capital or capital sector labor. Alternatively, if a labor supply spurt is expected in the future, it may be efficient first to build up the capital sector and then to maintain excess capital capacity in <u>both</u> sectors, until labor becomes available.

3) Just as there is the possibility of efficient excess capital capacity, there may be efficient unemployment of labor resources. This may arise from one of several causes. First, as in the case of capital, necessary conditions (1.13) and (1.14) imply a positive marginal value to sectoral labor as long as there is some future period in which all the sector's labor is utilized.

Thus there is no inconsistency between these conditions and a plan that, say, hoards labor in a sector that is currently capital-short but where relative productivity is secularly increasing. Second, unlike capital goods, we have not assumed that labor must be committed to one or the other sector as soon as it becomes available. The value of maintaining a pool of uncommitted resources when there is uncertainty has already been discussed in our first Chapter, and the same arguments apply here. Finally, in this model, maximum output and maximum employment may be incompatible goals in the long run. Indeed, for certain values of the parameters, at least, the strategy that maximizes employment - the committing of all resources to the capital goods sector - corresponds exactly to the strategy that <u>minimizes</u> consumption and the level of utility.

4) With positive mobility costs, efficient excess capacity and unemployment can exist in the economy at the same time. This is because with nonzero mobility costs it is not worthwhile to move labor between sectors to take advantage of short-lived opportunities. Movement will occur only if the <u>long-run</u> value of the worker in the alternative sector exceeds his value in his present sector plus mobility costs.

Because we have assumed that (given the availability of input stocks) the variable costs of production are zero, no optimal plan in this model will ever include contemporaneous excess capacity and unemployment in the same sector. This defect is remedied in our model of Chapter 3. There, changes in the marginal costs of utilization permit the coexistence of layoffs and idle machines within a given sector.

II. Let us move this model economy into a monetary market framework. It is our object to characterize briefly agent behavior and to search for the correspondences between the market and planning version.

Suppose that consumers in this economy 1) are intertemporal optimizers, 2) have separable von Neumann-Morgenstern utility functions that depend on net consumption and real balances only, and 3) in each period make joint decisions about consuming, building up real balances, purchasing shares of firms in a stock market, and working. Then we may write the i-th consumer's problem as:

$$(2.1) \qquad \max_{M_{it}, s_{it}, v_{it}} W_i = \sum_{j=t}^{T} \beta_i^{j-t} E(U_i(C_{ij}(\Theta) - v_{ij}(\Theta), m_{ij}(\Theta)))$$

such that

$$M_{it} = M_{i,t-1} + H_{it} + p_t Y_{it} + \tilde{s}_{it}\tilde{R}_t - (\tilde{s}_{it} - \tilde{s}_{i,t-1})\tilde{X}_t - p_c c_{it}$$

and $\qquad Y_{it} = Y_{it}(v_{it})$

where β_i = a constant discount factor

c_{ij} = real consumption (in period j)

m_{ij} = real balances

v_{ij} = real mobility costs incurred

M_{ij} = nominal balances

H_{ij} = nominal transfers

Y_{ij} = real labor income

$\tilde{s}_{ij}$ = the vector of fractions of firms held by the i-th individual $(\sum_i \tilde{s}_{ij} = (1,1,\ldots,1))$

$\tilde{R}_j$ = the vector of firm dividend payments

$\tilde{X}_j$ = the vector of firm market values

p_j = the price of consumption goods

The stochastic process generating Θ_t will be assumed known, for simplicity; however, the results can easily accomodate Bayesian priors. The current value of Θ - and hence, current dividends and market values - are assumed known. Note that the consumption decision is implicit.

The necessary conditions for an interior solution can be derived:

(2.2) (Money holdings)

$$-\frac{1}{p_t}\frac{\partial u}{\partial c}(c_t,m_t) + \frac{1}{p_t}\frac{\partial u}{\partial m}(c_t,m_t) +$$

$$\beta \int_{\Theta_{t+1}} {}_tf_{t+1}(\Theta_{t+1}) \cdot \frac{\partial u}{\partial c} (c_{t+1}(\Theta_{t+1}),$$

$$m_{t+1}(\Theta_{t+1})) \cdot \frac{1}{p_{t+1}(\Theta_{t+1})} d\Theta_{t+1} = 0 \quad .$$

(2.3) (Stock holdings)

$$\frac{1}{p_t} \frac{\partial u}{\partial c} (c_t, m_t) \ (\tilde{R}_t - \tilde{X}_t) + \beta \int_{\Theta_{t+1}} {}_tf_{t+1}(\Theta_{t+1}) \cdot$$

$$\frac{\partial u}{\partial c} (c_{t+1}(\Theta_{t+1}), m_{t+1}(\Theta_{t+1})) \cdot$$

$$\frac{1}{p_{t+1}(\Theta_{t+1})} \tilde{X}_{t+1}(\Theta_{t+1}) \ d\Theta_{t+1} = 0 \ .$$

where ${}_tf_{t+1}$ is the distribution of Θ_{t+1} given information in period t, and the i-subscripts have been suppressed.

We can rewrite these conditions as:

$$(2.4) \qquad 1 = \frac{\frac{\partial u}{\partial m} (c_t, m_t)}{\frac{\partial u}{\partial c} (c_t, m_t)} + \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot 1 \cdot d\Theta_{t+1}$$

$$(2.5) \qquad \tilde{X}_t = \tilde{R}_t + \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot \tilde{X}_{t+1}(\Theta_{t+1}) \ d\Theta_{t+1}$$

where $\alpha_t(\Theta_{t+1}) =$

$$\beta \ {}_tf_{t+1}(\Theta_{t+1}) \ \frac{p_t \cdot \frac{\partial u}{\partial c} (c_{t+1}(\Theta_{t+1}), m_{t+1}(\Theta_{t+1}))}{p_{t+1}(\Theta_{t+1}) \ \frac{\partial u}{\partial c} (c_t, m_t)}$$

can be viewed as a generalized time-state discount factor. This discount factor is of some interest, because it is easily shown that for any asset with current value p_t, return r_t, and uncertain future values $p_{t+1}(\Theta_{t+1})$, an optimizing consumer who holds some of the asset will set

$$(2.6) \qquad p_t = r_t + \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot P_{t+1}(\Theta_{t+1})\, d\Theta_{t+1}$$

In particular, (2.4) has the interpretation that the current price of money (equal to one) equals current services of money plus the discounted future value of money (equal to one in all states). Similarly, (2.5) says that current firm market values equal current dividends plus future market values discounted by time and state.

Note that, even if individuals have different wealth, utility functions, and priors, in equilibrium the market will insure that certain weighted integrals of the time-state discount factors (linear combinations, in the discrete-time case) will be equal for all individuals. In the futures-market-equivalent case, when there are as many independent assets as future states, consumption will be adjusted so that

$$(2.7) \qquad \alpha_{it}(\Theta_{t+1}) = \alpha_{jt}(\Theta_{t+1})$$

for all (i,j) and all values of Θ_{t+1}. This implies that gains from trade have been exhausted and the economy is at an efficient point.

Let us turn now to the consumer's work decision. Recall that a worker has three options: 1) He may be unemployed, earn zero, and live off lump sum transfers and wealth. 2) He may work in the consumption goods sector and earn w_{ct}. 3) He may work in the capital goods sector and earn w_{kt}. If he chooses to work in a sector where he is not presently located, he incurs real expense v_t. There is no disutility to labor, but the worker has a maximum labor endowment.

Since there is no disutility to labor, the worker need only choose the option that yields the preferred expected income stream. Obviously, initial employment status of the worker makes a great deal of difference; an unemployed worker is much more likely to move into the capital goods sector, say, than is a worker who already has a job in the consumption goods sector. Looking first at the unemployed or newly entering worker, and using our time-state discount factor notation, we can write the necessary conditions for the worker to be indifferent among his options:

$$(2.8) \qquad v_t = \frac{w_{ct}}{p_t} + \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \min(v_{t+1}, v^*_{c,t+1}(\Theta_{t+1}))\, d\Theta_{t+1}$$

$$v_t = \frac{w_{kt}}{p_t} + \int_{\theta_{t+1}} \alpha_t(\theta_{t+1}) \min(v_{t+1}, v^*_{k,t+1}(\theta_{t+1}))\, d\theta_{t+1}$$

where $v^*_{c,t+1}(\theta_{t+1})$ is the mobility cost at which an unemployed worker would just be indifferent to entering the consumption sector in period t+1, given θ_{t+1}. ($v^*_{k,t+1}(\theta_{t+1})$ is defined analogously.) This has the interpretation that, at indifference, the current mobility cost must be equal to the current real wage plus expected mobility cost savings gained by moving now rather than later. Note that the more "intuitive" condition

$$(2.9) \qquad U'(c_t)v_t = U'(\bar{c}_t)\frac{w_t}{p_t} + \sum_{i=t+1}^{T} \beta\, E(U'(c_i(\theta))\frac{w_i}{p_i}(\theta))$$

which equates the present value of real wages to the mobility cost, is true <u>only</u> if there is expected to be positive unemployment in every future period.

The conditions under which an employed worker would be indifferent between staying where he is and moving to the other sector are the same as (2.8), except that real wage differentials take the place of real wages.

Firms in this economy are competitive, may be either consumption or capital goods producers (though they may never switch sectors), and have the same fixed-coefficients technologies postulated for their respective sectors in Part I. Because the technology is fixed-coefficients, marginal productivities depend only on the endowments of the sector as a whole; it does not matter how the initial capital stock is distributed among firms.

We assume the existence of a stock market but no futures markets. Without futures markets profit maximization is not well-defined. We shall suppose that firms instead maximize their current stock market value. This implies that firms have knowledge of how the market evaluates potential income streams; i.e., firms must know some aggregate version of equation (2.5). Tentatively we write down the firm's problem as

$$(2.10) \qquad \max_{I_t, L_t} \left(X_t = R_t + \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot X_{t+1}(\Theta_{t+1})\, d\Theta_{t+1} \right)$$

where, recall, X_t is stock market value and R_t is net dividend payments.

Before working with (2.10), we must make several points:

1) The expression $\alpha_t(\theta_{t+1})$ is supposed to represent some market aggregate of the time-state discount factors of individuals. However, except in the futures-markets-equivalent case when there are as many independent assets as states, the aggregate α_ts will be underidentified (i.e., there are less integral restrictions than states). We will assume that the firm picks any set of α_ts consistent with existing restrictions.

2) We have not assumed the existence of a bond market (although it would not be a great complication to do so). Firms must therefore make purchases of new capital out of current earnings, creating the possibility of negative dividends. This is not a difficulty. With perfect bonds markets, negative dividends are the same as a combination of positive dividends and increased firm debt (Modigliani-Miller). Without bonds markets, there is no reason to restrict an offered income stream to nonnegative components. We can therefore express dividends as:

(2.11) (Consumer goods firms)

$$R_{ct} = p_t \cdot \min(K_{ct}, L_{ct}) - w_{ct}L_{ct} - p_{kt}I_{ct}$$

(2.12) (Capital goods firms)

$$R_{kt} = p_{kt} \cdot \min(K_{kt}/a, L_{kt}/b) \cdot \theta_t -$$

$$w_{kt}L_{kt} - p_{kt}I_{kt} \quad .$$

3) In their maximization, competitive firms must take the time-state discount factors they employ as parameters. Specifically, a firm would not take into account any change in the discount factor caused by the firm's impact on aggregate consumption.

With these caveats, we replace R_t in (2.10) with the expressions in (2.11) and (2.12), treat the α_ts as constant, and maximize with respect to L_t and I_t. This yields the necessary conditions for an interior solution:

(2.13) (Consumer goods firms)

$$p_{kt} = \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot \max(p_{t+1}(\Theta_{t+1}) - w_{c,t+1}(\Theta_{t+1}),0) + p_{k,t+1}(\Theta_{t+1}))\, d\Theta_{t+1}$$

$$L^*_{ct} = \begin{cases} K_{ct} & w_{ct} \leqq p_t \\ 0 & w_{ct} > p_t \end{cases} \quad .$$

(2.14) (Capital goods firms)

$$p_{kt} = \int_{\Theta_{t+1}} \alpha_t(\Theta_{t+1}) \cdot (\max(\tfrac{1}{a}(p_{k,t+1}(\Theta_{t+1}) \cdot \Theta_{t+1} - bw_{k,t+1}(\Theta_{t+1})),0) + p_{k,t+1}(\Theta_{t+1}))\, d\Theta_{t+1}$$

$$L^*_{kt} = \begin{cases} K_{kt} & w_{kt} \leqq p_{kt} \cdot \theta_t \\ 0 & w_{kt} > p_{kt} \cdot \theta_t \end{cases} .$$

Interior solutions equate the current price of capital to the time-state-discounted sum of next period's return to capital and next period's capital price. Also, an interior solution implies that expected marginal returns to capital next period are equal in the two sectors. This is not true ex post or if investment takes place only in one sector (corner solution).

We have calculated the optimizing behavior of agents in a particular market economy. As one might suspect, there is a strong duality between the conditions we derived in this part and those derived for the planned economy in Part I. It is interesting to ask when those conditions will be identical so that the market economy will duplicate the physical allocations of the planned economy.

It turns out that there are two necessary conditions, one of which is very close to assuming the futures-markets-equivalents case in the market economy. First, the objective functions of the two economies must be comparable. Second, the aggregate time-state discount factors must be uniquely determined. We can get these conditions by assuming that everyone has the same utility function, the same consumption, and that the utility function is of the form

$$(2.15) \qquad U^M(c_t, m_t) = U^P(c_t/L_t) + w_t(m_t)$$

where U^M is the consumer's utility function and U^P is the planner's objective function. These assumptions are heroic, of course, especially because complicated transfer payments are needed to give workers in different sectors the same consumption. It is a worthwhile exercise, however, because with these assumptions and some algebraic manipulation one can show the equivalence of the planner's necessary conditions (1.11) - (1.14) to the market economy conditions (2.13), (2.14), and (2.8). From the algebra we also get

$$(2.16) \qquad \lambda_{1t} = \frac{U'(\bar{c}_t)}{p_t} p_{kt}$$

$$\text{and} \qquad \lambda_{2t} = \frac{U'(\bar{c}_t)}{p_t} w_t + \beta E\left(\frac{U'(c_{t+1})}{p_{t+1}} \cdot \min(v_{t+1}, v^*_{t+1})\right) - U'(\bar{c}_t) v_t$$

which are market-oriented expressions for the planning problem Lagrange multipliers associated with a unit of capital or labor in period t.

The market economy does not have to duplicate the planned economy to be efficient. For efficiency we need the time-state discount factors to be unique. This is virtually equivalent to the full futures market case. In general, then, the market economy will contain inefficiencies.

Much needs to be said about the general equilibrium properties of the market model, especially about the determination of wages and prices, but I wish to defer that. Instead, let us recall some of the interesting properties of the planning model - notably, that it may exhibit efficient excess capacity and unemployment. By analogy of the necessary conditions, it is clear that the market economy will behave similarly. Thus it appears that we have a market economy model with 1) inefficiencies, and 2) unemployment of resources, where 1) and 2) are separate phenomena. The inefficiencies are due to the inability of agents to trade over all future periods and contingencies. The resource unemployment arises from speculation in the markets for capital and labor stocks. The latter phenomenon will be present in even a completely efficient economy.

What forms will this speculative resource unemployment take? From the necessary conditions we see that firms may add capital even in a period of excess capacity, both in anticipation of future opportunities and to beat expected increases in capital goods prices. Similarly, labor stocks may be held in reserve by firms. Workers may speculatively choose unemployment either by entering a sector where there is currently no work, or by remaining uncommitted while awaiting new information. In general, each of the forms of efficient unemployment discussed in the planning case has its market analogue.

We have given what is essentially an extended example, in which hoarding and speculation that lead to unutilized resources are given an efficiency justification. This outcome is expected to be theoretically robust, in the sense that it will follow from any economic model that obeys the broad and plausible description given in the introduction to this paper. The empirical importance of these phenomena, on the other hand, is not clear. These effects may possibly be swamped by some other source of unemployment (i.e., Keynesian disequilibrium). Resolution of this is left to future work.

CHAPTER THREE

INCOMPLETE LABOR CONTRACTS, CAPITAL, AND THE SOURCES OF UNEMPLOYMENT

Introduction

Recent years have seen the appearance of a relatively large number of theoretical papers on contracting between economic agents. This literature received its impetus from an influential reinterpretation of Keynes, offered by Barro and Grossman (1971), Clower (1965), and Leijonhufvud (1968), among others, which recast the traditional model into a theory of disequilibrium (quantity-constrained) trading. The formulation of Barro-Grossman *et al*. is appealing; it is, however, tenable only in a world where prices and wages do not adjust instantaneously to clear markets. This is something of a difficulty, as the persistent failure of a market to clear implies that profitable opportunities exist which are not being exploited. Contract theory was a response to this problem; it has attempted to resolve the difficulty by showing that, in a regime where there is contracting, short-term opportunities may rationally be foregone in the interest of longer-term benefits.

Much of the contracting literature has concentrated on the labor market, the market where the occurrence of serious and persistent disequilibria seems most credible. Most of these papers, though not all, as we shall see, have appealed to the risk-sharing motive for contracting. Azariadis (1975), Baily (1974), D. Gordon (1974), and Grossman (1975) have argued that labor contracts are preferred to spot markets because they allow employers to sell income insurance to (more risk-averse) employees. The contract form that provides the preferred insurance, they claim, is one that keeps wages fixed and varies employment. This is supposed to explain sticky wages and, consequently, the failure of the labor market to clear. The contracts that have these effects, moreover, need not be formal agreements; the contracts may be *implicit*, a part of the accepted way of doing business.

The "implicit contracting" theory has not convinced everyone. A counterargument to this and other theories of non-Walrasian allocations induced by contracts is offered by Barro (1977). Barro's idea may be stated as follows: Any contract that is supposed to be "optimal" must, *a fortiori*, provide for a Walrasian allocation of labor services in the short run; i.e., the marginal value of labor's product must equal the marginal value of leisure foregone. If this is not the case, *ex post* negotiations could make both parties better off. The claim of optimality for fixed-wage contracts, for instance, must be flawed (in this view), unless provisions for *ex post* employment adjustments are included. This is approximately true even if there are costs of *ex post* adjustment; it would still pay both parties to eliminate large deviations from the Walrasian solution.

The present paper is an attempt to reconcile these two positions and to relate labor contracting and the apparent disequilibrium phenomena observed in the labor market. Our approach is as follows. We concede to Barro the point that no fully optimal contract could produce a non-Walrasian allocation of labor (this is virtually by definition). However, we suggest that real-world labor contracts may be "incomplete", i.e., they may face exogenous restrictions on their form or content. Because certain contract provisions are not available (just as certain markets are "missing" in the classic Arrow-Debreu framework), real-world contracts are of a "second-best" nature. One result may be non-Walrasian labor allocations. The incomplete contracting framework allows a pinpointing of the differences in assumptions that cause Barro and the contract theorists to reach opposing conclusions.

Beyond clarification of the contracting debate, the incomplete contracting device is of independent use in analyzing unemployment in the labor market.

We show first that, under a certain type of contract incompleteness, the coexistence of a positive wage and involuntary unemployment need not imply that there are unexploited private opportunities for arbitrage. Further, the inclusion of incomplete contracts in a model with capital and variable utilization rates reveals that there are many forms of unemployment - some efficient, some inefficient - consistent with perfect information and zero unexploited private opportunities.

The paper is in two parts.

Part I studies contracting in the labor market. The distinction between complete and incomplete contracts is motivated and used to discuss the contracting debate. A simple model demonstrates how, with incomplete contracts, a positive wage can persist in the face of involuntary unemployment.

Part II introduces a capital stock with a fixed-proportions technology and variable utilization rates. It is shown that there are numerous potential sources of unemployment, even within a given sector, that are consistent with rational behavior.

I. Complete and Incomplete Labor Contracts

We begin our discussion of labor market contracting by asking what form an ideal contracting instrument would take. We propose the following minimum properties: The ideal contract must 1) be enforceable upon both parties, 2) admit of any type of transaction, and 3) dictate a well-defined outcome (or procedure) for every distinguishable state of nature. A contract that has these properties we will call a _complete_ contract. One justification for setting the complete contract up as a standard is that, in a labor market with competition on both sides and complete contracts, a Walrasian allocation of labor will be enforced in every period and state; else, more profitable contracts would be available. The existence of complete contracts is sufficient (though not necessary) for Barro's view to be correct.

If contracts are complete, then they merely form a "veil" under which an essentially Walrasian result obtains. However, there is reason to think that real contracts, especially in the labor market, are _incomplete_, i.e., they lack one or more of the above properties. Incompleteness stems from exogenous restrictions on the form or content of the contract; the analogy is to "missing markets" in the Arrow-Debreu model. Some possible sources of contract incompleteness are listed below.

1) Because of prohibitions agains slavery or indenture, and because of difficulties in setting a legal standard of worker compliance, labor contracts are typically not fully enforceable on workers. This incompleteness causes observed contracts to differ from the idealized model in several ways. Contracts must be structured to make voluntary compliance attractive. This may involve staggering benefits towards the end of the working life (through seniority rules, for example), setting up artifical barriers to mobility, or giving workers firm-specific training which is not easily used elsewhere. Such provisions may be inefficient. The magnitude of inefficiency will depend, among other things, on the presence of natural worker mobility costs. If mobility costs are sufficiently high, nonenforceability is not a problem.

We note that, even if there are no mobility costs, the nonenforceability restriction does not reduce contracts to the spot market case. Firms may still desire to offer one-way contracts, which bind the firm but not the worker. The advantage of one-way contracts is that, by allowing the firm to make a commitment to deliver certain benefits in the future, some gains from trade not available in the auction market may be realized. For example, a firm-owner with a high discount rate, by committing himself to give greater benefits than other firms in the future, can reduce his current labor costs and increase his utility. The firm-owner with a low discount rate is not helped by a one-way contract; he cannot offer higher benefits now in exchange for low benefits later, since he cannot bind his workers to stay with him in the later period.

2) Law, custom, union practices, etc., do not permit certain transactions between employers and individual workers. An outstanding example of this is minimum wage laws. Other examples include health and safety regulations, insurance and pension rules, statutory work weeks. While such arrangements are probably desirable on net, they do act to set a lower bound on the effective "wage offer" of a potential new worker. Thus, some employer-worker matches desired by both parties are prevented.

3) A third contract incompleteness stems from asymmetrical or incomplete information about states of nature. Information gaps may introduce moral hazard or adverse selection problems that make a fully contingent contract impractical. Hall and Lillien (1977) have made a study of this situation. They

suggest that information failure is an important determinant of the form that contracts actually take.

When there are potential moral hazard problems on both the supply and demand sides, say Hall and Lillien, a fully efficient contract is impossible. As a second-best measure, to increase short-run efficiency in the use of labor, real labor contracts provide for the "internalization" by one party, usually the firm, of both the costs and benefits of variations in labor hours. A typical contracting pattern is as follows: Periodic negotiations between the workers and the firm establish 1) a base level of compensation, B, and 2) a supply-of-labor-to-the-firm function, $V(x)$. As business conditions change, the firm is allowed to vary the number of labor hours required (x) unilaterally. However, the labor supply function has the property that, no matter how many labor hours are required, workers always receive just enough compensation to keep them indifferent between their current compensation-hours package and the base level of compensation, B. This arrangement makes sure that the firm's labor-hours decision will, if profits are maximized, equate the marginal product and marginal disutility of labor (as embodied in $V(\cdot)$). If labor supply shocks are small relative to those affecting productivity or the demand for output, then short-run allocative efficiency is realized. Note that the firm has no incentive to lie about its true demand for labor, as workers have no incentive to lie about their supply curve. Occasional renegotiation shifts the base compensation level, to keep it in line with current supply conditions in the economy.

We have suggested three ways in which real labor contracts may be incomplete, relative to an ideal contract; we do not pretend that this short list is complete. The problems that prevent the realization of the complete contract are of a nature similar to those that make the economy as a whole different from the Arrow-Debreu ideal: enforcement difficulties, institutional constraints, informational asymmetries, transactions costs. As in the case of missing markets, traders who must use incomplete contracts search for institutional or other arrangements in order to achieve the best possible allocation.

The incomplete contracting model can provide a framework in which to study the contracting debate. If we want to think of decision makers as rational and efficient, the benefit of the doubt must be given to Barro: Agents in the labor market will try to use contracts to achieve Walrasian allocations of labor. Writers who claim that optimal contracts are responsible for non-Walrasian allocations must show two things: First, that there are plausible reasons for assuming the agents are restricted away from complete contracts in a certain way. Second, that within the restricted class agents are permitted to use, the optimal contract leads to a non-Walrasian allocation of labor.

Let us see how the implicit contracting theorists fare under these criteria. The class of possible contracts they allow is a small one - those in which compensation is linear in labor hours (i.e., there is a fixed hourly wage). Other methods of compensation - e.g., lump-sums, or payments nonlinear in hours - are not considered. These writers satisfy our second criterion by proving fixed-wage-variable-employment theorems about the class of contracts they admit. However, they do not justify their severe restriction on contract form (first criterion). This is important, as their whole argument rests on the necessity of using <u>one</u> instrument (the wage) to perform two functions (risk-sharing and labor allocation). It is not clear why contracts must be so limited.

In contrast, Hall and Lillien are explicit about why they restrict the class of contracts they consider (informational asymmetries, moral hazard), fulfilling our first condition. Their discussion of the optimal contract within that class (second criterion) is non-rigorous, but (to us anyway) still plausible.

We will try to meet the two criteria ourselves in the next section, when we argue for a non-Walrasian result due to incomplete contracting.

Contracts and an apparent labor market disequilibrium. In this section we will show that some of the forms of contract incompleteness described above can create a situation that looks like excess supply in the labor market. This apparent disequilibrium is really an equilibrium, however, as there are no unexploited opportunities for private profit to motivate its elimination.

We will introduce a simple model in which there are two types of workers, "trained" and "untrained". By "trained" we mean something broader than "having acquired a certain set of technical skills." We will think of a trained worker in this model as one who is experienced in the primary labor market; who knows the rules and customs of the workplace; who has demonstrated the ability to show up on time, follow instructions, etc. An untrained worker is to be thought of as a new entrant or secondary market worker who may have (let us say) the same native ability as a trained worker, but is without primary market experience.

The following assumptions form the model:

1) Trained workers are each affiliated with a specific firm. Untrained workers make up a central pool.

2) The output of a trained worker in period i is X_i. X_i follows a random walk over time.* The productivity of an untrained worker is normalized to zero.

3) An initial investment of D is required to "train" an untrained worker and bring him to a firm. An initial cost of d is needed to move a trained worker from one firm to another. Assume $0 < d < D$.

The difference in costs D-d includes not only the costs of imparting technical skills on the job but also the costs of social adjustment and of "screening" (the cost imposed when a certain fraction of previously inexperienced workers turn out to be unacceptable). We make the crucial (but realistic) assumption that the firm must undertake at least some of the additional training required by a new worker. This is equivalent to assuming that, by such devices as diplomas, a new

*We can bound X_i away from zero by assuming that X_i=a implies X_{i+1}=a, where a is a small positive number.

worker on his own initiative cannot make himself a perfect substitute for an experienced worker.*

4) Contracts are incomplete in that a) workers cannot bind themselves to stay with a given firm for more than one period; and b) there is a legal minimum wage of zero. We assume that the zero minimum wage of provision b) is effective. In particular, firms are not in general allowed to set themselves up as joint educational and productive enterprises, accepting "tuition" from new workers. In most cases, such a plan would look only like an evasion of the wage law; it might also be thwarted by the limited access of secondary or new workers to capital markets. The existence of apprenticeships does not contravene our assumption, as long as the worker does not actually pay the firm in order to work.

5) Firms operate in a competitive labor market for a fixed number of workers. New firms have free entry. Firms maximize the present discounted value of their expected profits, where β is the common discount factor. Workers maximize the present discounted value of their expected wages, with an arbitrary (within the unit interval) discount factor.

---

*The truth of this assumption is apparent to any reader of help-wanted ads.

We wish to find labor market equilibrium for this model. Let $w^t(X)$ be the wage paid a trained worker when X is the current level of productivity. Since one can always get a trained worker from another firm by paying his mobility cost (d) and a wage infinitesimally higher than w^t, competition ensures that

$$(1.1) \qquad d = \sum_{i=0}^{\infty} \beta^i (E(x_i - w^t(X_i)) = \frac{X_0 - w^t(X_0)}{1 - \beta}$$

which implies

$$(1.2) \qquad w^t(X) = X - (1-\beta)d$$

The first-period wage for an untrained worker, $w^u(x)$, satisfies

$$(1.3) \qquad D = (X_0 - w^u(X_0)) + \frac{\beta}{1-\beta} (X_0 - w^t(X_0))$$

so that

$$(1.4) \qquad w^u(X) = X - D + \beta d$$

Alternatively, (1.4) can be written

$$(1.5) \qquad w^u(X) = w^t(X) + (1-\beta)d - D + \beta d$$

$$= w^t(X) + (d-D)$$

(1.5) shows that there is a wage differential between trained and untrained workers <u>in each period</u> equal to the difference in training costs. The full difference must be made up in a single period because workers cannot bind themselves to stay with a given firm for more than one period.

It is possible for X to take values such that $w^t(X) > 0$, $w^u(X) < 0$. This will happen if

$$(1.6) \qquad (1-\beta)d < X < D-\beta d$$

When X is in this range, the non-negativity restriction on wages implies that no untrained workers will be hired. Trained workers will be kept on at wage $w^t > 0$. X falls in this range with greater likelihood the larger is D and the smaller is d. Since X is a random walk, a drawing of X_i in this range implies a relatively high probability that X_{i+1} will also satisfy (1.6).

This situation has all the characteristics of excess supply in the labor market. There is a pool of unemployed workers who, no matter what their discount rate, would be willing to work at a starting wage of zero. Workers already on the job are being paid a positive wage. Nevertheless, no untrained workers are hired, and there is no tendency for the wages paid to firm-affiliated workers to fall. Given the restrictions on contractual arrangements, no profitable opportunities are being left unexploited. This apparent disequilibrium (it is, in fact, an equilibrium) will persist until a drawing of X falls outside the bounds of (1.6).

The problem in this market is that only starting wages - rather than lifetime wages - can move to clear the market for entering workers. The new worker cannot offer a lower lifetime wage than those already employed because he cannot bind himself indefinitely to a specific firm. Because the starting wage must be non-negative (or above some minimum), it may not be able to go low enough to clear the market for new workers. At the same time that new workers cannot find a job at zero wage, old workers (who have training costs already sunk in them) are being paid a positive wage to keep them from defecting to other firms.

The restriction that makes this model work is the assumption that workers cannot provide all of their own "training" or, alternatively, pay the firm for training by taking a negative wage. Without this restriction, workers would bear all training costs and the externality would be eliminated. However, given our broad definition of "training" (perhaps "experience" would be a better word), we feel our assumptions are credible.

If one accepts this model, it provides an interesting appendix to the contracting debate. We have shown a non-Walrasian outcome in the labor market which is due not to contracts but (in some sense) to the <u>absence</u> of contracts (the unavailability of certain contracting provisions). Moreover, the non-Walrasian result occurs not in the allocation of labor of workers already affiliated with a firm (this is where previous writers have concentrated their efforts), but in the market for workers that no one has yet hired.

II. Sources of Unemployment - a Model with Capital and Contracting

The original purpose of the contracting literature was to explain the coexistence in the labor market of 1) unemployment and 2) prevailing wages above the reservation levels of the unemployed, without relying on the existence of unexploited profit opportunities. In this section we put incomplete contracts into a model with capital to show that there are many sources of such unemployment, even within a given sector.*

Our model assumes neither imperfect information nor constraints on sales; these are neglected not because they are unimportant, but because their implications for unemployment have already been studied elsewhere.

The nature of production in our model is simple and formally restrictive: we assume a fixed-coefficients input relation between the services of capital and labor. This is not the same as a fixed relation between stocks, as we permit the utilization rates of capital and labor to vary independently. Limited

*This model complements the analysis of the last Chapter, which showed how unemployment of resources can occur because of differential development <u>between</u> sectors.

capital-labor substitution *ex post* is, we feel, a realistic description of actual technologies; moreover, it is required for excess capacity to be consistent with optimization. In this paper we also impose limited substitution *ex ante*. This is only for simplicity and has no important qualitative bearing on the results.

We look at an economy in which a single output good is produced by identical firms. We write the production function for period t:

$$(2.1) \qquad y_t = \min(KS_t, LS_t)$$

i.e., the production function is fixed-coefficients in services. KS_t and LS_t are the capital services and labor services, respectively, used by the firm in period t: units have been normalized so that the factor ratio is one. We will assume that the output good, y_t, can either be consumed directly or transformed by the firm into durable capital; more on this shortly.

Services of an input are equal to the stock of the input times its utilization rate:

$$(2.2) \qquad KS_t = K_t \cdot x_{kt}$$

$$LS_t = L_t \cdot x_{lt}$$

where

$$0 \leqq x_{kt} \leqq 1$$

$$0 \leqq x_{lt} \leqq 1$$

K_t and L_t, the number of machines owned and the number of people employed by the firm, are fixed in the short run. x_{kt} and x_{lt}, the fraction of the day the machine or worker is engaged in production, are firm decision variables. For a cost-minimizing firm, x_{kt} and x_{lt} will be related in the short run by

$$(2.3) \qquad x_{kt} = L_t/K_t \cdot x_{lt}$$

Thus a firm with one shovel and two workers, if it wants to use its workers eight hours a day (x_{lt} = .333), will have to keep its shovel in use sixteen hours a day (x_{kt} = .667).

Let us now consider the relation of the firm and its workers. In a given period, t, the firm has L_t employees on its roster. There is an incomplete labor contract of the type described in the last section: i.e., the contract is legally binding only on the firm; and it is of the Hall-Lillien form. Because it is a Hall-Lillien contract, compensation is made to depend on hours worked in such a way as to give employees the same basic utility level no matter how many labor hours are required by the firm. The base utility level is renegotiated each period and depends on the quality of worker alternatives. The (one-way) contract may be thought of as being either one period or many periods in length. The one-period contract is essentially the spot market case; here the base utility level must be at least as great as in the workers' best alternatives, net of mobility costs. In the multi-period case, the firm has succeeded in binding itself to deliver at least certain levels of utility to its workers in later periods; here the current base utility need not be as high as in the workers' alternatives, if workers value future provisions of the contract sufficiently highly.

The model includes both the single- and multi-period cases. In either situation there will be a current labor compensation function of the form

$$V_1(x_{1t}, s_t)$$

which, for any labor utilization rate x_{1t} and state of nature s_t, gives the quantity of goods required to keep the worker at his base utility level, $\bar{U}(s_t)$. If there is a positive utility to leisure, then $\partial V_1/\partial x_1 > 0$; if there is diminishing marginal utility to leisure, $\partial^2 V_1/\partial x_1^2 > 0$.

An example of a compensation function is in order. Suppose workers have the current utility function

$$U = \ln(y_t^w) + \ln(1-x_{1t})$$

where y_t^w is the quantity of goods received by the worker and x_{1t} is the labor utilization rate. A base utility level, $\overline{U}$, has been negotiated. Then $\overline{y} = \exp(\overline{U})$ is the level of compensation required to attain the base utility level when labor hours are zero. For a fixed state of nature we want a labor compensation function, $V_1^{cd}(x_{1t})$, such that

$$\overline{U} = \ln(V_1^{cd}(x_{1t})) + \ln(1-x_{1t}) \ ;$$

that is, utility is constant at the base level for any degree of labor utilization x_{1t}. Exponentiating both sides and solving for V_1^{cd}, we have

$$V_1^{cd}(x_{1t}) = \frac{\exp(\overline{U})}{1-x_{1t}} = \frac{\overline{y}}{1-x_{1t}}$$

In this, the Cobb-Douglas case, the firm must pay $\overline{y}$ in period t just to keep workers with the firm, even if they are temporarily laid off ($x_{1t} = 0$); otherwise the workers would change firms and would not be available for recall. Note that $dV_1^{cd}/dx_{1t} > 0$, $d^2V_1^{cd}/dx_{1t}^2 > 0$, and V_1^{cd} approaches infinity as labor time required approaches twenty-four hours a day ($x_{1t} = 1$).

Compensation of incumbent workers, we see, depends on labor market conditions and the nature of existing contracts. We must also specify how new workers are hired, i.e., how the firm expands its stock of labor, L. We shall treat labor as a quasi-fixed factor of production (see Oi (1962)). As in the last section, we assume that there is a certain fixed cost, D, that must be borne by the firm in order to bring a new worker "online". D may be thought of a real hiring and training costs and may have both general and firm-specific components.

The firm employs capital services as well as labor services. Recall that capital services are the product of the firm's stock of machines, K_t, and its capacity utilization rate, x_{kt}. In the short run K_t is fixed, but x_{kt} may be varied. We postulate a (real) per-machine operating-and-maintenance cost function

$$V_k(x_{kt}, s_t)$$

whose arguments are the capacity utilization rate and the state of nature. Total O&M costs increase with utilization, so $\partial V_k/\partial x_{kt} > 0$. Marginal costs also increase with utilization (maintenance is more difficult, inferior equipment is pressed into service); we take $\partial^2 V_k/\partial x_{kt}^2 > 0$. The V_k function is analogous to the V_1 function derived for labor utilization. The state of nature appears as an argument for V_k not only to represent technical unknowns like machine reliability but to capture unspecified market phenomena like changes in the real cost of fuel or replacement parts.

Between periods the firm can expand its capital stock. It does this by transforming some of its own previous-period output into machines at rate p_k, which, for simplicity, we will take as being technologically given and constant. A firm's capital stock is non-depreciating, bolted down, and cannot be transformed back into the output good.

We have now specified all uses for the output good, which allows us to write down the income identity. For each firm:

$$y_t \equiv c_t^f + V_1(x_{1t}, s_t)L_t + V_k(x_{kt}, s_t)K_t + p_k^* I_{kt} + D^* I_{1t} \tag{2.4}$$

where

c_t^f = consumption by owners of the firm in period t (profits)

I_{kt} = additions to the capital stock in period t

I_{lt} = additions to the labor stock in period t

and

$$p_k^* = \begin{cases} p_k & , \text{ if } I_{kt} \geqq 0 \\ 0 & , \text{ if } I_{kt} < 0 \end{cases}$$

$$D^* = \begin{cases} D & , \text{ if } I_{lt} \geqq 0 \\ 0 & , \text{ if } I_{lt} < 0 \end{cases}$$

<u>Statics</u>. Let us analyze the short-run properties of this model. Within a single period, t, the firm's input stocks (K_t, L_t), its utilization cost functions (V_k, V_l), and the current state of nature, s_t, are given. All that needs to be determined is the current level of output, y_t.

The factor utilization rates necessary to produce a given y_t are:

$$(2.5) \qquad x_{lt} = y_t/L_t$$

$$x_{kt} = y_t/K_t$$

Total cost is thus given by:

$$(2.6) \qquad TC(y_t) = V_l(y_t/L_t, s_t)L_t + V_k(y_t/K_t, s_t)K_t$$

Differentiating with respect to y_t to obtain the marginal cost function and setting this equal to one (the "price" of y_t, the numeraire good) yields optimal output y_t^* as an implicit function of the parameters:

$$(2.7) \qquad \partial V_l/\partial x_l(y_t^*/L_t, s_t) + \partial V_k/\partial x_k(y_t^*/K_t, s_t) = 1 \ .$$

For an example of short-run output determination we turn again to the Cobb-Douglas case. We have already derived a labor compensation function for a worker with Cobb-Doublas utility:

$$V_l^{cd}(x_{lt}, s_t) = \bar{y}^w(s)/(1-x_{lt})$$

Symmetrically, suppose per-machine operating costs are given by:

$$V_k^{cd}(x_{kt}, s_t) = g(s)/(1-x_{kt})$$

Then, using (2.7), output y_t^* is implicitly defined by:

$$\bar{y}^w/(1-y_t^*/L_t)^2 + g/(1-y_t^*/K_t)^2 = 1 \ .$$

Let us imagine for a moment that this firm has access to an unlimited and costless supply of labor services. Then the labor cost term is zero, and y_t^* for this example can be written as

$$y_t^* = (1-g^{1/2})K_t$$

Note that output and labor services employed are finite here, even when labor supply is costless and infinite. Moreover, if $g > 0$, there will be unused capital capacity, despite the costless labor supply.* These two propositions are frequently true of this class of technologies, in the short run.

Dynamics. We have so far not specified the agent objective functions in this model economy. To do dynamics, we must be more explicit. We assume that there are no futures or contingency markets. Hence, profit maximization is not well defined. One way to characterize firm behavior in this situation is to develop a stock market story; this is the approach taken in Chapter 2. Here let us assume that there are two classes of identical individuals: firm-owners and workers. Firm-owners (there is one owner per firm) have intertemporal expected utility functions

$$U^f = E\left(\sum_{i=t}^{T} \beta^{i-t} U(c_i^f)\right) \tag{2.8}$$

where, from (2.4) and (2.5)

$$c_i^f = y_i^* - V_1(y_i^*/L_i, s_i)L_i - V_k(y_i^*/K_i, s_i)K_i - p_k^* I_{ki} - D^* I_{1i} \tag{2.9}$$

*If $g \geq 1$, the machines use up more than they produce, and $y_t^* = 0$.

and

(2.10) $$I_{ki} = K_{i+1} - K_i$$

$$I_{li} = L_{i+1} - L_i$$

That is, the firm-owner's utility depends on his consumption, c^f, in each period and state of nature. That consumption is the production of his firm, less current payments for capital and labor services, less outlays for increasing the stocks of capital and labor.

Worker utility functions enter only through the form of the V_l (labor compensation) function, so they are not set out explicitly. We assume that workers do not save or invest, but consume all of their current compensation. Firm-owners are able to save by increasing their input stocks. Money is excluded from the model.

The firm-owner's optimizing problem is to maximize (2.8) with respect to (2.9) and (2.10). His choice variables are I_{ki} and I_{li}, his planned additions to his input stocks; y_i^*, his optimal level of current output, is already given by (2.7). For an interior solution, the two necessary conditions for an optimal path are:

(2.11) (Investment in capital stock)

$$U'(c_i^f)p_k = E(\beta U'(c_{i+1}^f)\{\partial V_k/\partial x_k(x_k^*)x_k^* -$$

$$V_k(x_k^*) + p_k^*\})$$

(2.12) (Investment in labor stock)

$$U'(c_i^f)D = E(\beta U'(c_{i+1}^f)\{\partial V_l/\partial x_l(x_l^*)x_l^* -$$

$$V_l(x_l^*) + D^*\})$$

where

(2.13) $$x_k^* = y_{i+1}^*/K_{i+1}$$

$$x_l^* = y_{i+1}^*/L_{i+1}$$

Parenthesized objects in (2.11) and (2.12) are arguments of functions; braces indicate multiplication. Expectations are with respect to information available in period i. Note that the envelope theorem allows us to ignore the effect of small changes in K_{i+1} and L_{i+1} on optimal output, y_{i+1}^*.

The decision to increase input stocks is seen to hinge on three considerations: 1) the cost of stock increments (p_k,D), relative to current consumption; 2) the effect of the increment on next-period production costs (the terms in V and V'); 3) the long-term value of the increment, given future plans (p_k^*,D^*).

The first factor does not require much analysis. The real investment costs, p_k and D, are exogenously fixed and are the same in all periods. *Ceteris paribus*, then, investment will be low in periods of low current production (because the marginal utility of consumption is higher in those periods). Observe that this leads to a serial correlation of investment levels even when the underlying stochastic disturbances are independent over time; low production and investment this period means lower-than-trend production next period, therefore lower-than-trend investment. ("Investment", recall, means hiring and training new workers as well as adding to capital.)

The second factor considered is next-period production costs. In either the capital or labor cases, the production cost savings due to an increase in input stock can be written as

$$(2.14) \qquad x^* \frac{\partial V}{\partial x}(x^*,s) - V(x^*,s)$$

The first term represents a positive saving, arising from the fact that a higher stock allows for a lower average utilization rate and, therefore, lower operating cost/labor compensation expenses per unit of stock. The second term is negative (a cost dissaving); it occurs because a factor stock increment increases the number of units that require current expenditure.

Since one term is positive and one negative, we have the possibility that increased input stocks might raise production costs. Let us examine this with our Cobb-Douglas example.

Once more, let $V_1^{cd} = \bar{y}^w/(1-x_1)$, $V_k^{cd} = g(1-x_k)$. The cost of producing some given output y is

$$(\bar{y}^w/(1-y/L))L + (g/(1-y/k))K$$

Minimizing this total cost expression with respect to L and K is done by setting the marginal cost-savings expressions, $x^* \, \partial V/\partial x - V$, equal to zero. This gives

$$(y/L) \; (\bar{y}^w/(1-y/L)^2) - \bar{y}^w/(1-y/L) = 0$$

and a similar expression for capital. As long as costs are nonzero, the solutions do not depend directly on $\bar{y}^w$ or g:

$$L^* = K^* = 2y$$

or

$$x_1^* = y/L^* = 1/2$$

$$x_k^* = y/K^* = 1/2$$

The general expression for the optimal utilization rate is $x^* = V(x^*)/(\partial V/\partial x(x^*))$.

In this example the firm-owner will, other things being equal, adjust his capital and labor stocks so as to keep utilization rates close to one-half. Current deviations of x from 1/2 will occur because of dynamic considerations, however - a topic we now consider.

We look at the final element of the stock increment decision - its long-term value, contingent on future plans. The necessary conditions suggest that the long-term values can be expressed as $U'(c^f_{i+1})p^*_k$ or $U'(c^f_{i+1})D^*$, which can be viewed as the savings in future investment gained by undertaking investment now. A more illuminating way to examine this is as follows: Imagine a firm-owner in period t who has solved his stochastic dynamic programming problem for periods t = 1,...,T. This gives him the future optimal values of his investment in capital and labor, contingent on current investment decisions and the contemporaneous state of nature. Denoting optimal values with an asterisk, we write:

$$(2.15) \qquad I^*_{ki} = I^*_{ki}(I_{kt}, I_{lt}, s_i)$$

$$I^*_{li} = I^*_{li}(I_{kt}, I_{lt}, s_i)$$

(Initial conditions, K_t, L_t, and s_t are taken as fixed.) The firm-owner's optimization problem in period t is to maximize (2.8) subject to

$$(2.16) \qquad c_t^f = y_t^* - V_l(y_t^*/L_t, s_t)L_t - V_k(y_t^*/K_t, s_t)K_t$$

$$- p_k I_{kt} - DI_{lt}$$

$$c_{i+1}^f = y_{t+1}^* - V_l(y_{t+1}^*/L_t + I_{lt}) s_{t+1})(L_t + I_{lt})$$

$$- V_k(y_{t+1}^*/(K_t + I_{kt}), s_{t+1})(K_t + I_{kt})$$

$$- p_k I_{k,t+1}^* - DI_{l,t+1}^*$$

and

$$c_i^f = y_i^* - V_l(y_i^*/(L_t + I_{lt} + \sum_{j=t+1}^{i-1} I_{lj}^*), s_i) \cdot$$

$$\{(L_t + I_{lt} + I_{lj}^*)\} -$$

$$V_k(y_i^*/(K_t + I_{kt} + \sum_{j=t+1}^{i-1} I_{kj}^*), s_i) \cdot$$

$$\{(K_t + I_{kt} + I_{kj}^*)\} - p_k I_{ki}^* - DI_{li}^*$$

for $i = t+2, \ldots, T$.

Solving the maximization problem, we can again take repeated advantage of the envelope theorem, which allows us to think of y^*, I_l^*, and I_k^* as fixed for small changes in I_{lt} and I_{kt}. For an interior solution in period t, we have the conditions

$$(2.17) \qquad U'(c_t)p_k = E\Bigg(\beta U'(c_{t+1})\{x^*_{k,t+1}\frac{\partial V_k}{\partial x_k}(x_{k,t+1}, s_{t+1})$$

$$- V_k(x_{k,t+1}, s_{t+1})\} +$$

$$\sum_i \beta^{i-t} U'(c_i)\{x^*_{ki}\frac{\partial V_k}{\partial x_k} - V_k\}\Bigg)$$

$$(2.18) \qquad U'(c_t)D = E\Bigg(\beta U'(c_{t+1})\{x^*_{1,t+1}\frac{V_1}{x_1} - V_1\}$$

$$+ \sum_i \beta^{i-t} U'(c_i)\{ x^*_{1i}\frac{\partial V_1}{\partial x_1} - V_1\}\Bigg)$$

where

$$(2.19) \qquad x_{k,t+1} = y^*_{t+1}/(K_t + I_{kt})$$

$$x_{ki} = y^*_i/(K_t + I_{kt} + \sum_{j=t+1}^{i-1} I_{kj})$$

$$x_{1,t+1} = y^*_{t+1}/(L_t + I_{1t})$$

$$x_{1i} = y^*_i/(L_t + I_{1t} + \sum_{j=t+1}^{i-1} I_{1j})$$

$i = t+2, \ldots, T$

and the arguments of V are omitted after the first instance.

The interpretation of these conditions is as follows. Suppose the firm-owner is considering making a small (positive) change in his current input stock investment plans. Because of the envelope theorem, he can examine the effects of this infinitesimal change while taking his future production and investment plans as fixed. The left-hand sides of the above expressions are the costs of the proposed change; the right-hand sides, the benefits. The costs of increasing current investment are the marginal utilities of consumption foregone. The benefits are the sum of future production cost savings, appropriately weighted by marginal utilities and the discount factor. The interior solution equates the costs and benefits of capital or labor stock investment.

Let us note two facts. First, marginal benefits to current investment in either input are diminishing, and typically go to zero for a large enough investment. Positive cost savings are possible only so long as the factor utilization rates are above their optimal values. The marginal cost savings possible diminish as the utilization rates decline (i.e., as stocks expand) and become negative if the stock becomes so large that utilization rates fall below the optima.

Second, the marginal benefits to current investment depend on all future labor compensation and capital operating costs, in the following way. A decline (say) in expected compensation or operating costs for some period in the future, call it t', increases the expected optimal output for that period for any given level of input stocks. The increase in the utilization rate in t' increases the marginal value of input stocks in t', and, therefore, the marginal value of investment in the initial period, t. In general, current investment in either capital or labor stocks bears an inverse relation to future costs of labor compensation or capital operation.

We can now see the "dynamic considerations" that thwart the firm-owner in his ideal of always maintaining optimal utilization rates. First, reaching the optimal stocks may require too great a sacrifice of consumption in the short term. Second, if current production costs do not move smoothly over time, no smooth growth path of stocks will always yield the optimal utilization rate. Instead, firm-owners may have to carry excess stocks when costs are high in order to have them available in periods when costs are low.

Sources of unemployment. The model we have analyzed can by used to identify sources of unemployment in a given sector.

Note first that this model can produce unemployment of labor resources in two distinct ways: through a low current labor utilization rate, x_{1t}, or through a low stock of firm-affiliated labor, L_t.

A low utilization rate may be interpreted variously as "short hours" or as "layoffs with recall". In this model, given input stocks, costs, and demand, - and, also given the assumption that shocks to the labor supply curve can be neglected, - labor utilization within the firm is Walrasian. This is because of the Hall-Lillien provision that insures equality of the marginal product and marginal disutility of labor for the workers covered by the contract.

The sources of layoffs are current changes in 1) the value of the output good, 2) marginal operating costs of capital, 3) costs of materials and other variable inputs, and 4) the opportunity cost of leisure. Of these, changes in the value of the output good are probably most important empirically, at least over the cycle. The other factors may be important secularly. For example, the work week has declined steadily with increases in the opportunity costs of working, while the oil shock may have permanently lowered utilization rates (at least for workers using the current vintage of capital).

The other form of labor resources unemployment predicted by the model occurs when L_t, the number of workers who are firm-affiliated, is less than the total available labor force. Our analysis has shown that the rate of increase of L_t depends on three sets of factors:

1) The cost of creating new jobs relative to current levels of consumption. Sometimes, given social discount rates, the capital stock cannot increase fast enough to employ a growing labor force. (This argument requires a lower bound on the _ex ante_ capital-labor ratio.) This may frequently be the case in developing economies; there is insufficient surplus for the number of manufacturing jobs to keep pace with labor supply.

This type of unemployment is much less likely in a developed economy; we expect that sufficient savings will be available to match the capital stock to the labor force. It is possible, however, that the recent unemployment experience of the United States is partly attributable to the short-term inability of the economy to absorb a substantial spurt in labor supply growth.

2) The impact of additional workers on short-run production costs. When the firm-owner puts new workers on the roster, he is incurring fixed costs in order to reduce his variable costs. Steeply increasing variable

costs (e.g., high overtime charges) motivate additional hires. Large fixed costs (for example, if training is expensive, or if laid-off workers still receive a large part of their base pay) reduce the number of workers put on the payroll.

The degree of risk-aversion in the firm-owner's single-period utility function affects this tradeoff. If the dominant stochastic factor in the environment is changes in produce demand, for example, more risk-averse awners will maintain smaller labor forces in the short run. This has the advantage of creating higher labor costs in high-utilization states and lower labor costs in low-utilization states relative to the large-labor-force firm, leaving the risk-averse owner a smaller variance of new consumption.

3) The long-term value of an additional worker, given future plans. The decision to expand the labor force instead of increasing utilization rates depends critically on the expected long-run situation. If student enrollments are secularly declining, for example, the school board would rather pay overtime to the existing staff than hire and train new teachers. Similarly, an upward trend in real energy prices promises lower future utilization rates, leading to lower optimal employment today. The magnitude of these effects depends

in part on the firm-owner's discount rate (his willingness to trade tomorrow's costs for today's.)

In general, both temporary and secular shifts in costs and in the demand for output keep capital and labor stocks "mismatched" in the short run. As stocks are adjusted to reflect average long-run needs, the short run will be characterized by alternating periods of high utilization and stock hoarding.

We see that even a relatively simple model yields a variety of sources of labor unemployment. Many of the types of unemployment we have analyzed would, in the real world. find their way into the national statistics, perhaps to be thought of as "disequilibrium" in the labor market. This, of course, is largely a problem with the way unemployment statistics are collected. The deeper questions are: How much of this unemployment would be perceived as "involuntary" by workers in the market? And what is the role of contracting in perceived disequilibrium?

It will be useful to separate the effects of contract incompleteness, as described in Part I, from the other sources of unemployment. Suppose that contracts are complete, in that workers can accept negative wages and bind themselves to stay with a specific firm. In this circumstance both principal types of unemployment of labor resources would still exist. The difference would be that any worker willing to undertake the costs of training (either directly or by paying the firm) could find a job. Walrasian allocations would result, both within the firm's internal labor market and in the market for new workers.

Despite this outcome, it is possible that workers might (erroneously) report that there is a disequilibrium in the labor market. It is true, for example, that a new worker will not be able to find a job by offering to work for a wage marginally below that of someone already employed. The already-employed worker's wage includes a return to the equity he holds in his own training. Since that training may not involve a diploma but rather (say) a knowledge of the firm's inventory, the unemployed worker may think of himself as identical to the fellow already in the job. He will feel "involuntarily unemployed" if his reservation wage is near the wage received by the employed worker.

If contracts are incomplete, the perception of labor market disequilibrium becomes more acute. As shown in Part I, the prospective entrant's starting wage offer must undercut the wage of experienced workers by the full differential cost of training. The entrant may in fact have no positive wage offer that would get him a job, even though employed workers are paid a positive wage. This would almost surely be construed as a disequilibrium. This situation is non-Walrasian and inefficient; however, it is not, as we have seen, a true disequilibrium, in that there are no unexploited opportunities which will tend to change the outcome.

We offer two conclusions. 1) With respect to the "contracting debate": Under any circumstances, the labor market is likely to experience unemployment, which may be perceived as involuntary. (Equivalently, wages will be thought to be "sticky".) If contracts are incomplete, allocations may be inefficient. Explanations of these phenomena, however, need not rely on "persistent disequilibria", or failure of the private sector to exploit opportunities. Labor contracting can be absolved from the claim that it "causes" serially correlated unemployment.

2) As a matter of policy-making, a fixed unemployment rate is probably not a good target for the economy as a whole to shoot for. Other research has already made this point (for example, the search literature). This paper has given reasons why the optimal work-leisure split may vary over time in a given sector. The previous Chapter discussed still other sources of variable (but efficient) utilization of resources. We conclude

that there is no _prima facie_ case for forcing the unemployment rate to stay in a narrow range.

While the appropriate goals of policy that affect the aggregate unemployment rate are not clear, manpower policy should be active. To the extent that contracts are incomplete, worker training has characteristics of a public good, and should therefore be subsidized.

General Bibliography

Abel, A., *Investment and the Value of Capital*, Ph.D. Thesis, M.I.T., 1978.

Aoki, Masanao, *Optimization of Stochastic Systems*, New York: Academic Press, 1967.

Arrow, K.J., and A.C. Fisher, "Environmental Preservation, Uncertainty, and Irreversibility", *Quarterly Journal of Economics*, Vol. 88 (May 1974), pp. 312-20.

Arrow, K.J., and Mordecai Kurz, *Public Investment, the Rate of Return, and Optimal Fiscal Policy*, Baltimore: The Johns Hopkins Press, 1970.

Arrow, K.J., "Optimal Capital Policy and Irreversible Investment", in J.N. Wolfe, ed., *Value, Capital, and Growth*, Chicago: Aldine Publishing Co., 1968, pp. 1-20.

Azariadis, C., "Implicit Contracts and Underemployment Equilibria", *Journal of Political Economy*, Vol. 83 (December 1975), pp. 1183-1202.

Baily, M.N., "Wages and Employment Under Uncertain Demand", *Review of Economic Studies*, Vol. 41, #1 (January 1974), pp. 37-50.

Barro, R.J., "Rational Expectations and the Role of Monetary Policy", *Journal of Monetary Economics*, Vol. 2 (January 1976), pp. 1-33.

Barro, R.J., "Long-term Contracting, Sticky Prices, and Monetary Policy", *Journal of Monetary Economics*, Vol. 3 (July 1977), pp. 305-16.

Barro, R.J., and H.I. Grossman, "A General Disequilibrium Model of Income and Employment", American Economic Review, Vol. 61, #1 (March 1971), pp. 82-93.

Bellman, R.E., and S.E. Dreyfus, Applied Dynamic Programming, Princeton, N.J.: Princeton University Press, 1962.

Clower, R., "The Keynesian Counter-Revolution: A Theoretical Appraisal", in F.H. Hahn and F.P.R. Brechling, eds., The Theory of Interest Rates, London, 1965.

Cyert, R.M., and M.H. DeGroot, "Rational Expectations and Bayesian Analysis", Journal of Political Economy, Vol. 82 (1974), pp. 521-36.

DeGroot, M.H., Optimal Statistical Decisions, New York: McGraw-Hill, 1970.

Fischer, S., "Long-term Contracts, Rational Expectations, and the Optimal Money Supply Rule", Journal of Political Economy, Vol. 85, #1 (February 1977), pp. 191-205.

Fischer, S., "Long-term Contracting, Sticky Prices, and Monetary Policy", Journal of Monetary Economics, Vol. 3 (July 1977), pp. 317-23.

Grossman, H.I., "Risk-shifting and Reliability in Labor Markets", Scandinavian Journal of Economics, 1977.

Hall, R.E., and D.W. Jorgenson, "Tax Policy and Investment Behavior", American Economic Review, Vol. 57, #3 (June 1967), pp. 391-414.

Hall, R.E., and D. Lillien, "Efficient Wage Bargains Under Uncertain Supply and Demand", unpublished ms., June 1977.

Hart, A.G., "Risk, Uncertainty, and the Unprofitability of Compounding Probabilities", in O. Lange *et.al.* eds., *Studies in Mathematical Economics and Econometrics in Memory of Henry Schultz*, Chicago: University of Chicago Press, 1942, pp. 110-18.

Howard, Ronald A., *Dynamic Probabilistic Systems*, New York: John Wiley and Sons, 1971.

Howard, Ronald A., "Dynamic Inference", *Research in the Control of Complex Systems*, T.R. #10, M.I.T., 1964.

Kaufman, G.M., "Sequential Investment Analysis Under Uncertainty", *Journal of Business*, Vol. 36, #1 (1963), pp. 39-64.

Leijonhufvud, A., *On Keynesian Economics and the Economics of Keynes*, New York, 1968.

Lucas, R.E., "Econometric Policy Evaluation: A Critique", in *The Phillips Curve and Labor Markets*, Karl Brunner, ed., supplement to the *Journal of Monetary Economics*, April 1976.

McCall, John J., "The Economics of Information and Optimal Stopping Rules", *Journal of Business*, Vol. 38, #3 (1965), pp. 300-17.

Merton, R.C., "Lifetime Portfolio Selection Under Uncertainty: the Continuous-time Case", *Review of Economics and Statistics*, Vol. 51, #2 (May 1969), pp. 247-57.

Murphy, Roy E., Adaptive Processes in Economic Systems, New York: Academic Press, 1965.

Oi, W., "Labor as a Quasi-fixed Factor:, Journal of Political Economy, Vol. 70 (December 1962).

Rothschild, Michael, "Searching for the Lower Price When the Distribution of Prices is Unknown", Journal of Political Economy, Vol. 82, #4 (July 1974), pp. 689-712.

Sage, Andrew P., Optimum Systems Control, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

Samuelson, Paul, "Lifetime Portfolio Selections by Dynamic Stochastic Programming", Review of Economics and Statistics, Vol. 51, #2 (May 1969), pp. 239-57.

Wachter, M.L., and O.E. Williamson, "Obligational Markets and the Mechanics of Inflation", Bell Journal of Economics, Vol. 9, #2 (Autumn 1978), pp. 549-71.

Williamson, O.E., M.L. Wachter, and J.E. Harris, "Understanding the Employment Relation: the Analysis of Idiosyncratic Exchange", Bell Journal of Economics, Vol. 6, #1 (Spring 1975), pp. 250-78.

Biographical Note

Ben Shalom Bernanke

Date of birth: December 13, 1953

Place of Birth: Augusta, Georgia

Marital status: Married to Anna Friedmann, May 29, 1978

Education: B.A., *summa cum laude*,
Harvard University, 1975

Awards: Phi Beta Kappa
National Science Foundation Graduate Fellow