NOTES ON MACROECONOMIC THEORY

by

Thomas J. Sargent

Discussion Paper No. 75-56, August 1975

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

PREFACE

This manuscript is a rough and unfinished set of lecture notes about macroeconomics. At many points, the notes are very tentative and amount to "thinking out loud." This is especially true of my remarks on consumption, a topic whose econometrics seem to me yet to be worked out satisfactorily. The notes on stochastic difference equations are very incomplete and are eventually intended to lay the foundation for a subsequent discussion of the business cycle theories implicit in modern macroeconometric models, together with an exposition of Tinbergen-Theil control theory.

The theory in these notes is "old fashioned," mostly being nonstochastic and embodying ad hoc theories of expectations. Perfect foresight models are discussed at several points, however, these being the counterparts of "rational" expectations models in a nonstochastic setting.

These notes summarize macroeconomic as opposed to general equilibrium models. By the latter one means complete and consistent models derived from explicit statements of the objectives, opportunities and information available to individual agents. In contrast, macroeconomics typically proceeds by postulating a set of behavior relations, relations often suggested by raw statistical correlations, and proceeds to analyze the implications of the interactions of a set of these relations. For example, large macroeconometric models are often constructed by combining the final products of a consumption function expert, an inventory expert, a financial sector expert, and so on. There is no assurance that there exists a deep, coherent general equilibrium model compatible with these

efforts: certainly there has rarely been one made explicit. The models
in these notes exhibit this limitation, a limitation that I think ought
to be taken seriously.

Helpful comments on these notes were made by several of my
colleagues at the Federal Reserve Bank of Minneapolis, especially Preston
Miller, Gary Skoog, and Arthur Rolnick. Much of the first four "chapters"
merely represents a writing down of things Neil Wallace has told me at
one time or another.

<div align="right">Thomas J. Sargent</div>

CONTENTS

Introduction

These pages present static and dynamic analyses of some standard macroeconomic models. By _static_ analysis we mean the analysis of events assumed to occur at a point in time. In effect, statics studies the alternative point-in-time or momentary equilibrium values for a set of _endogenous variables_ associated with alternative possible settings for the _exogenous variables_ at the particular point in time under consideration. Endogenous variables are those determined by the model at hand, while exogenous variables are those given from outside the model.

The task of _dynamics_ is to study the time paths of the endogenous variables associated with alternative possible time paths of the exogenous variables. Thus, in a dynamic analysis, the behavior of a model is studied as time is permitted to pass. In contradistinction, in a static analysis, attention is confined to events assumed to occur instantaneously; i.e., at a given moment.

A third kind of analysis, that of _stationary_ states, is a limiting form of dynamic analysis, and is directed toward establishing the ultimate tendencies of certain endogenous variables, such as the capital-output ratio, as time passes without limit and as certain critical exogenous variables remain constant through time. Stationary analysis ought not to be confused with statics.

The distinguishing feature of a static analysis is that it is capable of determining alternative values of the endogenous variables, taking as given only the values of the exogenous variables at that point in time, which may include values of endogenous and exogenous variables which were determined in the past and are thus given or predetermined at the present moment. As we shall see, some models for which a dynamic

analysis is possible simply cannot be subjected to static analysis. In order to perform static experiments, it is necessary partly to divorce current events from future events so that what happens in the future does not affect what happens now. This requires restricting the way in which people are assumed to form expectations about the future, and in particular requires that people not possess perfect foresight.

Generally, our models will consist of n _structural_ equations in n endogenous variables $y_i(t)$, i=1,...,n and m exogenous variables $x_i(t)$, i=1,...,m:

(1) $$g_i(y_1(t), y_2(t),...,y_n(t), x_1(t),...,x_m(t)) = 0, \quad i=1,...,n.$$

A _structural equation_ summarizes behavior, an equilibrium condition, or an accounting identity, and constitutes a building block of the model. In general, more than one, and possibly all n endogenous variables can appear in any given structural equation. The system of equations (1) will be thought of as holding at each moment in time t. Time itself will be regarded as passing continuously, so that t may be regarded as taking all values along the (extended) real line.

The exogenous variables $x_i(t)$, i=1,..., m are assumed to be right-continuous functions of time, and furthermore are assumed to possess right-hand time derivatives of at least first, and sometimes higher order at all points in time. By right-continuity of the functions $x_i(t)$ we mean

$$\lim_{\substack{t \to \bar{t} \\ t > \bar{t}}} x_i(t) = x_i(\bar{t}),$$

so that $x_i(t)$ approaches $x_i(\bar{t})$ as t approaches $\bar{t}$ from above; i.e., from

ii

the future. However, the function $x_i(t)$ can jump at $\bar{t}$, so that we do

<u>not</u> require

$$\lim_{\substack{t \to \bar{t} \\ t < \bar{t}}} x_i(t) = x_i(\bar{t}) \;.$$

For example, consider the function

$$x_i(t) = 0 \qquad , t < \bar{t}$$

$$x_i(t) = 1 \qquad t \geq \bar{t} \;,$$
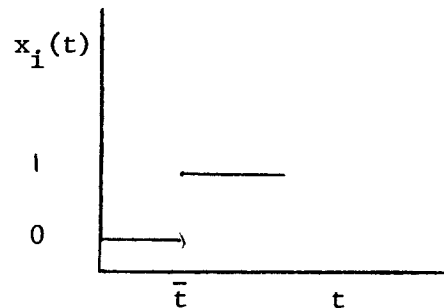
Figure 1

which is graphed in Figure 1. It is right-continuous everywhere even

though it jumps; i.e., is discontinuous, at $\bar{t}$.

The right-hand time derivative of $x_i(t)$, which is assumed to

exist everywhere, is defined as

$$\frac{d}{dt} x_i(\bar{t}) \equiv \lim_{\substack{t \to \bar{t} \\ t > \bar{t}}} \frac{x_i(t) - x_i(\bar{t})}{t - \bar{t}} \;.$$

For the function graphed in Figure 1, the right-hand derivative is zero

everywhere, even though the function jumps and hence isn't differentiable

at $t = \bar{t}$.

A model is said to be in static <u>equilibrium</u> at a particular

moment if the endogenous variables assume values that assure that equations

(1) are all satisfied. Notice that it is <u>not</u> an implication of this

definition of equilibrium that the values of the endogenous variables

are unchanging through time. On the contrary, since the values of the

exogenous variables will in general be changing at some nonzero rates

per unit time, the endogenous variables will also be changing over time.

Static analysis is directed toward answering questions of the

following form. Suppose that one of the exogenous variables $x_i(t)$ takes

a (small) jump at time $\bar{t}$ so that

$$\lim_{\substack{t \to \bar{t} \\ t < \bar{t}}} x_i(t) \neq x_i(\bar{t}) \quad .$$

Then the question is to determine the responses of the endogenous variables at $\bar{t}$. The distinguishing characteristic of endogenous variables is that each of them is assumed to be able to jump discontinuously at any moment in time in order to guarantee that system (1) remains satisfied in the face of jumps in the $x_i(t)$'s. Thus, to be endogenous from the point of view of statics, a variable must be able to change instantaneously. Notice that it is possible for the right-hand time derivative of a variable to be endogenous, i.e., to be capable of jumping discontinuously, even though the variable itself must change continuously through time (Figure 2 gives an example). One way to view the difference between the classical and Keynesian models is that in the former the money wage is a variable in static experiments, while in the latter the right-hand time derivative of the money wage is a variable but the level of the money wage is exogenous.
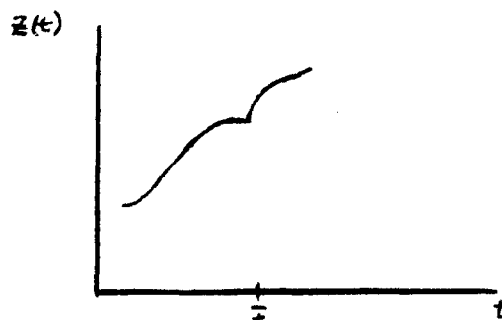
Figure 2: $z(t)$ is continuous, but the right-hand derivative jumps at $\bar{t}$.

To answer the typical question addressed in statics, the reduced form equations corresponding to the system (1) must be found. The reduced form equations are a set of equations, each expressing one

$y_i(t)$ as a function only of the $x_i(t)$'s:

(2) $y_i(t) = h_i(x_1(t), x_2(t),\ldots,x_m(t))$ $i=1,\ldots,n.$

We will generally assume that the functions $g_i(\ )$ in the structural equations (1) are continuously differentiable in all directions, that the n structural equations were satisfied at all moments immediately preceding the moment we are studying and that a certain function of the partial derivatives of (1), evaluated at the immediately preceding values of all variables, is not zero. To be more precise, we shall assume the hypotheses of the implicit function theorem.[*] Under these hypotheses, there exist continuously differentiable functions of the reduced form (2) which hold for $x_i(t)$'s sufficiently close to the initial (prejump) values of the $x_i(t)$'s. If these equations (2) are satisfied, we are guaranteed that the structural equations (1) are satisfied. For jumps in $x_i(t)$ sufficiently small, i.e., within the neighborhood identified in the implicit function theorem, the equations (2) hold and can be used to answer the characteristic question posed in static analysis. In particular, the reduced form partial derivative

(3) $\dfrac{\partial y_i(t)}{\partial x_j(t)} = \dfrac{\partial h_i}{\partial x_j(t)} (x_1(t),\ldots,x_n(t))$

gives the response of $y_i(t)$ to a jump in $x_j(t)$ that occurs at t. We are generally interested in the sign of the partial derivative of the reduced form.

---

[*]See, e.g. Taylor and Mann, <u>Advanced Calculus</u>, Second Edition, p. 363.

Rather than using the implicit function theorem directly to calculate the reduced form partial derivatives (3), it will be convenient to use the following alternative technique that always gives the correct answer. First, take the differential of all equations in (1) to obtain

(4) $\qquad \dfrac{\partial g_i}{\partial y_1} dy_1 + \ldots + \dfrac{\partial g_i}{\partial y_n} dy_n + \dfrac{\partial g_i}{\partial x_1} dx_1 + \ldots + \dfrac{\partial g_i}{\partial x_m} dx_m = 0,$

$\qquad\qquad\qquad i=1,\ldots n.$

all partial derivatives being evaluated at the initial values of the x_i's and y_j's. Then by successive substitution eliminate $dy_2, \ldots dy_n$ from the above system (4) of linear equations to obtain an equation of the form

(5) $\qquad dy_1 = f_1^{\ 1} dx_1 + f_2^{\ 1} dx_2 + \ldots + f_m^{\ 1} dx_m$

where the $f_j^{\ 1}$'s are functions of the partial derivatives appear in (4). Now equation (5) is the total differential of the reduced form for y_1, since dy_1 is a function only of $dx_1, \ldots, dx_m$. Taking the differential of the first equation of (2) gives

(6) $\qquad dy_1 = \dfrac{\partial h_1}{\partial x_1} dx_1 + \ldots + \dfrac{\partial h_m}{\partial x_m} dx_m .$

From (6) and (5) it therefore follows that

$\qquad f_j^{\ 1} = \dfrac{\partial h_1}{\partial x_j} \quad$ for $j=1,\ldots,n,$

so that the $f_j^{\ 1}$'s are the reduced form partial derivatives. Successive substitution in the system (4) will also, of course, yield the differentials of the reduced forms for the other endogenous variables, thereby enabling us to obtain the corresponding reduced form partial derivatives. The reduced form partial derivatives are often called "multipliers" in macroeconomics.

Notes on the "Classical" Model

Our model describes the determination of an economy's rate of output and the uses to which it is put. The economy produces a single good, which is produced at a rate per unit time of Y. This rate of output is divided among a real rate of consumption C, a real rate of investment I, a real rate of government purchases G, and a real rate of depreciation of capital δK:

(0) $\qquad Y = C + I + G + \delta K$

Equation (0) is the national income identity linking aggregate output and its components.

The economy is organized into three sectors. <u>Firms</u> employ capital and labor to produce output. The <u>government</u> taxes and purchases goods, issues money and bonds, and conducts open market operations. <u>Households</u> own the government's money and bond liabilities and all of the equities of firms. They make both a saving decision and a decision to allocate their portfolios of paper assets among bonds, equities, and money.

Firms

The economy consists of a large number of n perfectly competitive firms, each of which produces the same single good subject to the same production function. The rate of output of the ith firm at any instant is described by the instantaneous production function

$$(1) \qquad Y_i = F(K_i, N_i) \qquad\qquad i=1, \ldots, n$$

where Y_i is the output of the ith firm per unit time, K_i is the stock of capital employed by the ith firm, and N_i is employment of the ith firm. The variables Y_i, K_i, and N_i should each be thought of as functions of time. We have omitted a subscript i from the function F because it is assumed that all firms share the same production function. The production function is assumed to be characterized by positive though diminishing marginal products of capital and labor and a direct dependence of the marginal product of capital (employment) on employment (capital):

$$F_K, \; F_N > 0; \; F_{KK}, \; F_{NN} < 0, \; F_{KN} > 0.$$

The production function F is assumed to be linearly homogeneous in K_i and N_i, so that

$$\lambda Y_i = F(\lambda K_i, \lambda N_i) \quad , \qquad \lambda > 0.$$

By virtue of Euler's theorem on homogeneous functions we have

$$Y_i = \frac{\partial Y_i}{\partial K_i} (K_i, N_i) K_i + \frac{\partial Y_i}{\partial N_i} (K_i, N_i) \; N_i \; .$$

Also, by virtue of the linear homogeneity of F we have

$$\frac{\partial Y_i}{\partial K_i} = \frac{\partial F}{\partial \lambda K_i} (\lambda K_i, \lambda N_i) \; ;$$

setting $\lambda = 1/N_i$, we have

$$\frac{\partial Y_i}{\partial K_i} = \frac{\partial F}{\partial (\frac{K_i}{N_i})} (\frac{K_i}{N_i}, 1) \quad,$$

so that the marginal product of capital depends only on the ratio of capital to labor. Similarly, the marginal product of labor depends only on the ratio of capital to labor.

In this one-good economy, capital represents the accumulated stock of the one good which is available to assist in production. We assume that at any moment, the stock of capital is fixed both to the economy and to each individual firm. Assuming that capital is fixed to the economy amounts to ruling out once-and-for-all gifts of physical capital from abroad or from heaven and once-and-for-all decreases in the capital stock due to natural or human disasters. Assuming that capital is fixed to each firm at each moment in time amounts to ruling out the existence of a perfect market in the existing stock of capital in which individual firms can purchase or sell (or rent) capital, and so effect a discrete change in their stock of capital at a moment in time. The absence of a market in existing capital might be rationalized by positing that once is place, capital becomes completely specialized to each firm. Firms simply have no use for the existing capital of another firm, so that there is no opportunity for making a market in existing capital. Regardless of how the assumption is rationalized, however, ruling out trading of existing stocks of capital is a fundamental feature of the class of "classical" and "Keynesian" models that we will be describing.

While firms can't trade capital at a point in time, they are assumed to be able to vary employment instantaneously. Firms operate in

a competitive labor market in which at any moment they can hire all the labor they want at the going money wage w. Firms are perfectly competitive in the output market also, and each can sell output at any rate it wishes at the price of the one good in the model p.

The typical firm's profits Π_i are defined by

$$(2) \qquad \Pi_i = pF(K_i, N_i) - wN_i - (r+\delta-\Pi)pK_i$$

where r is the instantaneous rate of interest on government bonds, δ is the instantaneous rate of physical depreciation of capital, and Π is the anticipated rate of increase in the price of (newly produced) capital goods. In a sense to be defined below, $(r+\delta-\Pi)$ is the appropriate cost of capital that should be used to define the firm's profits. Were there a rental market in capital, $(r+\delta-\Pi)p$ would be the rental rate, expressed in dollars per unit time.

Each firm maximizes its profits per unit time with respect to the employment of labor, taking its capital stock as fixed momentarily. The firm's employment is then described by the first-order condition for maximization of (2),

$$\frac{\partial \Pi i}{\partial N_i} = pF_{N_i}(K_i, N_i) - w = 0$$

or

$$(3) \qquad F_{N_i}(K_i, N_i) = \frac{w}{p} \quad ,$$

which states that the firm equates the marginal product of labor to the real wage. Equation (3) is in the nature of a firm's demand function for labor which, given K_i, relates the firm's demand for employment inversely to the real wage. For each firm, equation (3) determines a capital-labor ratio, which is identical for all firms since all face a common real wage.

At any moment, the n firms have amounts of capital K_i, i=1, ..., n, which might differ across firms. Employment of labor then varies proportionately with K_i across firms.

Our assumptions about the identity of firms' production functions and their profit-maximizing behavior in the face of perfectly competitive markets for output and labor imply that there is a useful sense in which there exists an aggregate production function. The total rate of output of the one good in the economy is Y defined by

$$Y = \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} F(K_i, N_i) \quad .$$

By Euler's theorem, we have

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} (F_{K_i}(K_i, N_i)K_i + F_{N_i}(K_i, N_i)N_i)$$

But since the marginal products of capital and labor depend only on the capital-labor ratio, and since that ratio is the same for all firms, the marginal products of capital and of labor, respectively, are the same for all firms. Thus, we can write

$$\sum_{i=1}^{n} Y_i = F_K(\frac{K_i}{N_i}, 1) \sum_{i=1}^{n} K_i + F_N(\frac{K_i}{N_i}, 1) \sum_{i=1}^{n} N_i \quad .$$

Since the ratios K_i/N_i are the same for all n firms, they must be equal to the ratio of capital to employment for the economy, $\sum_{i=1}^{n} K_i / \sum_{i=1}^{n} N_i$. Consequently we have

$$Y = F_K(\frac{K}{N}, 1)K + F_N(\frac{K}{N}, 1)N$$

where $K = \sum_{i=1}^{n} K_i$ and $N = \sum_{i=1}^{n} N_i$. But by applying Euler's theorem to F, the above expression for Y can be written as the aggregate production

function

(4) $Y = F(K, N)$

Moreover, notice that $\frac{\partial Y}{\partial N}$ equals the marginal product of labor for each

firm while $\frac{\partial Y}{\partial K}$ equals the marginal product of capital for each firm. This

fact makes it legitimate to carry out our subsequent analysis solely in

terms of the aggregate production function (4) and the equality between

the real wage and $\partial Y/\partial N$:

(5) $F_N(N, K) = \frac{w}{p}$.

Equation (5) could be derived by maximizing economy-wide profits with

respect to employment.

Notice that (4) is a valid description of the relationship

among Y, N, and K only for a certain distribution of N_i across firms,

the one predicted by equation (3). That distribution is one that maxi-

mizes Y for any given N. If some other distribution of the N_i's across

firms is imposed, one that violates (3), then equation (4) will not des-

cribe the relationship among aggregate output Y, aggregate employment N,

and the aggregate capital stock K. For our work, however, it is suffi-

cient that (4) hold only in the sense described above.

It remains to describe the behavior of firms with respect to

the accumulation of captial over time. If there were a perfect market

in which firms could trade capital at each moment, firms would want to

purchase (sell) capital instantaneously as long as the marginal product

of capital exceeded (was less than) the real cost of capital, r+δ-Π.

However, such trading of capital has been ruled out. In its place we

posit a Keynesian investment demand function on the part of firms. This

function relates firms' demand to accumulate newly produced capital at some finite rate per unit time directly to the gap between the marginal product of capital and the cost of capital:

$$(6) \qquad \frac{dK}{dt} \equiv I = I(\frac{Y_K - (r+\delta-\Pi)}{r-\Pi}) \qquad\qquad I' > 0.$$

where $\frac{dK}{dt}$ is interpreted as a right-hand derivative. According to (6), firms invest at a higher rate the higher is the marginal product of capital and the lower is the real interest rate $r-\Pi$. Equation (6) describes aggregate investment demand for the economy, and is assumed to have been derived from individual firms' investment demand functions of the same form. We will find it convenient to write (6) in the compact form

$$(6') \qquad I = I(q-1) \qquad\qquad I' > 0$$

where q is defined by

$$(7) \qquad q = \frac{Y_K - (r+\delta-\Pi)}{r-\Pi} + 1 \equiv q(K, \ N, \ r-\Pi, \ \delta)$$

We will shortly provide an interpretation of q as an important relative price that might plausibly govern firms' demand to accumulate capital. Notice that q is a function of K and N by virtue of the dependence of the marginal product of capital on the labor-capital ratio.

Assets Owned by Households

There are three paper assets that households alone own: money, bonds, and equities. Money, the quantity of which is denoted by M, measured in dollars, is a paper asset that is supposed to be used as the medium of exchange. It is issued by the government and bears a nominal yield that is fixed at zero. By nominal yield we mean the yield of the asset in percent per unit time that can be obtained while leaving the nominal quantity of the asset intact. By the real yield of an asset we mean the yield in percent per unit time that can be obtained while setting aside enough resources to keep the real stock of the asset held intact. The nominal yield on money is fixed at zero because holding money gives rise to no payments of interest. However, the real yield on money in general is not zero. The real quantity of money is M/p, a quantity measured in units of output. The time derivative of M/p is

$$\frac{d\left(\frac{M}{p}\right)}{dt} = \frac{p\dot{M} - M\dot{p}}{p^2} = \frac{\dot{M}}{p} - \frac{M}{p}\frac{\dot{p}}{p} \quad .$$

To keep M/p intact over time, it is necessary to set the above derivative equal to zero, which gives

$$\frac{\dot{M}}{p} = \frac{M}{p} \cdot \frac{\dot{p}}{p}$$

or

$$\frac{\dot{M}}{M} = \frac{\dot{p}}{p}$$

Thus to keep real money balances M/p intact, it is necessary to add to nominal money balances at the rate $\dot{p}/p$. Consequently money has a real yield of $-\dot{p}/p$. That is, with $\dot{M}$ equal to zero, real balances depreciate at the rate $\dot{p}/p$ per unit time. People don't necessarily preceive the rate $\dot{p}/p$ at which prices are depreciating, since that would in general require perfect foresight. We denote the rate per unit time at which people expect the price level to increase as Π, which can differ from $\dot{p}/p$. The expected real rate of yield on money then equals $-\Pi$.

The second asset is a variable-coupon bond that is issued by the government. The bond is essentially like a savings deposit, changes in the interest rate altering the coupon but leaving the dollar value of bonds outstanding unchanged. We denote the nominal value of bonds outstanding B, which is measured in dollars. The bonds bear a nominal yield of r in percent per unit time. Thus the bonds throw off a stream of interest payments of $r \cdot B$ which is measured in dollars per unit time. The <u>real yield</u> on bonds is defined as r minus the percentage real rate at which households must buy bonds to keep their real value B/p intact. Like money, the real value of a fixed nominal quantity of bonds depreciates at the rate $\dot{p}/p$. Thus, the expected real rate of return associated with holding bonds is $r-\Pi$.

The third paper asset consists of equities, which are issued by firms in order to finance investment. We assume that firms issue no bonds and retain no earnings, so that all investment is financed by issuing equities. Only households hold, buy, and sell equities. By assuming that there is no market in which firms can purchase or sell physical capital, we are obliged also to rule out the possibility that

firms trade equities, which are the financial counterpart of physical capital. We assume also that households regard equities and bonds as perfect substitutes. This implies that their expected real yields will be equal, an equality enforced by investors' refusing to hold the lower yielding asset should the equality not hold. It follows that the nominal bond rate of interest is the pertinent yield for discounting expectations of firms' net cash flow (which equals expected aggregate dividends) in order to determine the value of firms' equities. At instant s, firms pay out a flow of dividends

$$p(s)Y(K(s), N(s)) - w(s)N(s) - \delta p(s)K(s) \quad ,$$

measured in dollars per unit time. Then the nominal value of firms' equities at instant t, denoted by $V(t)$, is

$$V(t) = \int_t^\infty [p(s)Y(K(s), N(s))-w(s)N(s)-\delta p(s)K(s)]e^{-r(s-t)}ds \quad .$$

We assume that firms and households expect the price level and wage rate to be following the paths

$$p(s) = p(t)e^{\Pi(s-t)}$$

$$w(s) = w(t)e^{\Pi(s-t)}$$

where Π is the anticipated rate of inflation. We further assume that the public expects the real rate of dividends to remain unchanged over time at the current rate. Then $V(t)$ can be written

$$V(t) = [p(t)Y(K(t), N(t))-w(t)N(t)-p(t)\delta K(t)] \int_t^\infty e^{-(r-\Pi)(s-t)}ds$$

or

$$V(t) = \frac{P(t)Y(t)-w(t)N(t)-p(t)\delta K(t)}{r-\Pi} \quad .$$

It is easy to see that $V(t)$ can be rewritten as

$$V(t) = \frac{p(t)\left(Y(t)-\frac{w(t)}{p(t)}N(t)-Y_K(t)K(t)\right)}{(r-\Pi)} + \frac{\left(Y_K-(r+\delta-\Pi)\right)p(t)K(t)}{r-\Pi} + p(t)K(t) \quad .$$

But by the marginal productivity condition for employment, and Euler's theorem, the first term in this expression equals zero, so that we have the following expression for the nominal value of equities:

$$(8) \qquad V(t) = \left[\frac{(Y_K-(r+\delta-\Pi))}{r-\Pi} + 1\right] p(t)K(t) \quad .$$

The value of equities varies directly with the gap between the marginal product of capital and the cost of capital. It is interesting to compute the ratio of the nominal value of equities to the nominal value of the capital stock evaluated at the price of newly produced capital p:

$$\frac{V(t)}{p(t)K(t)} = \frac{(Y_K-(r+\delta-\Pi))}{r-\Pi} + 1$$

$$\equiv q$$

which is the argument that appears in the aggregate investment schedule, equation (7). Thus, our investment demand schedule is one that relates firms' demand for capital accumulation directly to the ratio of the value of equities to the replacement value of the capital stock, q. This is the way Tobin formulates the Keynesian investment schedule.

Notice that the dividend-price ratio, which equals the earning--price ratio, is

$$\frac{pY-wN-p\delta K}{V} = r-\Pi \quad ,$$

which is the expected real rate of interest on bonds and equities.

If we add the expected rate of appreciation in the nominal value of exising equities, Π, to the earnings-price ratio, we obtain the nominal yield on equities, r. It is interesting also to calculate the time derivative of V(t). Writing V(t) as p(t)K(t)q(t), we have by logarithmic differentiation with respect to time

$$\frac{\dot{V}(t)}{V(t)} = \frac{\dot{p}(t)}{p(t)} + \frac{\dot{K}(t)}{K(t)} + \frac{\dot{q}(t)}{q(t)} \quad .$$

Suppose q is constant over time so that $\dot{q} = 0$. Then the nominal value of equities changes for two reasons. First, existing equities appreciate in nominal value at the rate $\dot{p}/p$, while investment leads to the issuing of new equities at the rate $\dot{K}/K$.

At each moment in time, households allocate their existing wealth between, on the one hand, bonds and equities which they view as perfect substitutes, and on the other hand, money. Households' total real wealth is denoted W and defined by

$$(9) \qquad \frac{V+B+M}{p} = W$$

Households desire a division of their wealth between M and B+V that is described by the pair of asset demand functions:

$$(10) \qquad \frac{M^D}{p} = m(r, Y, W)$$

$$(11) \qquad \frac{B^D+V^D}{p} = b(r, Y, W)$$

where D superscripts denote desired quantities. The demand schedules are constructed in such a way that at each value of r, Y, and W, total wealth is allocated between M and B+V, so that

(12) $\qquad \dfrac{B^D+V^D+M^D}{p} = W$

for all r, Y, and W. This implies that the partial derivatives of (10) and (11) are related in certain definite ways. Thus, take the total differential of equation (10) and add it to the total differential of equation (11).

$$d\left(\dfrac{M^D+B^D+V^D}{p}\right) = (m_r+b_r)dr + (m_Y+b_Y)dY + (m_W+b_W)dW \quad .$$

Subtract the total differential of equation (12) from the above expression to obtain

$$0 = (m_r+b_r)dr + (m_Y+b_Y)dY + (m_W+b_W-1)dW \quad .$$

This equality can hold for all values of dr, dY and dW if and only if

$$m_r+b_r = 0$$

$$m_Y+b_Y = 0$$

and

$$m_W+b_W = 1 \quad .$$

We shall assume that these restrictions characterize the asset demand functions (10) and (11).

Portfolio equilibrium requires that households be satisfied with the division of their portfolios between bonds and equities, on the one hand, and money on the other:

$$\dfrac{M^D}{p} = \dfrac{M}{p}$$

and

$$\frac{B^D + V^D}{p} = \frac{B+V}{p} \quad .$$

But notice that (9) and (12) together imply that either one of the above equations is sufficient to describe portfolio equilibrium. For suppose that

$$\frac{M}{p} = \frac{M^D}{p}$$

Then subtracting this equality from (12) gives

$$\frac{B^D + V^D}{p} = W - \frac{M}{p} = \frac{B+V}{p} \quad ,$$

so that the demand for the stock of bonds and equities equals the supply. This is an example of Walras' Law: if demand functions build in balance sheet constraints, and if individuals are content with their holdings of all assets but one, then they must be satisfied with their holdings of that last asset too.

We choose to characterize portfolio equilibrium by equality between the supply and demand for money balances:

$$\frac{M}{p} = \frac{M^D}{p} = m(r, Y, w) \quad .$$

We assume that $m_r < 0$, $m_Y > 0$, and $m_w = 0$. It follows that $b_r > 0$, $b_Y < 0$, and $b_W = 1$, so that at given r and Y households desire to hold any increments in real wealth entirely in the form of bonds and equities. We can write our condition for portfolio equilibrium as

$$(13) \qquad \frac{M}{p} = m(r, Y) \quad , \qquad\qquad m_r < 0, \ m_Y > 0 \quad .$$

Real output Y enters (13) as a proxy for the rate of transacting in the economy. We posit that the higher is the rate of transacting, the higher

is the demand for real money balances. The nominal interest rate r equals the difference between the real yield on bonds and equities, $r-\Pi$, and the real yield on money, $-\Pi$. We posit that the larger is the difference between those real yields, the greater is the incentive to economize on money balances in order to hold one of the higher yielding assets.

We can summarize the portfolios of the three sectors of the economy with the following three balance sheets:

Government		Firms		Households	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	B	qpK	V	V	
	M			B	
	net worth			M	net worth

Firms hold neither bonds nor money, while the government owns neither capital nor equities. Households own only paper assets.

The Government

The government collects taxes net of transfers at the real rate T per unit time and makes expenditures at the real rate G per unit time. It will be assumed that government purchases share with consumption the characteristic that they lead to no accumulation of stocks. Goods purchased by the government are used up immediately and do not augment the capital stock. Real taxes net of transfers T are assumed to be collected in a way that makes T independent of real income and the price level. The government sets T and G subject to the flow budget constraint

$$(14) \qquad G = T + \frac{\dot{B}}{p} + \frac{\dot{M}}{p} \quad ,$$

where G, T, $\dot{M}/p$, and $\dot{B}/p$ are all measured in goods per unit time.

The government also conducts open market operations, which are once-and-for-all exchanges of bonds for money at a moment. These exchanges are made subject to the constraint

$$(15) \qquad dM = -dB \quad .$$

Households

Households make two distinct sets of decisions. First, given their stock of wealth at any point in time, they decide how they wish to allocate it among alternative assets. This decision is described by the portfolio equilibrium condition (13). Second, they make a distinct decision about how fast they wish their wealth to grow, that is, they choose a saving rate. This decision determines how they divide their disposable income between consumption and saving.

Households' perceived disposable income represents the rate of income they receive which they expect to be able either to consume or save. Consumption leads to no accumulation of stocks, while saving causes households' wealth in the form of paper assets to grow at some rate per unit time. Households' demand for consumption is summarized by a consumption function that relates their intended real consumption C directly to their perceived real disposable income Y_D and inversely to the real interest rate on bonds and equities:

(16) $\qquad C = C(Y_D, r-\Pi) \ , \ 1 > C_1 > 0, \ C_2 < 0$

where consumption plus saving equal perceived real disposable income:

$$S + C = Y_D \quad .$$

The expressions C_1 and C_2 in (16) denote the partial derivative of C with respect to the first and second arguments, respectively. It is assumed that C_1, the marginal property to consume out of perceived real disposable income is positive but less than unity. Desired consumption is posited to increase in response to an increase in the real rate of interest.

Perceived real disposable income, Y_D, is equal to the real value of wage payments <u>plus</u> dividend payments (which equal economy-wide firms' net cash flow, since firms retain no earnings) <u>minus</u> total real tax collections net of government transfer payments <u>minus</u> the perceived rate of capital loss on the real value of the public's net claims on the government <u>plus</u> the rate at which the real value of equities is increasing <u>minus</u> the real rate at which firms are issuing equities to finance investment. The real rate of wage payments is $\frac{w}{p} \cdot N$. Firms' real dividends are paid at the rate $Y - \frac{w}{p} N - \delta K$. The anticipated rate of capital loss on the real value of government debt is $\frac{M+B}{p} \Pi$. Firms issue equities at the real rate $\dot{K}$ to finance their investment, while the real value of equities actually increases at the rate $q\dot{K} + K\dot{q}$. We will assume that $\dot{q}$ is expected to be zero, so that the real value of equities is expected to increase at the rate $q\dot{K}$. Then perceived real disposable income is

$$(17) \qquad Y_D = Y - \delta K - T - \frac{M+B}{p} \Pi + (q-1)\dot{K} \quad .$$

This concept of disposable income turns out to equal the rate at which society expects that it could consume while leaving its real wealth defined by (9) intact. To show this, we differentiate (9) with respect to time to obtain

$$\dot{W} = \frac{\dot{M}+\dot{B}}{p} - \frac{M+B}{p} \frac{\dot{p}}{p} + q\dot{K} + K\dot{q} \quad .$$

Next replace $\dot{p}/p$ with the public's expectation of it, and replace $\dot{q}$ with the value the public expects it to be, which we have assumed is zero. We obtain

$$\dot{W}_e = \frac{\dot{M}+\dot{B}}{p} - \frac{M+B}{p} \Pi + q\dot{K} \quad .$$

where $\dot{W}_e$ is the rate of change of wealth expected by the public. By virtue of the government's flow budget constraint $(\dot{M}+\dot{B})/p = G-T$, while by virtue of the national income identity, $G = Y-\dot{K}-\delta K-C$. Using these equalities, $\dot{W}_e$ can be written

$$\dot{W}_e = Y-\delta K-T - \frac{M+B}{p} \Pi +(q-1)\dot{K}-C \quad .$$

Using (17), the above equation can be written as

$$C+\dot{W}_e = Y_D \quad ,$$

which verifies that the concept of disposable income Y_D corresponds to the rate at which households can consume while expecting that their real wealth is being left intact (i.e., that $\dot{W}_e = 0$).

Labor Supply

We have now set down enough equations to determine the six variables Y, N, C, I, p and r in our Keynesian model, a model that views the money wage rate w as exogenous at a point in time. In our classical model, on the other hand, w is a variable that must be determined by the model at each moment so that we stand in need of one more equation. The classical model includes a labor supply curve that describes the labor-leisure preferences of workers:

$$(18) \qquad N^S = N(\frac{w}{p}) \qquad\qquad N' > 0$$

where N^S is the volume of employment offered by workers at instant t. It is postulated that the supply of labor is an increasing function of the real wage. The description of the workings of the classical labor market is completed by imposing the condition that actual employment N must equal the volume of employment forthcoming at the existing real wage, N^S. Substituting actual employment N for N^S in (18) then yields

$$(19) \qquad N = N(\frac{w}{p}) \quad .$$

The Complete Model

The classical model can now be summarized as consisting of equations (0), (4), (5'), (6'), (13), (16), and (19), seven equations potentially able to determine seven variables at any moment. For convenience we write down these equations here, renumbering them:

I. $\dfrac{w}{p} = F_N$

II. $N = N(\dfrac{w}{p})$

III. $Y = F(K, N)$

IV. $C = C(Y - T - \delta K - \dfrac{M+B}{p} \Pi + (q(N, K, r-\Pi, \delta) - 1)I, r-\Pi)$

V. $I = I(q(K, N, r-\Pi, \delta) - 1)$

VI. $Y = C + I + G + \delta K$

VII. $\dfrac{M}{p} = m(r, Y)$

We have replaced the variable in equation IV and V with the function $q(K, N, r-\Pi, \delta)$ which equals q by equation (7). The model then consists of seven equations which we will view as determining the seven variables N, $\dfrac{w}{p}$, Y, C, I, r, and p. All of these variables are permitted to jump discontinuously as function of time in order to satisfy the seven equations at each moment. The parameters of the model consist of the exogenous variables T, G, δ, Π, M and the additional parameters determining the shapes of the underlying functions. Notice that the anticipated rate of inflation, which is assumed to be unanimously held, is one of the parameters, and that in particular it is not a variable that depends on the actual rate of inflation. Were Π to depend on $\dot{p}/p$, defined as

the right hand time derivative of the logarithm of p, then the above

seven equations would involve eight variables--$\dot{p}$ would be the eighth--

and would not form a complete model. For similar reasons we have assumed

that the public's expectation of $\dot{q}$ is exogenous, and more particularly

equal to zero. If the expected $\dot{q}$ were a function of the actual $\dot{q}$ we

would have to determine the actual right hand time derivative of q in

order to determine the values of our other variables at any moment. More

of this later. For now, we simply assume that Π and the expected $\dot{q}$ are

both exogenous, and note that on this assumption the model possess the

same number of equations as variables it must determine at a point in

time.

We begin our analysis of the classical model by in effect linear-

izing the system of equations I through VII around initial equilibrium

values of the variables. We assume that such an initial equilibrium exists.

(Later we will consider some problems posed by Keynesians that may call

into question the existence of an equilibrium in a classical model like

ours.) Then to obtain a linear system we proceed to obtain the total

differentials of equation I through VII, the differentials being under-

stood as representing deviations from initial equilibrium values of the

variables:

i. $\quad d\left(\frac{W}{p}\right) = F_{NN}dN + F_{NK}dK$

ii. $\quad dN = N'd\left(\frac{W}{p}\right)$

iii. $\quad dY = F_{N}dN + F_{K}dK$

iv. $\quad dC = C_1 dY - C_1 dT - C_1 \delta dK - C_1 \frac{M+B}{p} d\Pi - C_1 \Pi\left(\frac{dM+dB}{p} - \frac{M+B}{p}\frac{dp}{p}\right) + C_1\left[(q-1)dI + Iq_N dN + \right.$

$\left. Iq_K dK + Iq_{r-\Pi}dr - Iq_{r-\Pi}d\Pi\right] + C_2 dr - C_2 d\Pi$

v $\quad dI = I'q_N dN + I'q_K dK + I'q_{r-\Pi} dr - I'q_{r-\Pi} d\Pi$

vi $\quad dY = dC + dI = dG + \delta dK$

vii $\quad \dfrac{dM}{p} - \dfrac{M}{p}\dfrac{dp}{p} = m_r dr + m_Y dY$

We have assumed that δ is always constant, so that $d\delta = 0$. The derivatives of q with respect to N, K, and r-Π, which appear in (iv) and (v) are obtained by differentiating equation (7):

$$dq = \frac{[F_{KN}dN + F_{KK}dK - (dr - d\Pi)](r-\Pi) - [F_K - (r+\delta-\Pi)](dr-d\Pi)}{(r-\Pi)^2}$$

$$= \frac{1}{r-\Pi}(F_{KN}dN + F_{KK}dK - q(dr-d\Pi)).$$

Thus, we have

$$q_N = \frac{1}{r-\Pi}F_{KN} > 0 \ ,$$

$$q_K = \frac{1}{r-\Pi}F_{KK} < 0 \ ,$$

$$q_{r-\Pi} = \frac{-q}{r-\Pi} < 0 \ ,$$

so that q is an increasing function of employment and a decreasing function of capital and the real rate of interest.

Notice that equation (iv) can be simplified somewhat due to the constraint on open market operations, dM+dB = 0. It is then of some interest to write equations (i) through (vii) in the following matrix form:

$$
\begin{bmatrix}
1 & -F_{NN} & 0 & 0 & 0 & 0 & 0 \\[2mm]
-N' & 1 & 0 & 0 & 0 & 0 & 0 \\[2mm]
0 & -F_N & 1 & 0 & 0 & 0 & 0 \\[2mm]
0 & -C_1 Iq_N & -C_1 & 1 & -C_1(q-1) & -(C_1 Iq_{r-\Pi}+C_2) & -C_1 \Pi\frac{M+B}{p^2} \\[2mm]
0 & -I'q_N & 0 & 0 & 1 & -I'q_{r-\Pi} & 0 \\[2mm]
0 & 0 & 1 & -1 & -1 & 0 & 0 \\[2mm]
0 & 0 & n & 0 & 0 & m_r & \frac{M}{p^2}
\end{bmatrix}
\begin{bmatrix}
d\left(\frac{W}{p}\right) \\[2mm]
dN \\[2mm]
dY \\[2mm]
dC \\[2mm]
dI \\[2mm]
dr \\[2mm]
dp
\end{bmatrix}
=
\begin{bmatrix}
F_{NK}dK \\[2mm]
0 \\[2mm]
F_K dK \\[2mm]
-C_1 dT-C_1\delta dK-C_1\frac{M+B}{p}d\Pi +Iq_K dK \\[1mm]
-(C_1 Iq_{r-\Pi}+C_2)d\Pi \\[2mm]
I'q_K dK-I'q_{r-\Pi}d\Pi \\[2mm]
dG+ dK \\[2mm]
\frac{dM}{p}
\end{bmatrix}
$$

Inspection of the 7x7 matrix on the left side of the above equation reveals a peculiar characteristic of the classical model that is very important. In particular, note that only two variables appear in the first two equations: $d\left(\frac{w}{p}\right)$ and dN. All of the other variables have zero coefficients in the first two equations. As a consequence, these two equations form an independent subset that determines $d\left(\frac{w}{p}\right)$ and dN, no contribution being made by the remaining variables to the determination of these two. Similarly, the first three equations also form an independent subset capable of determining dY as well as dN and $d\frac{w}{p}$ independently of the other four equations in the model. The above system is an example of a "block recursive" system of equations. In such a system interdependence is not general, the system being solvable sequentially, since at least one subset of equations involves a closed subset of the variables. However, once the values of the variable in the subset are determined, they may influence, though not be influenced by, the value taken by the remaining variables.* The fact that the classical model has this property is very important.

----

*A system of equations is said to be block recursive if it can be written in the form

$$A \ x = b$$

where

$$A = \begin{bmatrix} A^{11} & A^{12} & \ldots & A^{1N} \\ A^{21} & A^{22} & \ldots & A^{2N} \\ \vdots & & & \\ A^{N1} & A^{N2} & \ldots & A^{NN} \end{bmatrix}, \quad x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^N \end{bmatrix}, \quad b = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^N \end{bmatrix}$$

where A^{IJ} are matrices; and where $A^{IJ} = 0$ for all $I = 1, \ldots, N$ and $J > I$. In the case of the classical model, A^{12} would consist of the 3x4 matrix in the upper right hand corner of the matrix in the text. Then the first three variables can be determined from

$$x' = (A^{11})^{-1} b^1 .$$

We solve the subsystem formed by equations (i) and (ii) by substituting (ii) into (i), which yields:

$$d\left(\frac{W}{p}\right) = F_{NN} \cdot N' \ d\left(\frac{W}{p}\right) + F_{NK} dK$$

or

$$(20) \qquad d\left(\frac{W}{p}\right) = \frac{F_{NK}}{(1-F_{NN}N')} \ dK \quad .$$

Since we have assumed that $F_{NK} > 0$, $F_{NN} < 0$, $N' > 0$, it follows that $F_{NK}/(1-F_{NN}N') > 0$. Equation (20) states that a change in the capital stock at a point in time, could it occur, would drive the real wage upward.

Substituting (20) into (ii) yields

$$(21) \qquad dN = N' \frac{F_{NK}}{(1-F_{NN}N')} \ dK.$$

Since $N' > 0$, (21) implies that an increase in the capital stock at a point in time would increase employment. It would do so by causing an increase in the demand for labor, in turn causing an increase in the real wage, which would increase the quantity of labor supplied.

Substituting (21) into (i), the total differential of the production function, yields

$$(22) \qquad dY = \left(\frac{F_N N' F_{NK}}{(1-F_{NN}N')} + F_K\right) dK \quad .$$

Since F_K and F_N are both positive, it follows that a once-and-for-all increase in capital would produce an increase in the rate of output, both because the marginal product of capital is positive, and because the increase in capital would increase both the marginal product of capital and the number of workers employed.

Equations (20), (21), and (22) show that given the production function and the labor supply schedule, only a once-and-for-all change in the capital stock can bring about a once-and-for-all change in the rate of output: K is the only variable that is exogenous at a point in time that enters into the determination of the levels of output, employment, and the real wage at a point in time. Other exogenous variables may affect the time rates of change of those variables, but not their levels at a point in time. We shall assume from now on that capital can only be accumulated by investing, thus ruling out once-and-for-all jumps in the capital stock. From this it follows that given the production function and the labor supply schedule, output, employment, and the real wage are constants at a point in time, independent of what government policy and the public's expectations are.

The situation is depicted in Figure 1, which shows the production function, its derivative with respect to employment, which is the demand schedule for employment, and the supply schedule for labor. Employment and the real wage are determined at the intersection of the demand and supply schedules for labor, while output is determined by substituting the equilibrium level of employment into the production function.
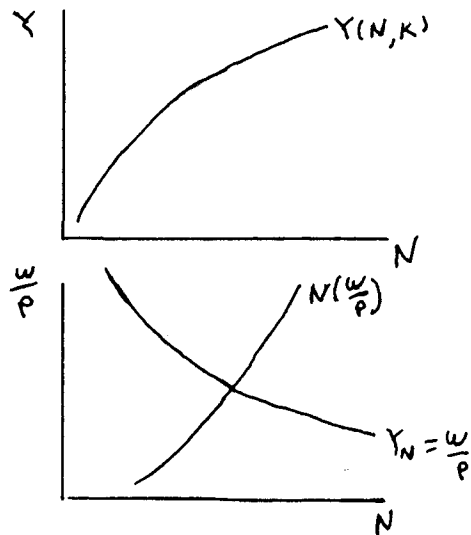

Figure 1

Equations (i), (ii), and (iii) may be thought of as the determinants of aggregate supply. The remaining four equations have the role of ensuring that aggregate demand is made equal to aggregate supply, aggregate demand doing all of the adjusting. In most discussions of the classical model, it is only equations (iv), (v), and (vi) that are involved in equilibrating aggregate demand and supply, the money market equilibrium equation (vii) playing no role. This requires that dp not appear in equations (iv), (v), and (vi), since only then will those three equations form an independent subset in dC, dI, and dr, given the values of dY, $d(\frac{W}{p})$ and dN that emerge from the aggregate supply subset.* This condition will be met only if it happens that the coefficient on dp in equation (iv), which is $- C_1 \Pi \frac{M+B}{p^2}$, happens to equal zero. This will be so either if Π happens to be zero or if M+B is zero, B being negative and representing government loans to the public.

Since it is commonly assumed that real consumption is independent of the price level, we begin our analysis by tentively assuming that M+B = 0 initially. The substituting (iv) and (v) into (vi), and setting dK = dY = dN = 0,* we have

$$- C_1 dT + [C_1 I q_{r-\Pi} + C_2 + (1 + C_1(q-1))I'q_{r-\Pi}]dr$$

$$- [C_1 I q_{r-\Pi} + C_2 + (1 + C_1(q-1))I'q_{r-\Pi}]d\Pi + dG = 0$$

Solving for dr, we have

_____

*Alternatively, notice that if the coefficients on dp are all zero, the formal requirements for block recursiveness with respect to dC, dI, and dr are fulfilled.

*We know from solving i, ii, and iii that dY = dN = 0 if dK = 0.

(23) $dr = \dfrac{C_1}{H} dT - \dfrac{1}{H} dG + d\Pi$

where $H = [C_1 Iq_{r-\Pi} + C_2 + (1 + C_1(q-1)I'q_{r-\Pi}] < 0$

We assume that the partial derivative of disposable income with respect to the interest rate obeys

$$\dfrac{-C_2}{C_1} > Iq_{r-\Pi} + (q-1)I'q_{r-\Pi} = \dfrac{\partial Y_D}{\partial r} \bigg|_{dT=dY=d\Pi=dp=0}$$

which requires that, if it is positive, it not be too large in absolute value. This condition is sufficient, although not necessary, to guarantee that H is negative. As we shall see below, H must be negative if the model is to be "stable." The magnitude H has a straight forward interpretation. It is the total derivative of aggregate demand, $C+I+G+\delta K$, with respect to the interest rate. Manipulation of (23) then produces the following expression for the partial derivatives of r with respect to the exogenous variables T, G, and Π;

$$\dfrac{\partial r}{\partial T} = \dfrac{C_1}{H} < 0 ,$$

$$\dfrac{\partial r}{\partial G} = \dfrac{-1}{H} > 0 ,$$

$$\dfrac{\partial r}{\partial \Pi} = 1.$$

These are partial derivatives of the "reduced form" for r, which express r solely in terms of the exogenous variables, the endogenous variables all having been eliminated by substitution. The interest rate rises in response to an increase in government purchases or a decrease in the rate of tax collections. If Π rises, r rises by the same amount, leaving $r-\Pi$ unaltered.

To determine the effect of changes in T, G, and Π on net investment I, we substitute (23) into (v) and solve for the partial derivatives of I with respect to those exogenous variables:

$$\frac{\partial I}{\partial T} = I'q_{r-\Pi}\frac{\partial r}{\partial T} > 0$$

$$\frac{\partial I}{\partial G} = I'q_{r-\Pi}\frac{\partial r}{\partial G} < 0$$

$$\frac{\partial I}{\partial \Pi} = I'q_{r-\Pi}\frac{\partial r}{\partial \Pi} - I'q_{r-\Pi} = 0.$$

We see that net investment I is stimulated by an increase in taxes or a decrease in government expenditures. A change in Π leaves the real rate of interest $r-\Pi$ unaltered, and so has no effect on investment.

The effects of changes in T, G, and Π on consumption can be studied by substituting (23) into (iv) and computing the following partial derivatives:

$$\frac{\partial C}{\partial T} = C_1+(C_2+C_1Iq_{r-\Pi}+C_1(q-1)I'q_{r-\Pi})\frac{\partial r}{\partial T}$$

$$\frac{\partial C}{\partial G} = (C_2+C_1Iq_{r-\Pi}+C_1(q-1)I'q_{r-\Pi})\frac{\partial r}{\partial G}$$

$$\frac{\partial C}{\partial \Pi} = (C_2+C_1Iq_{r-\Pi}+C_1(q-1)I'q_{r-\Pi})\frac{\partial r}{\partial \Pi} - (C_2+I'q_{r-\Pi}+(q-1)I'q_{r-\Pi}) = 0.$$

If q is not too much smaller than 1, then $(C_2+C_1Iq_{r-\Pi}+C_1(q-1)I'q_{r-\Pi})$ is negative, and $\partial C/\partial G$ is negative while $\partial C/\partial T$ is also negative. The expression $Iq_{r-\Pi}+(q-1)I'q_{r-\Pi}$ is the partial derivative of disposable income with respect to the interest rate. We have assumed that this derivative satisfies

$$-\frac{C_2}{C_1} > Iq_{r-\Pi}+(q-1)I'q_{r-\Pi} \quad ,$$

so that the derivative of disposable income with respect to the interest rate is less than $-C_2/C_1$, which guarantees that $\partial C/\partial G < 0$. It also follows that $\partial C/\partial T < 0$.

In this version of the classical model—in which M+B is equal to zero initially—the interest rate bears the entire burden of adjusting the level of aggregate demand so that it equals the aggregate supply of output determined by equations I, II, and III. We have already seen that given the capital stock, aggregate supply is independent of the other exogenous variables. If from a position of initial equilibrium, one of the exogenous variables changes in such a manner that it induces an increase in aggregate demand at the initial interest rate, the interest rate must rise to diminish desired consumption and investment by enough to restore equality between aggregate demand and the unchanged rate of aggregate supply determined by I, II, and III. That is why the interest rate must rise, for example, in response to increases in aggregate demand induced by increases in government expenditures or reductions in taxes. The rise in the interest rate in turn generally induces changes in rates of capital accumulation and consumption. Only in the case of a change in Π does the net effect of the change fail to affect investment and consumption. That is because a change in Π produces an equivalent changes in r which leaves $r-\Pi$, the real rate of interest, unaltered.

The determination of the interest rate in the classical model is easily illustrated graphically. First note that since disposable income is allocated between savings and consumption, we have

$$S+C = Y-T-\delta K+(q-1)I.$$

Substituting the rational income identity (0) into the above equation

and rearranging yields an alternative form of the equilibrium condition

between aggregate demand and aggregate supply,

(24)   G+qI = S+T.

Equation (24) states that in equilibrium, real government expenditures

plus real investment evaluated at the stock market value of equities must

equal saving plus the rate of tax collections. A version of (24) is

sometimes interpreted as stating that "injections" (G+qI) must equal

"leakages" (S+T)if the economy is to be in equilibrium.

The role of the interest rate in making aggregate demand equal

the predetermined aggregate supply is shown graphically in figure 2, which
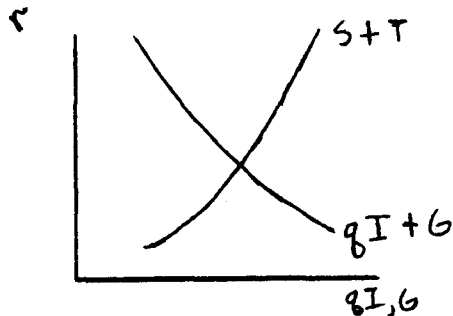
depicts S+T as an increasing function
of interest and qI+G as an inverse func-
tion of the interest rate. We know
that qI is an inverse function of the
interest rate, since

$$\frac{\partial(q(K, N, r-\Pi)I(q(K, N, r-\Pi))}{\partial r} = q_{r-\Pi}I+qI'q_{r-\Pi} < 0 \ ,$$

as long a I > -qI', which we assume is true.

We have depicted savings plus exogenous tax collections as

varying directly with the interest rate, which need not be true. For

saving is defined by

$$S = Y_D - C(Y_D, \ r-\Pi) \ .$$

The derivative of saving with respect to the interest rate is

$$\frac{\partial S}{\partial r} = \frac{\partial Y_D}{\partial r} - C_1 \frac{\partial Y_D}{\partial r} - C_2$$

$$= \frac{\partial Y_D}{\partial r} (1-C_1) - C_2 \ .$$

We have assumed that $\partial Y_D / \partial r$ obeys

$$\frac{-C_2}{C_1} > \frac{\partial Y_D}{\partial r} \ ,$$

which implies only that

$$\frac{\partial S}{\partial r} < - \frac{C_2}{C_1} (1-C_1) - C_2$$

or

$$\frac{\partial S}{\partial r} < - \frac{C_2}{C_1} \ .$$

So our assumptions do not rule out a saving schedule that depends inversely

on the rate of interest. Both S+T and G+qI also depend on output, since

I varies with employment and S with output. But output is predetermined

from the point of view of interest rate determination, and it will not vary

in response to variations in G or T. The equilibrium interest rate is

determined at the intersection of the qI+G and the S+T curves. Using the

graph, it is easy to verify that increases in government expenditures drive

the interest rate upward, while increases in tax collections drive it

downward. An increase in Π is easily verified* to shift both the S+T curve

_____

*For example, given G+qI, how much must r change in order
to stay on the G+qI curve when Π changes? Taking the differential
of the G+qI schedule and setting $d(G+qI) = dN = dK = 0$, we have

$$(Iq_{r-\Pi} + dI'q_{r-\Pi})dr = (Iq_{r-\Pi} + qI'q_{r-\Pi})d$$

or

$$dr = d\Pi \ ,$$

which establishes that the qI + G curve shifts upward by the amount
of the increase in Π.

and the G+qI curve upward by the amount of the increase in Π. The result
is that r rises by the increase in Π, the equilibrium G+qI being left
unchanged.

An equivalent "loanable-funds" interpretation can readily be
placed on the process of interest rate determination. Substituting the
government's flow budget constraint into equation (24) yields the alter-
native form of that equation,

$$qI + \frac{\dot{M}}{p} + \frac{\dot{B}}{p} = S.$$

The left-hand side of this equation is the expected time rate of growth
of the economy's real stock paper assets. The right-hand side is the
rate at which the public desires to add to its stocks of assets, that is,
desired saving. In equilibrium, the expected rate of growth of the
economy's paper assets must just equal the rate at which the public wishes
to add to its assets. The expected real growth rate of government issued
financial assets, $(\dot{M} + \dot{B})/p$, is equal to the government's deficit. Given
that rate and given total taxes, T, the interest rate adjusts to ensure
that desired saving exceeds the expected rate of increase in the real value
of equities, qI, exactly by the real rate at which the government is
expanding the public's claims upon it in the form of financial assets.
Thus, for example, given real tax receipts, T, an increase in government
expenditures raises $\frac{\dot{M} + \dot{B}}{p}$; in order for equilibrium to be restored, the
interest rate must rise, thus diminishing investment and stimulating
saving, to such a point that the new higher real rate of addition to
government issued financial assets is consistent with the public's saving
and investment plans. The above equation succinctly summarizes the tra-

ditional rationale for levying taxes as opposed to financing government expenditures by printing bonds or money: It is to protect the rate of growth of physical capital by limiting the extent to which private saving is diverted to accumulating claims on the government instead of (claims on) physical capital.

The differentials of the interest rate and output having been determined in equations (i) through (vi), the role of equation (vii) is simply to determine the differential of the price level:

$$(25) \qquad \frac{dp}{p} = \frac{dM}{M} - {}^m r\frac{p}{M} \, dr - M \frac{p}{M} dY$$

Thus, if the money supply is the only exogenous variable that changes, only the price level is affected, and it changes proportionately with the money supply. On the other hand, the price level does respond to the changes in output and the interest rate that emerge from equations (i) through (vi), increases in the interest rate driving the price level upward, while increases in output drive it downward. However, as long as neither dp nor dM appears in equations (i) through (vi), "money is a veil," having no effects on output employment, the real wage, consumption, investment, or the interest rate. Recall that in order to eliminate dp from the equation for dC, we assumed that M+B was zero initially.

Stability

It is important to verify that the equilibrium positions we have described are stable, i.e., that there is a tendency to return to them if the system is displaced from them. Otherwise the comparative static exercises we have performed are of little practical interest. The need to examine the stability of static equilibria is the heart of Samuelson's "correspondence principle."

In the classical model it is usual to argue that the commodity price level adjusts when there is a discrepancy between aggregate demand and supply of goods, while the interest rate adjusts when there is inequality between the supply of real balance and the demand. We summarize this by the following two differential equations, where s is the time index:

$$\frac{dp}{ds} = \sigma(C(Y_D, r-\Pi)+I(q-1)+G+\delta K-Y(K, N)), \sigma' > 0, \ \sigma(0) = 0 \ .$$

$$\frac{dr}{ds} = \beta(m(r, Y)-\frac{M}{p}), \ \beta' > 0, \ \beta(0) = 0 \ .$$

The first equation states that the price level rises when flow aggregate demand exceeds flow aggregate supply, while the second states that the interest rate rises when the demand for real balances exceeds the supply, or, what is the same, when the real supply of bonds and equities exceeds the demand for them. Taking the first order part of the Taylor's expansion of these equations around equilibrium values of the variables gives the approximation:

$$\frac{dp}{ds} = \sigma'\{C_2+I'q_{r-\Pi}+C_1[(q-1)I'q_{r-\Pi}+Iq_{r-\Pi}]\} \ (r-r_0)$$

$$\frac{dr}{ds} = \beta'm_r(r-r_0)+\beta'\frac{M}{p^2}(p-p_0)$$

where zero subscripts denote initial equilibrium values and where we continue to assume that M+B equals zero initially, so that changes in the price

level doesn't affect Y_D. This system can be written in matrix form as

$$
\begin{bmatrix} \dfrac{dp}{ds} \\[2ex] \dfrac{dr}{ds} \end{bmatrix}
=
\begin{bmatrix} 0 & \sigma'H \\[2ex] \beta'\dfrac{M}{p^2} & \beta'M_r \end{bmatrix}
\begin{bmatrix} p-p_0 \\[2ex] r-r_o \end{bmatrix}
$$

where as in (23), H is defined by

$$
H = C_2 + I'q_{r-\Pi} + C_1 q_{r-\Pi}((q-1)I'+I) \ .
$$

This system is an example of a linear differential equation of first order of the form

$$
\frac{dx}{ds} = Ax
$$

where x is an (nx1) vector and A is an (nxn) matrix. The solution of the system has the general form

$$
x_j(s) = k_{j0} + \sum_{h=1}^{n} k_{jh} e^{\lambda hs}
$$

where $x_j(s)$ is the value of the jth component of x at instant s, the k_j's are constants, and the λ_h, h=1, ..., n are the roots of the characteristic equation

$$
\left| A - \lambda I \right| = 0
$$

where the vertical bars denote the determinant. The system will be stable if the eigenvalues, the λ's, have negative real parts. For our system the characteristic equation is

$$
\begin{vmatrix} -\lambda & \sigma'H \\[3ex] \beta'\dfrac{M}{p^2} & \beta'm_r-\lambda \end{vmatrix} = 0
$$

$$\lambda^2 - \beta' m_r \lambda - \sigma' \beta' \frac{M}{p}2 \, H = 0$$

Necessary and sufficient conditions that the eigenvalues have negative real parts are that the coefficients on λ and the constant term both be positive.* Thus, necessary and sufficient conditions for stability are

$$\beta' m_r < 0$$

$$\sigma' \beta' \frac{M}{p}2 \, H < 0 \; .$$

Since σ', β', and $\frac{M}{p}2$ all exceed zero, our stability conditions are

$$m_r < 0$$

$$H = C_2 + I' q_{r-\Pi} + C_1 q_{r-\Pi} ((q-1)I'+I) < 0.$$

The interest elasticity of the demand for money must be negative while the interest slopes of consumption and investment must sum to a negative number.

Notice that if $m_r = 0$ the model is not stable. If m_r equals zero the characteristic equation becomes

$$\lambda^2 - \sigma' \beta' \frac{M}{p}2 \, H = 0$$

whose solution is

$$\lambda = \pm \sqrt{\sigma' \beta' H \frac{M}{p}2}$$

Since H is negative, the two roots are complementary imaginary numbers, i.e., the real parts of the roots are zero. Consequently where $\lambda_1 = bi$, $\lambda_2 = -bi$, the solution for r(s), for example, will have the form

---

*See Paul Samuelson, <u>Foundations of Economic Analysis</u>, pp. 430-431.

$$r(s) = k_{10} + k_{11}e^{bis} + k_{12}e^{-bis}$$

$$= k_{10} + k_{11}(\cos b s + i \sin bs) + k_{12}(\cos b s - i \sin bs)$$

$$= k_{10} + (k_{11} + k_{12})\cos bs + (k_{11} - k_{12}) i \sin bs.$$

The solution consists of a nondamped, nonexplosive sinusoidal. There is no tendency for the system to approach equilibrium as "time" passes, regardless of the adjustment speeds.* Remember, however, that it is a linear approximation to the actual dynamic system which possesses this borderline instability with nondamped but nonexplosive oscillations. In this case, the behavior of the system itself, as opposed to the linear approximation, may be either stable or unstable. The character of the nonlinearities of the system is what will determine whether the system is stable. The appropriate way to analyze the matter is to consider higher order terms in the Taylor's series approximation to the dynamic system.

The possible instability of the model where $m_r = 0$ means that comparative static excercises that are performed assuming $m_r = 0$ must be interpreted carefully. On the correspondence principle, such excercises only make sense if we interpret them as describing how the system will behave as m_r approaches zero from below.

---

*The role of real and imaginary parts of the roots in governing convergence to equilibrium can be seen by rewriting the solution of $x_j(s)$ as

$$x_j(s) = k_{jo} + \sum_{h=1}^{n} k_{jh}e^{(a_h + b_h i)s}$$

where $\lambda_h = a_h + b_h i$. Then

$$x_j(s) = k_{jo} + \sum_{h=1}^{n} k_{jh}e^{a_h s} (\cos b_h s + i \sin b_h s).$$

If all $a_h < 0$, the oscillations of the system (if any) are damped.

If all $a_h = 0$, there is no damping.

The differential equations set out above should not be construed as describing the evolution of the system in calendar time. For as we have already pointed out, the variables are assumed to jump instantaneously to satisfy the equilibrium conditions. Consequently, we will interpret the adjustment processes described above as processes in which σ' and β' both approach infinity. On this interpretation, the above calculations retain their sense, while we can keep the view that adjustments of the variables of the model occur instantaneously, provided that the model is stable.

The Model Where M+B $\neq$ 0

Up to now, we have assumed that M+B initially equals zero, implying that expected real disposable income and consumption are both independent of the price level. We now investigate the implications of abandoning that special assumption, and assume M+B > 0. Inspection of equations (i) through (vii) indicates that, given dY, dN, and $d\frac{w}{p}$ determined by equations (i), (ii), and (iii), equations, iv, v, vi, and vii form an interdependent system that determines dC, dI, dr, and dp. Consequently, we should expect dM, which appears in (vii), now to influence dI, dr, and dC.

To solve the system, we first substitute (iv) and (v) into (vi) to obtain the differential of the r-p locus that makes aggregate demand equal to aggregate supply:

$$(26) \qquad 0 = -C_1 dT + dG + \frac{M+B}{p} C_1 \Pi \frac{dp}{p} + Hdr$$

$$-[H + C_1 \frac{M+B}{p}]d\Pi \quad ,$$

where as in equation (23) H is defined by H = $(C_2 + C_1 Iq_{r-\Pi} + [1 + C_1(q-1) I'q_{r-\Pi}])$ < 0. Notice that the slope in the p-r plane of the locus of points that make aggregate demand equal to aggregate supply depends on the sign of (M+B) $\cdot$ Π.

Equation (25) gives the differential of the locus of points that make the demand for real balances equal to the supply. Substituting (25) into (26) gives an equation that is the differentaial of the reduced form equation for r:

$$(27) \qquad [H - m_r \frac{M+B}{M} C_1 \Pi]dr = C_1 dT - dG - \frac{M+B}{p} C_1 \Pi \frac{dM}{M} + [H + C_1 \frac{M+B}{p}]d\Pi \quad .$$

We assume that

$$[H - m_r \frac{M+B}{M} C_1 \Pi] < 0 ,$$

a condition that guarantees that aggregate demand falls in response to an increase in r. This is a condition of stability. We use (27) to compute the following reduced form partial derivatives:

$$\frac{\partial r}{\partial G} = \frac{-1}{H - m_r \frac{M+B}{M} C_1 \Pi} > 0$$

$$\frac{\partial r}{\partial T} = \frac{C_1}{H - m_r \frac{M+B}{M} C_1 \Pi} < 0$$

$$\frac{\partial r}{\partial \Pi} = \frac{H + C_1 \frac{M+B}{p}}{H - m_r \frac{M+B}{M} C_1 \Pi} .$$

We note that if M+B is set equal to zero, the above expressions agree with our earlier expressions for those three partial derivatives. We note that if M+B exceeds zero and Π is positive, an increase in government expenditures or a decrease in taxes will increase the rate of interest more than if M+B were zero provided that

$$[H - m_r \frac{M+B}{M} C_1 \Pi] < 0 ,$$

as we have assumed. Thus, suppose government expenditures rise. This tends to increase the interest rate which in turn, through equation (vii), causes the price level to rise. But the rise in the price level in turn reduces the real value of government debt, and given an exogneous, positive Π, also reduces the anticipated rate of real capital loss on those assets. Hence expected real disposable income increases, which in turn increases the desired rate of consumption, driving the interest rate up

even farther. A similar explanation holds for the response of the interest rate to an increase in taxes. It is clear that the signs of those effects ought to depend on the signs of Π and $(M+B)/p$, since if one of them is negative, an increase in p will produce a decrease in expected real disposable income.

Next notice that an increase in Π no longer leads to an equivalent increase in the interest rate. This is due to the effect of changes in Π on expected real disposable income. It is possible for the interest rate to rise more or less than the given rise in Π.

We now evaluate the reduced form partial derivative of r with respect to M:

$$\frac{\partial r}{\partial M} = \frac{-\frac{M+B}{p} C_1 \Pi}{[MH - m_r(M+B)C_1\Pi]}$$

If M+B or Π equals zero, we see that an increase in the money supply has no effect on the interest rate, as before. If Π and M+B are both positive, however, then an increase in M drives r upward. The increase in money drives prices upward, which reduces the real value of anticipated capital losses on paper assets, thus raising perceived real disposable income, driving intended consumption upward. The interest rate must rise to equate aggregate demand with aggregate supply. Thus money has ceased to be a veil, being capable of affecting the real interest rate and hence the rate of capital accumulation. That in turn gives the monetary authority some control over the economy's rate of growth over time. But clearly, money remains important as far as concerns determining the level of output.

The Model with $\Pi = \dfrac{\dot{p}}{p}$

Up to now we have always taken as exogenous those expectations that pertain to the future values of some of the endogenous variables. In doing so we have ruled out the existence of "perfect foresight" or for that matter the existence of any systematic relationships between expectations and the actual subsequent behavior of those endogenous variables about which expectations are being formed. We have done so in spite of the fact that, given assumed paths of the exogenous variables over future time, the model is supposed to be capable of determining the value of the endogenous variables at each moment. Here we show that making the assumption that expectations of future endogenous variables are exogenous is necessary in order to perform comparative static exercises, that is, excercises that involve computing alternative equilibrium values of endogenous variables at one and only one point in time.

Suppose that we have a model that can be summarized by the system of m reduced form equations in T exogenous variables:

(28) $z_1(t) = f_1(x_1(t), \ \ldots \ x_T(t))$

$z_2(t) = f_2(x_1(t), \ \ldots, \ x_T(t))$

$\vdots$

$z_m(t) = f_m(x_1(t), \ \ldots, \ x_T(t))$

where $z_j(t)$ is the value of the jth endogenous variable at instant t and $x_i(t)$ is the value of the ith exogenous variable at time t. The exogenous variables do not respond to movements of the endogenous variables at a point in time, but may depend on past values of the endogenous variables. For example, the capital stock is related to past rates of investment by:

(29) $\qquad K(t) = \int_{t_0}^{t} I(s)ds + K(t_0)$;

but the current capital stock is independent of the current rate of invest-ment, as long as the latter is finite. The evolution of the capital stock over time is thus determined by the model, but at each moment the capital stock is predetermined.

The system of equations (28) determines the values of the endo-genous variables at any given moment. Furthermore, if we specify arbitrary time paths for exogenous variables or, where appropriate, laws of motion over time that like equation (29), link exogenous variables to past values of endogenous variables, then the system of equations (28) is capable of determining the endogenous variables at each moment over some time interval. The model thus describes the 'dynamic' behavior of the endogenous variables, despite the fact that we have heretofore restricted ourselves to comparing alternative equilibrium positions at a point in time. Those static exer-cises were meaningful since the model was capable of determining the equilibrium values of the variables at a moment taking into account only the values of the exogenous variables at that moment.

Now suppose we propose appending the equations (28) the relation-ship

(30) $\qquad x_1(t) = z_m(t) \equiv \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{z_m(t+\varepsilon) - z_m(t)}{\varepsilon}$

where $x_1(t)$ is a variable, previously treated as exogenous, that equals the public's expectation of the time rate of change of $z_m(t)$ over the immediate future. For example z_m might be the logarithm of the price level, so that $x_1(t)$ would correspond to the anticipated rate of inflation.

Notice that the derivative in (30) is a right-hand derivative, so that (30) embodies a notion of "perfect foresight."

If equation (30) is added to (28), we have a system that is no longer capable of determining the values of the endogenous variables at a single point in time given the remaining exogenous variables $x_2(t)$, ..., $x_T(t)$ at that point in time. For in order to determine any endogenous variable at a point in time, it is necessary also to determine $\dot{z}_m(t)$, the right-hand derivative of z_m at that same moment in time. But to determine $\dot{z}_m$ requires in effect that we determine z_m over at least a small interval of the future. The equation for $z_m(t)$ is found by substituting (30) into the last equation of (28):

$$(31) \qquad z_m(t) = f_m(\dot{z}_m(t), x_2(t), \ldots, x_T(t)) \quad .$$

At any given moment in time, this equation determines only a locus of $z_m(t)$, $\dot{z}_m(t)$ combinations that satisfy it. As long as both $z_m(t)$ and $\dot{z}_m(t)$ are viewed as being determined by the model at the current instant, we aren't able to arrive at a solution.

It might be thought that a way out of our problem is to differentiate (31) with respect to time in the right-hand direction in order to get an equation that determines $\dot{z}_m(t)$. But this yields

$$(32) \qquad \dot{z}_m(t) = \frac{\partial fm}{\partial x_1} \ddot{z}_m(t) + \frac{\partial fm}{\partial x_2} \dot{x}_2(t) + \ldots + \frac{\partial fm}{\partial x_T} \dot{x}_T(t)$$

where all time derivatives are right-hand ones. Equation (32) only determines a locus of $\dot{z}_m$, $\ddot{z}_m$ combinations that solve the equation, given values of $\dot{x}_2(t)$, ..., $\dot{x}_T(t)$. Following this line of reasoning will then lead us to follow an infinite progression calculating higher and higher order time

derivatives of z_m. In order to determine z_m under assumption (29), it is necessary to determine right-hand derivatives of all orders at that moment.

Equivalently, equation (31) is a differential equation in $z_m(t)$, one that might possibly be solved for a time path of $z_m(t)$ over a time interval $t_0 \leq t \leq t_n$, where t_n could be arbitrarily large. To compute a particular solution, a terminal condition on z_m would have to be supplied. Then z_m might be able to be expressed as a function of (expected) future values of the exogenous variables, e.g.

$$(33) \qquad z_m(t) = g\left(\int_t^{t_n} W_2(s-t)x_2(s)ds, \ldots, \int_t^{t_n} W_T(s-t)x_T(s)ds, z_m(t_n) \right)$$

for

$$t_0 \leq t \leq t_n \quad ,$$

where $W_2(\ )$, $\ldots$, $W_T(\ )$ are coefficients that are functions of the parameters of the equation system (28). Equation (33) expresses the current value of z_m over time in terms of the (expected) paths of the exogenous variables. In order to determine the equilibrium value of z_m at any one moment in time, it is necessary to calculate an equilibrium time path of z_m over some portion of the future.

To do the point-in-time comparative static exercises we have performed, it is necessary to divorce equilibrium values of the endogenous variables at one moment from any dependence on equilibrium values of those variables in subsequent moments. To do that, it is necessary to deny the existence of any relationships that, like equation (30), produce a dependence of current endogenous variables on subsequent values of endogenous variables. This explains why we have maintained that the public's expectations do not embody perfect foresight, or do not depend systematically on

actual subsequent events.   Instead, we have viewed them as exogenous at any moment.

To illustrate the point, consider the classical model described above in which it is assumed that M+B equals zero initially, so that "money is a veil." In that model, Y, N, $\frac{w}{p}$, C, I, and r are all ir¹ependent of the level of the money supply. In addition, the real rate of interest, $r - \Pi$, was seen to be independent of the rate of expected inflaton and the money supply. We could write this as

$$r = (r-\Pi)+\Pi$$

$$r = \rho+\Pi$$

where ρ equals $r-\Pi$ and is independent of Π and M. Given r and Y, the job of equation (vii) is to determine the price level:

$$\frac{M}{p} = m(r, Y)$$

or

VII' $$\frac{M}{p} = M(\rho+\Pi, Y) \quad .$$

Given M, Y, ρ, and Π, the equation VII' is capable of determining a unique price level.

But suppose that we require that $\Pi = \dot{p}/p$ where $\dot{p}$ is interpreted as a right-hand derivative. The variable $\dot{p}/p$ is viewed as being determined by the model. The equation

$$\frac{M}{p} = M(\rho+ \frac{\dot{p}}{p} , Y) \quad ,$$

with Y, M, and p given, is capable only of determining a locus of $(p, \dot{p}/p)$ combinations that guarantee portfolio balance. In particular, higher $\dot{p}/p$'s

are compatible with higher p's at a moment in time, so that the momentary

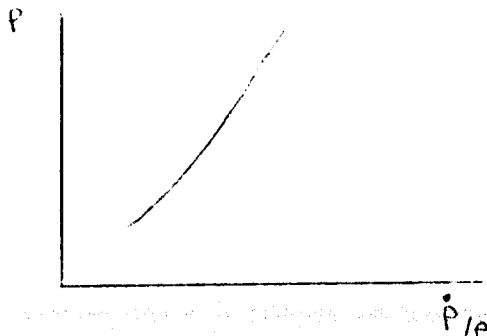$(\dot{p}/p, p)$ locus determined by VII' is like the one shown in figure   .

The value of the exogenous variables at a point in time and the restrictions in equations I through VII don't contain enough information to determine both p and $\dot{p}/p$ at that point in time.

Figure 3

It is useful to pursue this illustration using a concrete example of the model that was proposed by Cagan.* Ignoring the role of income and the real interest rate, Cagan assumed that the demand for money obeyed the equation

$$\log_e \frac{M(t)}{p(t)} = \alpha \Pi(t), \text{ where } \alpha < 0.$$

We suppose that the public's expectation $\Pi(t)$ equals the right-hand derivative $d/dt p(t)/p(t) = d/dt \log_e p(t)$. It is convenient to use the operator D defined by

$$Dx(t) = \frac{d}{dt} x(t)$$

$$D^n x(t) = \frac{d^n x(t)}{dt^n}$$

For many purposes, it is legitimate to manipulate D like an algebraic symbol, which accounts for its utility. So the above equation can be rearranged to read

(40) $D \log p(t) + \frac{1}{\alpha} \log p(t) = \frac{1}{\alpha} \log M(t)$.

We think of M as being the exogenous variable, so that this is a one-equation model that is supposed to determine p.

Given only the value of M at a point in time, equation (40) is incapable of determining either p or Dp at that point in time. That is because equation (40) is in effect one equation in two variables, p and Dp, at each point in time.

Rewrite (40) as

$$(D + \frac{1}{\alpha}) \log {}_e p(t) = \frac{1}{\alpha} \log {}_e M(t)$$

(41) $\qquad \log_e p(t) = \frac{1/\alpha}{1/\alpha + D} \log {}_e M(t)$.

or

(41a) $\qquad \log {}_e p(t) = \frac{1}{1+\alpha D} \log {}_e M(t)$.

Now recall the geometric expansion $1/(1-\lambda) = 1+\lambda+\lambda^2 + \ldots$, for $|\lambda| < 1$. Similarly, it is appropriate to interpret $1/(1+\alpha D)$ as the analogous expansion $1/(1+\alpha D) = 1-\alpha D + \alpha^2 D^2 - \alpha^3 D^3 + \ldots$
where $-\alpha D$ is treated as if it were λ in the above formula for the geometric expansion. So we have

$$\log {}_e p(t) = [1-\alpha D + \alpha^2 D^2 - \alpha^3 D^3 + \ldots] \log_e M(t)$$

(42) $\qquad \log {}_e p(t) = \log_e m(t) - \alpha D \log_e m(t) + \alpha^2 D^2 \log_e M(t) - \ldots$

which expresses $\log_e p(t)$ in terms of an infinite number of right-hand derivatives of $\log M(t)$ with respect to time at t. Equation (42) can be derived directly, without use of the D operator, by repeatedly differentiating the demand schedule for money with respect to time and eliminating successively higher derivatives of $\log p(t)$. This indicates that to determine $\log p(t)$ it is necessary simultaneously to determine time derivatives of all orders of $\log_e p(t)$, which requires taking into account time derivatives of $\log_e M(t)$

of all orders at time t. But if we know time derivatives of all orders of $\log_e p(t)$ at t doesn't the logic of Taylor's expansion imply that we know the values of $\log_e p(s)$ for $s \geq t$ also? The answer is yes, as the following alternative interpretation of the solution (41) readily shows. Notice that

$$\frac{1/\alpha}{1/\alpha+D} = -\frac{1}{\alpha} \int_t^\infty e^{(s-t)(1/\alpha+D)} ds$$

$$= -\frac{1}{\alpha} \cdot [\frac{e^{(s-t)(1/\alpha+D)}}{1/\alpha+D}]_{s=t}^{s=\infty}$$

So (41) can be written

$$(43) \qquad \log_e p(t) = [-\frac{1}{\alpha} \int_t^\infty e^{(s-t)(1/\alpha+D)} ds] \log_c M(t) .$$

Now consider $e^{(s-t)D}$. Taking its Taylor's expansion about $(s-t)D = 0$ we have

$$e^{(s-t)D} = 1+(s-t)D+\frac{(s-t)^2 D^2}{2!} + \ldots \quad .$$

So we have

$$e^{(s-t)D}x(t) = x(t)+(s-t)Dx(t)+(s-t)^2 D^2 x(t) = x(t+s-t) = x(s),$$

by Taylor's expansion. We will exploit this by writing (43) as

$$\log_e p(t) = -\frac{1}{\alpha} \int_t^\infty e^{(s-t)/\alpha} e^{(s-t)D} \log_e M(t) ds$$

Since $e^{(s-t)D}M(t) = M(s)$, the above equation can be written as

$$(44) \qquad \log_e p(t) = -\frac{1}{\alpha} \int_t^\infty e^{(s-t)/\alpha} \log_e M(s) ds .$$

It can be verified by differentiation that (44) is a solution to the differential equation (40), where D is interpreted as a right-hand derivative

operator. Equation (44) is the particular solution of (40) corresponding to the "forcing" function M(t). As always we must add to (44) the solution to the homogeneous equation

$$[D+1/\alpha] \log_e p(t) = 0,$$

which is easily verified to be $ce^{-t/\alpha}$ where c is an arbitrary constant chosen to satisfy some condition. If we impose the "reasonable" condition that the public expects inflation to occur only if the money supply is increasing, this has the effect of making c=0 so that (44) is the appropriate solution.

According to equation (44), the value of $\log_e p(t)$ at t depends on the value of the entire money supply path from t until forever. So values of the endogenous variable p(t) cannot be determined solely from the values of the exogenous variables at t. Instead, the entire time path of M(s) for s $\geq$ t must be supplied. We simply can't do static or point in time exercises. We are thrust into a dynamic analysis. Of course, a dividend of this is that we do have the solution for the entire time paths of the endogenous variables.

An Alternative Definition of Disposable Income

Under the definition of disposable income used up to now it has "mattered," in the sense of affecting the equilibrium values of some of the variables at a point in time, how the government finances its expenditures, G, as between taxes on the one hand, and printing bonds and money on the other hand. At the same time, it has not mattered in what proportion the government finances its deficit, $p(G-T)$, by printing bonds or money.* (We know that the division between $\dot{M}$ and $\dot{B}$ hasn't mattered because, while G and T both appeared in the system of equations I through VII, neither $\dot{M}$ nor $\dot{B}$ did.) These features of our model stem directly from the definition of expected real disposable income that we have employed. Here we show how these aspects of the model are altered when an alternative and more "illusion free" concept of disposable income is utilized.

The alternative definition of disposable income is derived from a definition of wealth that subtracts the discounted value of future interest payments on the current stock of government debt from the definition of wealth used to derive our previous disposable income concept. That definition of real wealth was

$$W = qK + \frac{M}{p} + \frac{B}{p}$$

Our new definition is

$$W_1 = qK + \frac{M}{p} + \frac{B}{p} - \frac{DTL}{p}$$

---

*Of course, the rate of growth over time of M+B is equal to the nominal deficit per unit time. How the deficit is financed affects the evolution of M and B over time, and does matter once time is permitted to elapse.

where DTL denotes the present value of the future interest payments on the current interest-bearing government debt. The rationale behind this definition is that these debt service charges are ultimately charges that the public is liable to pay through higher taxes in the future. The interest on the current government debt is rB. The present value of future interest charges on the current debt is

$$DTL = \int_{t_o}^{\infty} r \, B \, e^{-r(t-t_o)} dt$$

$$= B \ .$$

Thus, W_1 can now be written as

$$W_1 = qK + \frac{M}{p}$$

Expected real disposable income is now defined as the rate at which households can consume while expecting to leave real wealth W_1 intact. The following definition of disposable income satisfies this definition:

$$Y_D' = Y - T - \delta K - \frac{\dot{\overline{DTL}}}{p} - \frac{M}{p}\Pi + (q-1)I$$

$$= Y - T - \delta K - \frac{\dot{B}}{p} - \frac{M}{p}\Pi + (q-1)I$$

Substituting the national income identity and the government budget constraint into the above definition of disposable income yields

$$Y_D' = C + qI + \frac{\dot{M}}{p} - \frac{M}{p}\Pi$$

which verifies tht Y_D' is the amount individuals can consume while expecting that they are leaving their real wealth W_1 intact.

The concept Y_D' differs from Y_D in two ways: it excludes expected capital losses on the interest-bearing portion of the public

debt, since interest-bearing government debt is not a part of wealth; and it deducts $\frac{\dot{B}}{p}$, which equals the real rate of increase of discounted tax liabilities that results from printing bonds at the rate $\dot{B}$ per unit time.

Now consider the effects of replacing Y_D with Y_D' in the consumption function (IV). The consumption function is the only relationship in which T appeared in the earlier version of our model. We immediately notice an important change resulting from using our new definition of disposable income. In the old version of the model neither $\dot{M}$ nor $\dot{B}$ appeared in any of the equations determining the levels of our seven variables. That is what sufficed to show that it didn't matter how the government financed its deficit. Now, however, $\dot{B}$ appears in our new version of equation IV

$$(IV') \qquad C = C(Y - T - \delta K - \frac{\dot{B}}{p} - \frac{M}{p} \Pi + (q-1)I, \ r-\Pi)$$

We can immediately show that now, given the level of G, it doesn't matter whether the government uses taxes or bond printing to finance its expenditures. If it offsets a change in taxes of dT by a change in bond-printing just sufficient to continue financing the current level of G, there is no effect on disposable income, and so no change in aggregate demand, given $r-\Pi$. For holding G and $\dot{M}/p$ constant, we have from the government budget constraint

$$dT = -d \ (\frac{\dot{B}}{p})$$

Such a simultaneous change in taxes and bond-printing is easily seen to leave ˜ disposable income Y_D', unaltered, and so to leave aggregate demand unaffected. It follows that none of our seven variables is affected

by changes in the extent to which government expenditures are financed by
T on the one hand or $\dot{B}/p$ on the other.

Under certain very special conditions it is possible to show
that a more general conclusion follows, namely that it doesn't matter
whether the government finances its expenditures by collecting taxes,
printing bonds, or printing money. In particular, we assume that the
interest elasticity of the demand for money is zero, i.e., $m_r = 0$. We
also assume that at each instant $\Pi = \dot{p}/p$. From our discussion in section
, we know that this last assumption cannot be made lightly. We will
proceed by assuming that while $\Pi = \dot{p}/p$ at any moment, people expect this
same rate of inflation to prevail over the indefinite future. This means
that we don't equate the expected magnitude $\dot{\Pi}$ with the right-hand time
derivative of $\dot{p}/p$. Instead, we simply assume that $\dot{\Pi}$ is zero. Thus people
are assumed to foresee the first time derivative of the logarithm of the
price level, but not the second time derivative. This assumption makes
$\dot{\Pi}$ exogenous and permits us to determine both p and $\dot{p}/p$ at an instant,
given the exogenous variables and their right-hand time derivatives.

Differentiating the portfolio balance condition with respect
to time and solving for $\dot{p}/p$ we have

$$\dot{p}/p = \frac{\dot{M}}{M} - m_{\frac{Y\underline{M}}{p}}\frac{\dot{Y}}{Y}\ .$$

The system that we describe is one in which $\dot{Y}$ can be viewed as being
determined by the right-hand time derivatives of equations I through III.
In addition, the system is one in which $\dot{K}$, which is determined by equations
IV, V, and VI, is independent of Π. It then follows that $\dot{Y}$ will be inde-
pendent of Π and $\dot{M}/M$. Taking the total differential of the above equation
for $\dot{p}/p$, which equals Π, gives

$$d\Pi = d\left(\frac{\dot{M}}{M}\right) - \frac{m_Y\left(\frac{m}{p}d\dot{Y} - \dot{Y}m_r d\Pi\right)}{\left(\frac{M}{p}\right)^2} \quad .$$

Since $d\dot{Y}$ will always be zero for the experiments that we will describe, and since $m_r = 0$, we write the above differential as

$$d\Pi = d\left(\frac{\dot{M}}{M}\right)$$

$$= \frac{d\dot{M}}{M} - \frac{\dot{M}}{M}\frac{dM}{M} \quad .$$

Since we will be studying a case in which $dM = 0$, we have

$$d\Pi = \frac{d\dot{M}}{M} \quad .$$

Now we seek to evaluate the effect on disposable income $Y_D{}'$ of an increase in money creation that results from financing more of government expenditures by money creation and less by taxes and bond-printing. One effect on disposable income comes about as the result of the higher rate of expected inflation that accompanies the higher actual rate of inflation caused by the higher rate of money creation. The effect on disposable income of the higher rate of money creation is found by evaluating $d\left(\frac{M}{p}\Pi\right)$, which is the expected capital loss on the real money supply. We have

$$d\left(\frac{M}{p}\Pi\right) = \Pi m_r\, dr + \frac{M}{p}\, d\Pi = \Pi m_r\, d\Pi + \frac{M}{p}\, d\Pi$$

$$= \Pi m_r\, \frac{d\dot{M}}{M} + \frac{d\dot{M}}{p}$$

But since we are assuming that $m_r = 0$, we have $d\left(\frac{M}{p}\Pi\right) = \frac{d\dot{M}}{p}$.

Now note that

$$d\left(\frac{\dot{M}}{p}\right) = \frac{d\dot{M}}{p} - \frac{\dot{M}}{p}\frac{dp}{p}$$

and that since $m_r = 0$ the rise in Π does nothing to cause p to change.

Thus we have that

$$d(\frac{M}{p}\Pi) = d(\frac{\dot{M}}{p}) \ .$$

Taking the total differential of the government budget constraint gives

$$dG = dT + d(\frac{\dot{M}}{p}) + d(\frac{\dot{B}}{p}) \ ,$$

or since $d(\frac{M}{p}\Pi) = d(\frac{\dot{M}}{p})$,

$$dG = dT + d(\frac{\dot{B}}{p}) + d(\frac{M}{p}\Pi) \ .$$

It follows that on our assumptions disposable income is not affected by changing from one means of financing government expenditures to another, given the level of government expenditures, for

$$dY_d' = d(Y +(q-1)I - \delta K) - dT - d(\frac{\dot{B}}{p}) - d(\frac{M}{p}\Pi)$$

$$= d(Y +(q-1)I - \delta K) - dG.$$

Thus it doesn't matter how the government finances its expenditures, whether by taxes, money creation, or bond creation.

This is a system in which the traditional rationale for taxing rather than printing bonds and money breaks down. For notice that since $S+C = Y_D'$, we have

$$S+C = Y-\delta K - T + (q-1)I - \frac{M}{p}\Pi - \frac{\dot{B}}{p} \ .$$

Substituting the national income identity and the government budget constraint into the above equation and rearranging yields

$$S = qI + \frac{\dot{M}}{p} - \frac{M}{p}\Pi.$$

Since $d(\frac{\dot{M}}{p}) = d(\frac{M}{p}\Pi)$, it follows that $dS = qdI$, i.e., that saving equals investment at the margin. Whether or not the government finances its expenditures by taxing has no impact on the economy's rate of capital accumulation and hence its rate of physical growth.

Neutrality

Macroeconomic models like the one we have been studying possess a characteristic known as neutrality. Suppose we have a macroeconomic model that consists of m structural equations relating m endogenous variables z_1, ..., z_m and n exogenous variables x_1, ..., x_n. Suppose the first m_1 endogenous variables are expressed in units involving the monetary unit of account, e.g., dollars or dollars per unit of output or dollars per man hour. The remaining $m-m_1$ endogenous variables are measured in "real" units, units not involving the monetary unit of account. Similarly suppose that the first n_1, exogenous variables are measured in units of money, while the remaining $n-n_1$ are not. The system of m structural equations can be written as

$$g_1(z_1, \ldots, z_n, x_1, \ldots, x_n) = 0$$

$$g_2(z_1, \ldots, z_m, x_1, \ldots, x_n) = 0$$

(34)
$$\vdots$$

$$g_m(z_1, \ldots, z_m, x_1, \ldots, x_n) = 0$$

Suppose that for the particular values of the exogenous variables $x_1^{\,o}$, ..., $x_n^{\,o}$ the system is in equilibrium provided that the value of the endogenous variables assume the particular values $z_1^{\,o}$, ..., $z_m^{\,o}$. Those values of the z's and the x's satisfy the system of equations (34). The model is said to possess the property of _neutrality_ if the particular values $\lambda z_1^{\,o}$, ..., $\lambda z_{m_1}^{\,o}$, $z_{m_1+1}^{\,o}$, ..., $z_m^{\,o}$ and $\lambda x_1^{\,o}$, ..., $\lambda x_{n_1}^{\,o}$, $x_{n_1+1}^{\,o}$, ..., $x_n^{\,o}$ for any scalar $\lambda > 0$, also satisfy the equilibrium conditions (34). That is, the system is a neutral one if, on starting

from an initial position of equilibrium, multiplying all those endog-
enous variables that are measured in the monetary unit by a positive
scalar leaves the system in equilibrium.

It is easily verified that the model we have been studying
possesses the property of neutrality. Thus, consider the model formed
by equations I through VII. Multiply all variables measured in nominal
or dollar units by the scalar λ. If the system was initially in
equilibrium at a price level of p_0 and a wage rate of w_0, it will now
have an equilibrium with price λp_0 and wage λw_0, while the equilibrium
values of Y, N, C, I, and r are unchanged.

A system will possess the property of neutrality if it can be
written so that each equation involves only magnitudes measured in dollar-
free or "real" units. Thus, if every where there appears a dollar-denomi-
nated magnitude, like the stock of money or bonds or the wage rate, it is
divided through by another dollar denominated variable, like the price
level or the wage rate, then the model is one that describes relations
among real magnitudes. If this is true, then doubling or tripling the
nominal magnitudes cannot have any influence on the equilibrium values
of those real variables. The system formed by equations I through VII
is an example of one in which variables measured in dollars never appear
except when they are divided by the price level. Consequently multiply-
ing M, B, w, and p by any $\lambda > 0$ does not change the equilibrium values
of the real variables determined by that system. The real variables
determined by the model are Y, N, C, I, r, w/p, and M/p.

Dichotomy

A macroeconomic model is said to dichotomize if a subset of equations can determine the values of all real variables with the level of the money supply playing no role in determining the equilibrium value of any real variable. Given the equilibrium values of the real variables, the level of the money supply helps determine the equilibrium values of all nominal variables that are endogenous, but cannot influence any real variable. In a system that dichotomizes, the equilibrium values of all real variables are independent of the absolute price level. In such a system, "money is a veil."

As we have seen, the system formed by equations I through VII does not dichotomize unless $\Pi(M+B)$ equals zero initially. Now consider the system discussed in section   , which is formed by replacing IV with IV', which we repeat here for convenience:

IV' $C = C(Y - T - \delta K - \dfrac{\dot{B}}{p} - \dfrac{M}{p}\Pi + (q-1)I, \ r-\Pi)$

Substituting $m(r, Y)$ for M/p in IV' gives

IV'' $C = C(Y - T - \delta K - \dfrac{\dot{B}}{p} - \Pi m(r, Y) + (q-1)I, \ r- \Pi) \quad .$

The system formed by equations I, II, III, IV'', V, VI, and VII is one that dichotomizes under certain assumptions about financing the government debt. To take a simple special case first. suppose $\dot{B}$ equals zero so that the deficit is financed by printing money. Then equations I, II, III, IV'', V, and VI alone are capable of determining all real variables, it not being necessary to know the money supply and the price level in order to determine the equilibrium values of the real variables. More generally, suppose that

M and $\dot{M}$ always vary proportionately, so that, for example, a doubling of the level of the money supply is accompanied by a doubling of the right-hand time derivative of the money supply. Under this constraint, the system under study does dichotomize. This can be seen most easily if the government flow budget constraint is used to rewrite IV'' as

$$C = C(Y - T - \delta K - (G-T) + \frac{\dot{M}}{p} - m(r, Y)\Pi + (q-1)I, r-\Pi)$$

Suppose the system is initially in equilibrium. Then doubling M, $\dot{M}$, p, and w and keeping all real variables unchanged leaves the system in equilibrium. No change in aggregate demand at the initial r and Y is induced by such proportional changes in M, $\dot{M}$, p, and w.

Obviously, "neutrality" and "dichotomy" are distinct concepts. A system in which there is neutrality need not dichotomize, while a system in which "money is a veil" need not be one that satisfies our definition of a neutral system. To take an example of a system that dichotomizes but in which neutrality fails, consider the system that consists of equations I, II, III, V, VI, VII and the following replacement for IV:

IV'' $C = C(Y - T - \delta K + (q-1)I, r-\Pi , M+B)$

$$1 > C_1 > 0, C_2 < 0, C_3 > 0 .$$

Under the constraint an open-market operations, this is a system that dichotomizes. Increasing the money supply subject to dM + dB = 0 increases the price level and money wage proportionately but leaves all real variables unaffected. The differentials of the first six equations involve only the differentials of real variables which is what yields the dichotomy. On the other hand, neutrality is not a characteristic of the system since

increasing M, B, w, and p proportionately will not leave the system in

equilibrium if we start from a position of equilibrium. Instead, the

increase in M+B increases intended consumption, causing the interest rate

to rise and the price level and money wage to rise more than proportionately

with the increase in the stocks of money and bonds.

Conclusions

In the classical model, the level of employment is determined
in the labor market. The assumption of perfectly flexible money wages
and prices implies that the labor market "clears", a real wage being deter-
mined at which the quantity of labor demanded by firms exactly equals that
which workers are willing to supply. There can be no involuntary unem-
ployment, since the volume of employment is always equal to the labor
supply forthcoming at the prevailing real wage. Everyone who wants to
work at the existing real wage is employed. If for some reason unemploy-
ment were to emerge because of deficient aggregated demand, the unemployed
workers would bid down the money wage. Since firms are competitive, this
would lead in the first instance to proportionate reductions in the price
level, which would in itself leave the real wage unchanged and do nothing
to alleviate unemployment. But the reduction in the price level would
increase the supply of real balances, causing the interest rate to fall.
That would increase the level of aggregate demand for goods and services
and cause the level of employment to rise.

Aggregate demand plays no role in determining the volume of
employment and output, although it does affect their rates of growth
over time. Instead, exogenous changes in aggregate demand cause what-
ever adjustments in the interest rate that are necessary to bring the
demand for goods into equality with the supply of goods, which is deter-
mined in the labor market. The interest rate plays the role of totally
alleviating any "Keynesian" problems of deficient aggregate demand.

Finally, the role of money depends on the precise definition of
perceived real disposable income. Whatever that definition, however, the

money supply plays no role in determining the levels of employment and
output, although as we have seen, under some circumstances it may affect
their rates of growth over time.

NOTES ON THE "KEYNESIAN" MODEL

The standard "Keynesian" model has as its outstanding characteristics that the full employment of labor is not automatic and that the volume of employment at a point in time is in part determined by monetary and fiscal policies. These features stand in sharp contrast to those of the classical model, in which full employment is guaranteed and fiscal and monetary variables have no impact on output and employment at a point in time, though they may influence the growth rates of employment and output.

We shall consider the Keynesian model as consisting of the following six equations:

$$(1) \quad Y = Y(K,N)$$

$$(2) \quad \frac{w}{p} = Y_N$$

$$(3) \quad C = C(Y-T- \delta K - \frac{M+B}{p} \pi, \ r-\pi)$$

$$(4) \quad I = I(q(K,N,r-\pi,\delta)-1)$$

$$(5) \quad C + I + G + \delta K = Y$$

$$(6) \quad \frac{M}{p} = m(r,Y)$$

There are two differences, one inessential and the other essential, between this Keynesian model and the classical model presented above.

To take the inessential difference first, the consumption function in the above model incorporates a simpler concept of disposable income than the one used in our classical model. In particular, the disposable income concept $Y - T - \delta K - \frac{M+B}{p} \pi$ ignores any effects of discrepancies between the equity-market and reproduction-cost value of the stock of capital, i.e., it ignores the effects of movements of q away from unity. In effect, the disposable income concept used here corresponds to a wealth concept

$$w = \frac{M+B}{p} + K,$$

in the sense that it is the rate that individuals perceive they can consume while leaving intact their real wealth so defined. We use this simplified definition of disposable income to streamline the presentation somewhat. Little harm is thereby done, since the ingredients for analyzing things under the more complicated definition of income have been presented with the classical model.

The essential difference between the classical model and the Keynesian model is the absence of the classical labor supply curve combined with labor market equilibrium condition from the latter. Since there is one less equation in the Keynesian model, it can only determine six endogenous variables instead of the seven that are determined in the classical model. To close the Keynesian model, the money wage w is regarded as a parameter, one that at a point in time can be regarded as being given from outside the model, perhaps from the past behavior of itself and other endogenous or exogenous variables. While w is regarded as a datum, unaffected by changes in the other parameters of the model, we need not assume that it is constant through time. The time deriva-

tive dw/dt can be nonzero and might very well itself be functionally
related to the variables of the model.[1] All that we require is that
the value of dw/dt implied by any such relationship be finite so that w
cannot jump at a point in time as a result of its interactions with
other endogenous and exogenous variables.

It bears emphasizing that the equation that we have deleted in
moving from the classical to the Keynesian model, $N = N^S(w/p)$, is a
combination of the two equations

$$N^S = N^S(\frac{w}{p})$$

and

$$N = N^S,$$

the first being a supply schedule, the second being an equilibrium
condition. Note that we continue to require that employment satisfy the
labor demand schedule (2). We shall think of the labor supply schedule
$N^S = N^S(w/p)$ as being satisfied and helping to define the unemployment
rate. But since N^S nowhere appears in equations (1) - (6), it plays no
role in determining the equilibrium position of the model. The force of
removing the classical equation $N = N^S(w/p)$ should be thought of as
removing the labor market equilibrium condition. Indeed, so long as we
don't require the equality $N = N^S$, the labor supply schedule $N^S = N^S(w/p)$
can be appended as a seventh equation of the Keynesian model, one that
is easily seen not to interact with the first six equations. Usually,
the model is assumed to reach equilibrium in a position satisfying
$N < N^S$, so that there is an excess supply of labor.

_____

[1] The Phillips curve is an example of what we have in mind.

Removing the labor supply schedule--labor market equilibrium condition and making the money wage exogenous are the essential changes that must be made in the classical model in order to arrive at the Keynesian model. The asset structure, the constraints on the behavior of the government, and the assumptions underlying the remaining six equations are all as they were in the classical model. Thus, the Keynesian model consists of six equations in the six endogenous variables, p, Y, N, r, C, and I. The exogenous variables are w, M, G, T, π, δ, and K and the parameters determining the shapes of the various schedules.

We begin our analysis of the model by totally differentiating equations (1) - (6). In what follows we will assume $dK = d\delta = 0$. Then we obtain the following system of six equations in the differentials of our six variables:

i $\qquad dY = Y_N \, dN$

ii $\qquad \dfrac{dw}{w} - \dfrac{dp}{p} = \dfrac{Y_{NN}}{Y_N} \, dN$

iii $\qquad dC = C_1 dY - C_1 dT - C_1 \dfrac{M+B}{p} \, d\pi - C_1 \pi \left(\dfrac{dM+dB}{p} - \dfrac{dp \,(M+B)}{p^2} \right)$

$\qquad\qquad + C_2 dr - C_2 d\pi$

iv $\qquad dI = I'q_N dN + I'q_{r-\pi} \, dr - I'q_{r-\pi} \, d\pi$

v $\qquad dC + dI + dG = dY$

vi $\qquad \dfrac{dM}{p} - \dfrac{dp}{p} \dfrac{M}{p} = m_r dr + m_Y dY$

The differential of consumption can be simplified by recalling the constraint on open market operations, $dM + dB = 0$. The system can then be written as the following matrix equation:

$$
\begin{bmatrix}
1 & -Y_N & 0 & 0 & 0 & 0 \\[6pt]
0 & \dfrac{Y_{NN}}{Y_N} & 0 & 0 & 0 & -\dfrac{1}{p} \\[10pt]
-C_1 & 0 & 1 & 0 & -C_2 & -C_{1\pi}\dfrac{M+B}{p^2} \\[10pt]
0 & -I'q_N & 0 & 1 & -I'q_{r-\pi} & 0 \\[6pt]
1 & 0 & -1 & -1 & 0 & 0 \\[6pt]
m_Y & 0 & 0 & 0 & m_r & \dfrac{M}{p^2}
\end{bmatrix}
\begin{bmatrix}
dY \\[6pt] dN \\[6pt] dC \\[6pt] dI \\[6pt] dr \\[6pt] dp
\end{bmatrix}
=
\begin{bmatrix}
0 \\[6pt]
\dfrac{dw}{w} \\[10pt]
-C_1 dT - \left(C_2 + C_1\dfrac{M+B}{p}\right)d\pi \\[10pt]
I'q_{r-\pi} \\[6pt]
dG \\[6pt]
dM/p
\end{bmatrix}
$$

Inspection of the matrix on the left reveals that the system is _not_ block recursive. That is, it is impossible to find an independent subset of equations that determines a subset of variables. Instead, interdependence is general, interactions occurring among all six variables. As a consequence, we should expect that a change in any one parameter will in general affect the equilibrium values of all of the variables at a point in time.

In order to solve the system, we will utilize Hicks' IS-LM curve apparatus. This simply entails adopting the strategy of collapsing equations (i-vi) into a system of two equations in dY and dr, this being accomplished by eliminating dN, dp, dC, and dI by substitution. First we obtain the total differential of the IS curve, a schedule that shows the locus of combinations of the interest rate and output that satisfy equation 5, the aggregate demand-aggregate supply equality.

Substituting iii and iv into v yields

(7) $\quad C_1 dY - C_1 dT - C_1 \frac{M+B}{p} d\pi + C_1 \pi \frac{M+B}{p} \frac{dp}{p} + C_2 dr - C_2 d\pi$

$\quad + I'q_N dN + I'q_{r-\pi} dr - I'q_{r-\pi} d\pi + dG = dY$

Solving (i) for dN and (ii) and (i) for $\frac{dp}{p}$ yields

$\quad dN = \frac{1}{Y_N} dY,$

$\quad \frac{dp}{p} = \frac{dw}{w} - \frac{Y_{NN}}{Y_N^2} dY$

Substituting these two expressions into (7) yields

$\quad C_1 dY - C_1 dT - C_1 \frac{M+B}{p} d\pi + C_1 \pi \frac{M+B}{p} (\frac{dw}{w} - \frac{Y_{NN}}{Y_N^2} dY) + C_2 dr$

$\quad - C_2 d\pi + I'\frac{qN}{YN} dY + I'q_{r-\pi} dr - I'q_{r-\pi} d\pi + dG = dY$

which is the total differential of the IS curve. Rearranging yields,

(8) $\quad (1 - C_1 - I'\frac{q_N}{Y_N} + C_1 \pi \frac{M+B}{p} \frac{Y_{NN}}{Y_N^2}) dY = -C_1 dT + dG + (C_2 + I'q_{r-\pi})dr$

$\quad + (-I'q_{r-\pi} - C_2 - C_1 \frac{M+B}{p}) d\pi + C_1 \pi \frac{M+B}{p} \frac{dw}{w}$

As with our analysis of the classical model, we shall begin by assuming that M+B equals zero. Then (8) simplifies to

(9) $\quad (1 - C_1 - I'\frac{q_N}{Y_N}) dY = -C_1 dT + dG + (C_2 + I'q_{r-\pi})dr$

$\quad + -(I'q_{r-\pi} + C_2)d\pi$

The slope of the IS curve is given by

(10) $\quad \frac{dr}{dY}\bigg|_{IS} = \frac{1 - C_1 - I'\frac{q_N}{Y_N}}{C_2 + I'q_{r-\pi}}$

The denominator of the right side is negative, while the numerator may be of either sign. It is positive if

$$1 - C_1 > I' \frac{q_N}{Y_N}$$

Now $1 - C_1$ is the marginal propensity to save out of disposable income. The expression $I'{}^{q_N}/Y_N$ represents the marginal propensity to invest out of income; to see this, differentiate the investment schedule partially with respect to Y:

$$\frac{\partial I}{\partial Y} = \frac{\partial I}{\partial q} \frac{\partial q}{\partial N} \frac{\partial N}{\partial Y} = I' q_N \frac{1}{Y_N}$$

Accordingly, the IS curve will be downward sloping only if the marginal propensity to save exceeds the marginal propensity to invest out of income. We will henceforth assume that this condition is satisfied. (The model can be stable even if this condition is violated, provided at least that the IS curve is less steep than the LM curve.)

To determine how the IS curve shifts when the parameters T, G, and π change, we use (9) to evaluate the partial derivatives of r with respect to each parameter, dY being set equal to zero:

$$\left.\frac{\partial r}{\partial T}\right|_{IS} = \frac{C_1}{C_2 + I' q_{r-\pi}} < 0$$

$$\left.\frac{\partial r}{\partial G}\right|_{IS} = \frac{-1}{C_2 + I' q_{r-\pi}} > 0$$

$$\left.\frac{\partial r}{\partial \pi}\right|_{IS} = 1 \; .$$

The IS curve shifts upward when taxes decrease or government expenditures increase; when anticipated inflation increases, the IS curve shifts upward by the amount of the increase in anticipated inflation.

The LM curve depicts the combinations of interest rates and output which create equality between the demand and supply for real balances. The differential of the LM curve is derived by using (i) and (ii) to eliminate dp from (vi). Thus,

$$\frac{dp}{p} = \frac{dw}{w} - \frac{Y_{NN}}{Y_N^2} \, dY$$

Substituting the expression on the right for dp/p in (vi) yields

$$\frac{dM}{p} - \frac{M}{p} \frac{dw}{w} + \frac{M}{p} \frac{Y_{NN}}{Y_N^2} \, dY = m_r dr + m_Y dY$$

Solving for dr yields

$$(11) \qquad dr = \frac{1}{m_r} (\frac{dM}{p} - \frac{M}{p} \frac{dw}{w} + (\frac{Y_{NN}}{Y_N^2} \frac{M}{p} - m_Y) dY)$$

The slope of the LM curve is thus given by

$$\frac{dr}{dY}\bigg|_{LM} = \frac{1}{m_r} (\frac{Y_{NN}}{Y_N^2} \frac{M}{p} - m_y) > 0$$

Notice that as m_r approaches zero, the LM curve approaches a vertical position while as $m_r \rightarrow -\infty$, as is supposed in the case of the "liquidity trap", the slope of the LM curve approaches zero. By using (11) it is easily established how the LM curve shifts when w or M changes. Thus, setting dY = 0 in (11) and solving for dr we have

$$dr = \frac{1}{m_r} \frac{dM}{p} - \frac{M}{p} \frac{dw}{w}$$

This expression establishes that for a given level of income, if the money supply rises, the interest rate must fall (provided $-\infty < m_r$) in

order to maintain equilibrium between the demand and supply for money.
Similarly, when money wages fall, the interest rate must also fall, again
provided $-\infty < m_r$, in order to keep the money market in equilibrium.

An interest rate and an income level that satisfy the six
equations of the model are determined at the intersection of the IS and
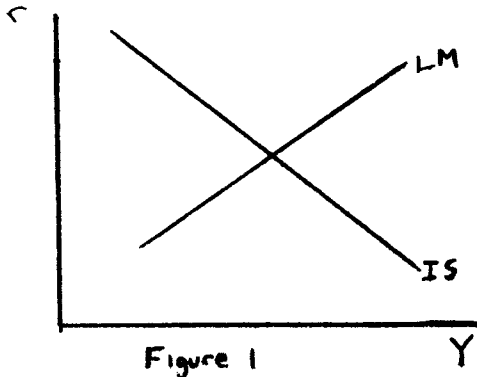LM curves, as depicted in Figure 1. As long as the IS curve is downward
sloping and the LM curve is upward
sloping, we are assured that the
equilibrium is unique (if it exists).
Using our knowledge of how the IS and
LM curves shift when the various para-
meters of the model change, it is easy

Figure 1

to ascertain the effects of those changes on the equilibrium rate of
output and rate of interest. For example, if the level of government
expenditures rises, the IS curve shifts upward and to the right, while
the LM curve is left unchanged. This drives both income and the inter-
est rate upward, provided that the LM curve is not vertical--in which
case the interest rate alone rises--or horizontal--in which case only
the rate of output rises. An increase in the money supply shifts the LM
curve toward the right, which drives the interest rate down and output
upward, provided that the LM curve is not horizontal.

We can solve the system analytically by substituting (11) into
(9). We obtain

$$(1 - C_1 - I' \frac{q_N}{Y_N})dY = - C_1 dT + dG - (C_2 + I'q_{r-\pi})d\pi$$

$$+ \frac{(C_2+I'q_{r-\pi})}{m_r} = (\frac{dM}{p} - \frac{M}{p}\frac{dw}{w} + (\frac{Y_{NN}}{Y_N^2}\frac{M}{p} - m_Y)dY)$$

or

(12) $\qquad (1 - C_1 - I'\frac{q_N}{Y_N} - \frac{(C_2 + I'q_{r-\pi})}{m_r} (\frac{Y_{NN}}{Y_N{}^2} \frac{M}{p} - m_Y))dY = - C_1 dT$

$$+ \ dG - (C_2 + I'q_{r-\pi})d\pi \ + \frac{(C_2 + I'q_{r-\pi})}{m_r} (\frac{dM}{p} - \frac{M}{p}\frac{dw}{w})$$

This is the total differential of the reduced form for Y. On our assumption that $1 - C_1 > I'\frac{q_N}{Y_N}$, the coefficient on dY is positive (notice that it may be positive even if that condition is violated). Let H stand for the coefficient on dY in (12). Then the partial derivatives of Y with respect to T, G, π, M, and w are given by:

$$\frac{\partial Y}{\partial T} = \frac{-C_1}{H} \leq 0$$

$$\frac{\partial Y}{\partial G} = \frac{1}{H} \geq 0$$

$$\frac{\partial Y}{\partial \pi} = \frac{I'q_{r-\pi} + C_2}{H} \geq 0$$

$$\frac{\partial Y}{\partial M} = \frac{C_2 + I'q_{r-\pi}}{m_r pH} \geq 0$$

$$\frac{\partial Y}{\partial w} = \frac{(I' - C_2) M}{m_r P \ wH} \leq 0$$

Notice that as $m_r \to 0$, $H \to \infty$ so that dY/dG, dY/dT, and dY/dπ approach zero. On the other hand, as $m_r \to - \infty$, dY/dM and dY/dw both approach zero. Also, if $C_2 = I' = 0$, then dY/dM = dY/dw = 0. Thus, if neither consumption nor investment expenditures respond to variations in the interest rate, changes in the money supply and the money wage have no effects on output. All of those special cases are easy to analyze

graphically using the IS and LM curves.[2] More generally, we might characterize the relative potency of monetary and fiscal policy by a "bang-per-buck ratio" namely

$$\frac{dY}{dG} \Big/ \frac{dY}{\frac{dM}{p}} \, ,$$

or the ratio of the derivative of real output with respect to real government expenditures to the derivative of real output with respect to the nominal money supply divided by the price level. This ratio is given by

$$\frac{dY}{dG} \Big/ \frac{dY}{\frac{dM}{p}} = \frac{m_r}{C_2 + I'q_{r-\pi}}$$

---

[2] Notice that if $m_r = -\infty$, from (12) we have

$$\frac{dY}{dG} = \frac{1}{1 - C_1 - I'\dfrac{q_N}{Y_N}}$$

and

$$\frac{dY}{dT} = \frac{-C_1}{1 - C_1 - I'\dfrac{q_N}{Y_N}}$$

The expression for dY/dG is the "super-multiplier" of Hicks, the reciprocal of the difference between the marginal propensity to save and the marginal propensity to invest out of income. If I' equals zero, so that investment does not respond to variations in income, the above expressions are identical with the standard simple Keynesian multiplier formulas. Also, notice that if I' equals zero, the response of output to a change in government expenditures matched by an equal change in taxes is given by

$$\frac{dY}{dG}\bigg|_{dT=dG} = 1,$$

which is a version of the so-called balanced-budget multiplier.

On this definition, changes in real government expenditures are relatively more potent the more responsive to the interest rate is the demand for real balances and the less responsive to real interest are consumption and investment demand. (Our bang-per-buck measure of the relative potency of monetary and fiscal policies is of very little practical use, since what is important for policymaking in the real world is the relative uncertainty associated with using the two kinds of policy. We are ignoring uncertainty here and hence, except in certain limiting cases, can't say anything about the relative merits of relying on monetary as opposed to fiscal policy to achieve desired real output and price objectives.)

Having determined the equilibrium interest rate and output, the equilibrium levels of N, p, C, and I are easily determined. Thus, N is determined from (1), p from (2), I from (4) and C from (3).

Stability

As in the classical model, it is usual to posit that if aggregate demand exceeds aggregate supply, the price level is bid up, while if the demand for real balances exceeds the supply, the interest rate is bid up. In the Keynesian model, however, output varies with the price level as the system moves along the aggregate supply schedule

$$p = p\ (Y,w,K),$$

the schedule that we implicitly derived above by eliminating N from equations (1) and (2). As shown above, the partial derivative of p with respect to Y along this curve is given in

$$\frac{dp}{p} = -\ \frac{Y_{NN}}{Y_N^2}\ dY$$

or

(13) $\qquad dY = -\dfrac{w}{p^2} \dfrac{Y_N}{Y_{NN}} \, dp.$

During the adjustment to equilibrium, we shall insist that the system remain continuously on the aggregate supply curve so that (13) is satisfied continuously.

We posit the following differential equations for p and s, where s is a "time" index:

$$\frac{dp}{ds} = d[C(Y-T- \delta K, r-\pi) + I(q(N,K,r-\pi,\delta)-1)+G+ \delta K-Y]$$

$$\alpha' > 0, \quad \alpha(0) = 0$$

$$\frac{dr}{ds} = \beta [m(r,Y) - {}^M/_p] \qquad\qquad \beta' > 0, \quad \beta(0) = 0$$

Taking a linear approximation to these equations around equilibrium values and using (13) we have

$$\frac{dp}{ds} = \alpha' \; (1-C_1-I'\frac{q_N}{Y_N}) \; \frac{w}{p^2} \; \frac{Y_N}{Y_{NN}} \; (p-p_0) + \alpha'(C_2 + I'q_{r-\pi})(r-r_0)$$

$$\frac{dr}{ds} = \beta' \; (\frac{m}{p^2} - m_Y \frac{w}{p^2} \frac{Y_N}{Y_{NN}}) \; (p-p_0) + \beta'm_r \; (r-r_0),$$

where zero subscripts denote initial equilibrium values. This is a first-order differential equation whose characteristic equation is

$$\begin{vmatrix} a- \lambda & b \\ c & d- \lambda \end{vmatrix} = \lambda^2-(a+d) \lambda + (ad-bc) = 0.$$

where $\qquad a = \alpha' \; (1-C_1-I'\dfrac{q_N}{Y_N}) \; \dfrac{w}{p^2} \dfrac{Y_N}{Y_{NN}}$

$\qquad\qquad b = \alpha' \; (C_2 + I'q_{r-\pi})$

$\qquad\qquad c = \beta' \; (\dfrac{m}{p^2} - m_Y \dfrac{w}{p^2} \dfrac{Y_N}{Y_{NN}})$

$\qquad\qquad d = \beta' \, m_r \qquad .$

Necessary and sufficient conditions for the system to be stable (i.e., have roots with negative real parts) are $-(a+d) > 0$ and $ad-bc > 0$. The first condition can be written

$$(14) \qquad \alpha'(1-C_1-I'\frac{q_N}{Y_N})\frac{w}{p^2}\frac{Y_N}{Y_{NN}} + \beta'm_r < 0$$

while the second condition is

$$m_r(1-C_1-I'\frac{q_N}{Y_N})\frac{w}{p^2}\frac{Y_N}{Y_{NN}} - (C_2 + I'q_{r-\pi})(\frac{M}{p^2} - m_Y\frac{w}{p^2}\frac{Y_N}{Y_{NN}}) > 0$$

which is equivalent with

$$(15) \qquad (\frac{1-C_1-I'\frac{q_N}{Y_N}}{C_2+ I'q_{r-\pi}}) - \frac{1}{m_r}(\frac{M}{w}\frac{Y_{NN}}{Y_N} - m_Y) < 0$$

or

$$\left.\frac{dr}{dY}\right|_{IS} - \left.\frac{dr}{dY}\right|_{LM} < 0,$$

so that a necessary condition for stability is that the LM curve is steeper than the IS curve. This condition is automatically satisfied when the LM curve is upward sloping and the IS curve is downward sloping. It can still be satisfied if the IS curve is upward sloping, provided that the LM curve is more steeply sloped. Since α' and β' are both positive, condition (14) will be satisfied so long as $1-C_1-I'q_N/Y_N > 0$, which is our condition for the IS curve to be downward sloping. If $1-C_1 - I'\ q_N/Y_N < 0$, condition (14) can still be satisfied, but it will depend then on the relative sizes of the adjustment speeds α' and β'.

Some Experiments

We utilize a graphical apparatus (Figure (2) with IS-LM curves in the upper panel, the production function in the lower right panel,
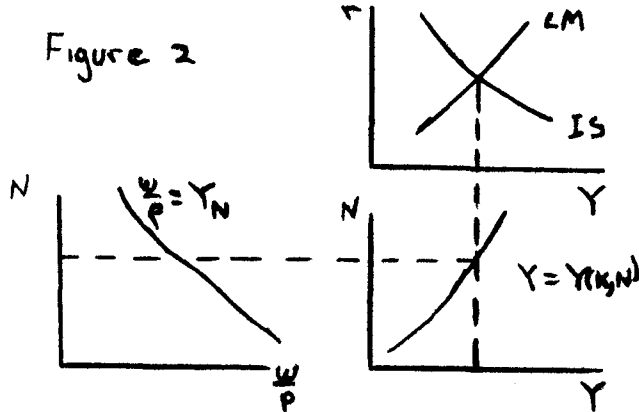
Figure 2

and the demand schedule for labor in the lower left panel. Equilibrium output and interest are determined at the intersection of the IS and LM curves. The production function then determines employment; then given N the demand schedule for labor determines the real wage w/p. Since w is fixed, p is thereby determined.

We analyze the effects of several shocks.

An increase in government expenditures: An increase in G shifts the IS curve outward, sparking an upward movement along the LM curve, thereby in general causing both r and Y to increase. The rise in Y requires a larger N which in turn requires a lower real wage and therefore, with w fixed, a higher price level.

Our stability analysis suggests that the following adjustment pressures materialize to propel the system (instantaneously) from the initial equilibrium to the new one. At the initial values of r and Y, the rise in G causes aggregate demand $(C+I+K+ \delta K)$ to exceed aggregate supply. That causes the price level p to be bid upward, causing the real wage to fall, since w is fixed, and thereby causing employers to expand employment N. So output is driven up. At the initial r and the expanded Y and p, the demand for nominal balances rises, creating an excess demand for money. Households attempt to acquire more money by disposing of bonds and equities but in the aggregate they can't, since

the quantities of money, bonds and equities, are fixed. What happens then is that the interest rate r is bid upward until households are satisfied with the amount of money in existence. The rise in r causes I and C to fall which helps reduce aggregate demand and moderate the magnitude of the increase in Y required to reestablish equilibrium.

An increase in the money supply: An increase in M shifts the LM curve downward, sparking a movement downward along the IS curve, causing r to fall and Y to rise. The increase in Y means that N must rise, w/p must fall, and p must rise. Both C and I rise, due to the effects of the lower interest rate and higher income.

The adjustment pressures can be described as follows. At the initial r, Y combination, the open market operation disturbs portfolio balance, leaving households with more money than they want. Consequently, they attempt to move into bonds and equities and out of money. The effect is to bid the interest rate down. But that in turn stimulates I and C, causing aggregate demand to exceed aggregate supply. This causes the price level to be bid up, forcing a fall in the real wage and creating incentives for entrepreneurs to expand employment. As N increases, so does output Y.

An increase in the money wage (the Keynes effect): An increase in w causes the LM curve to shift to the left, causing a movement upward along the IS curve. This means that r must rise and Y must fall. Then from the production function N must fall, and from the demand schedule for labor w/p must rise. From the portfolio equilibrium condition

$$\frac{M}{p} = m(r,Y),$$

we know that p must rise, since the rise in r and fall in Y cause m(r,Y) to fall; p must rise to cause M/p to fall; C and I both fall, both

because of the fall in Y and the rise in r.

The adjustment pressures work as follows. At the initial
(Y,r) combination, the rise in w causes a proportional increase in p, as
competitive producers simply pass on the wage increase. The rise in p
stimulates the demand for nominal balances, thereby upsetting portfolio
equilibrium and leaving an excess demand for money. The interest rate
is bid upward as households attempt to dispose of bonds and equities and
acquire money. The rise in r in turn restricts I and C, causing aggre-
gate demand to be less than aggregate supply. This causes prices to
start falling, causing the real wage to rise, employment to fall, and
output to decrease.

Thus, an increase in the money wage works very much like a de-
crease in the money supply. The adverse effect of an increase in the
money wage on employment and output is commonly known as the "Keynes
effect."

The reader is invited to analyze the effects of changes in
other exogenous variables, and to analyze the adjustment pressures
driving the system toward the new equilibrium in each case.

Analysis of "Cost-Push" and Demand-Pull" Inflation

It should be apparent that we could easily have solved the
system (i) through (vi) by collapsing it into two equations in another
pair of variables, for example dp and dY. Thus, the solution of (1) and
(2) can be taken to represent the aggregate supply schedule in the p - Y
plane. Its differential, derived from (i) and (ii) is given by

$$dY = \frac{Y_N^2}{Y_{NN}} \frac{dw}{w} - \frac{Y_N^2}{Y_{NN}} \frac{dp}{p}$$

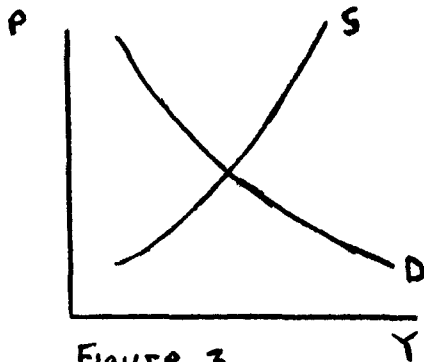Aggregate supply is seen to vary directly with the price level; at a

given price level, aggregate supply increases when the money wage falls. A typical aggregate supply schedule is depicted in Figure 3. An aggregate demand schedule in the p-Y plane can be derived from equations

Figure 3

(3) through (6). Its differential, derived from (iii) through (vi) is given by

$$dY \left(1 - C_1 - I' \frac{q_N}{Y_N} + \frac{C_2 + I'q_{r-\pi}}{m_r} m_Y \right) = -C_1 dT + dG$$

$$- (C_2 + I'q_{r-\pi}) d\pi + \frac{(C_2 + I'q_{r-\pi})}{m_r} \frac{dM}{p} - \frac{(C_2 + I'q_{r-\pi})}{m_r} \frac{M}{p^2} dp.$$

On our usual assumptions, the slope of the aggregate demand schedule is negative, such as is true of the one in Figure 3. The slope of the aggregate demand schedule is

$$\frac{dp}{dY} \Big)_{Dd} = - \left(\frac{1 - C_1 - I'\frac{q_N}{Y_N} + \frac{(C_2 + I'q_{r-\pi})}{M_r} m_Y}{\frac{(C_2 + I'q_{r-\pi})}{m_r}} \right.$$

$$= - \frac{[m_r(I - C_1 - I' \frac{q_N}{Y_N}) + (C_2 + I'q_{r-\pi}) m_Y]}{(C_2 + I'q_{r-\pi})}$$

It is interesting to compute the following limits:

$$\lim_{m_r \to 0} \frac{dp}{dY} \Big)_{Dd} = -m_Y$$

$$\lim_{m_r \to -\infty} \frac{dp}{dY} \Big)_{Dd} = - \infty$$

$$\lim_{(C_2 - I') \to 0} \frac{dp}{dY} \Big)_{Dd} = -\infty$$

As $m_r \to -\infty$, the aggregate demand curve becomes vertical in the p-Y plane. In this "liquidity trap" case, decreases in the money wage, which shift the aggregate supply curve outward in the p-Y plane, do nothing to stimulate real output. Except in limiting cases, increases in G, π, and M and decreases in T will in general shift aggregate demand upward and/or outward in the p-Y plane.

The equilibrium level of income and prices are determined at the intersection of the aggregate demand and supply curves, as depicted in Figure 3. We leave it to the reader to analyze the effects of changes in various parameters by making use of this graphical device. Needless to say, the apparatus will produce the same results as those produced by the IS-LM analysis.

The apparatus consisting of the aggregate demand curve and aggregate supply curve in the p-Y plane can readily be used to explain the distinction between increases in the price level caused by "demand-pull", on the one hand, and "cost-push" on the other. The aggregate supply curve, being the locus of p-Y combinations that solve equations I and II, depends on the capital stock, the money wage, and the parameters of the production function. Changes in these elements of supply, most particularly the money wage rate, are what are called "cost-push" factors that can cause the price level to move. Thus, an exogenous increase in the money wage will shift the aggregate supply curve upward, in general driving the price level up as well as the real wage and the interest rate, while causing employment, output, and firms' real profits to fall. We know that real profits ε_0 fall since

$$\varepsilon_0 = Y - \frac{w}{p} N - (r+\delta-\pi)K \; ;$$

taking the total differential gives

$$d \, \varepsilon_0 = (Y_N - \frac{w}{p})dN - N \, d(\frac{w}{p}) - Kdr = - (N \, d(\frac{w}{p}) + Kdr).$$

We know that both the real wage and interest rate rise when w increases, thus assuring that profits decrease. So do dividends.

"Demand-pull" causes of price changes stem from changes in parameters that affect the position of aggregate demand in the p-Y plane. These parameters include π as well as the government policy variables T, G, and M. Increases in G and M and decreases in T shift aggregate demand upward in the p-Y plane, causing output, employment, the price level, and real dividends to rise while the real wage falls. To see that real dividends rise, recall that real dividends are

$$\text{div} = Y - \frac{w}{p} N - \delta K.$$

Then

$$d(\text{div}) = (Y_N - \frac{w}{p})dN - Nd(\frac{w}{p}) = -Nd\frac{w}{p} \; ,$$

and so dividends move inversely with the real wage. When demand-pull factors push the price level up, the interest rate will rise if the cause is a rise in π or G or a decline in T, while it will fall if a rise in M is what causes the jump in p. Similarly, how real profits move depends on the particular source of demand-pull. For notice that real profits are

$$\varepsilon_0 = Y - \frac{w}{p} N - (r + \delta - \pi)K$$

$$= Y - \frac{w}{p} N - Y_K K + (Y_K - (r + \delta - \pi)K;$$

by Euler's theorem and the marginal productivity condition for labor, we have

$$\varepsilon_0 = (Y_K - (r + \delta - \pi))K.$$

Whether real profits rise or fall depends on what happens to the gap between the marginal product of capital and the real cost of capital. When that gap increases, as will happen when the money supply is increased, profits rise as does investment. But an increase in government expenditures will generally narrow that gap and cause profits and investment to fall.

It is clear that while both "cost-push" and "demand-pull" factors can cause the price level to rise, they do produce effects of opposite sign on the real wage, output, employment, and dividends. It is those differences that could in principle be used to identify particular inflationary episodes as being of the "cost-push" or "demand-pull" variety.

The Model Where $(M+B)\pi \neq 0$

Thus far, our analysis of the Keynesian model has been conducted on the assumption that $M+B$ equals zero, or alternatively on the assumption that perceived real capital gains on the public's holdings of government debt should not be a component of perceived real disposable income. If we relax that assumption, we see from equation (8) that the position of the IS curve will depend on the money wage. Thus, setting dY in (8) equal to zero, we find that dr must satisfy

$$dr = \frac{C_1\pi \frac{M+B}{p}}{I' - C_2} \frac{dw}{w} = -\frac{C_1\pi \frac{M+B}{p}}{C_2 + I'q_{r-\pi}} \frac{dw}{w}$$

if all the other parameters are constant and if the equality between aggregate demand and supply is to be maintained when money wages change. The denominator of this expression is positive, while the numerator may be positive or negative depending on the signs of π and $M+B$. If π and

M+B are both positive, dr is positively related to dw, the IS curve thus shifting upward when w rises. That is because given output, an increase in money wages leads to an increase in prices which depreciates the real value of anticipated capital losses on the public's holdings of bonds and money, thus __raising__ disposable income and aggregate demand. Under these conditions, then, an increase in money wages causes both the IS curve and the LM curve to shift upward. Interest rates are bound to rise but what happens to output depends on the relative shifts and relative slopes characterizing the IS and LM curves.

To determine the solution analytically when $\pi(M+B)$ does not necessarily equal zero, we substitute (11) into (8) to obtain

$$(1 - C_1 - I' \frac{q_N}{Y_N} + C_1 \pi \frac{M+B}{p} \frac{Y_{NN}}{Y_N^2} \, dY = - C_1 dT + dG$$

$$- (I'q_{r-\pi} + C_2 + C_1 \frac{M+B}{p}) \, d\pi + C_1 \pi \frac{M+B}{p} \frac{dw}{w}$$

$$+ \frac{(C_2 + I'q_{r-\pi})}{m_r} (\frac{dM}{p} - \frac{M}{p} \frac{dw}{w} + (\frac{Y_{NN}}{Y_N^2} \frac{M}{p} - m_Y) dY)$$

or

$$(1 - C_1 - I' \frac{q_N}{Y_N} + C_1 \pi \frac{M+B}{p} \frac{Y_{NN}}{Y_N^2} - \frac{(C_2 + I'q_{r-\pi})}{m_r} (\frac{Y_{NN}}{Y_N} \frac{M}{p} - M_Y)) dY =$$

$$- C_1 dT + dG - (I'q_{r-\pi} + C_2 + C_1 \frac{M+B}{p}) \, d\pi$$

$$+ (C_1 \pi \frac{M+B}{p} - \frac{(C_2 + I'q_{r-\pi})}{m_r} \frac{M}{p}) \frac{dw}{w} + \frac{(C_2 + I'q_{r-\pi})}{m_r} \frac{dM}{p}$$

Even provided that π is small, so that the coefficient on dY is positive—— a requirement of stability——an increase in w is seen to have ambiguous effects on output when $\pi(M+B)$ is positive since the coefficient on dw

may be either positive or negative.   This conclusion is consistent with
our analysis of the effects on the IS and LM curves of changes in w.

We leave it to the reader to analyze the effects of changes in
other parameters when (M+B) π is not zero.

The Classical Model Again

Suppose that we now relax the assumption that wages are fixed
at a point in time, allowing wages to change instantaneously in response
to deviations of employment from its full-employment level.   This can be
accomplished by reinstating the labor market equilibrium condition of
the classical model, N = N(w/p).   Once these changes are made, we are
back with the classical model.   In terms of the IS-LM apparatus, the
workings of the classical model are depicted in Figure 4.   Since the IS

Figure 4
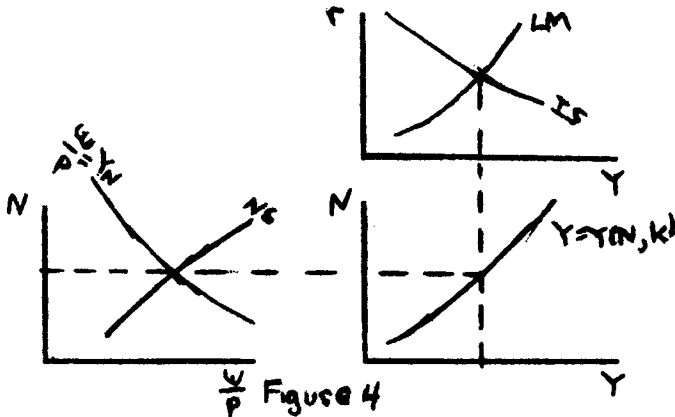
and LM curves are based on six of the
equations in the classical model,
they continue to describe the locus
of (r,Y) combinations that equilibrate
aggregate demand and supply and the
demand and supply for real balances,
respectively, <u>given</u> an assumed money wage.   Let us assume that π(M+B)
equals zero so that the IS curve is independent of the money wage.[3]
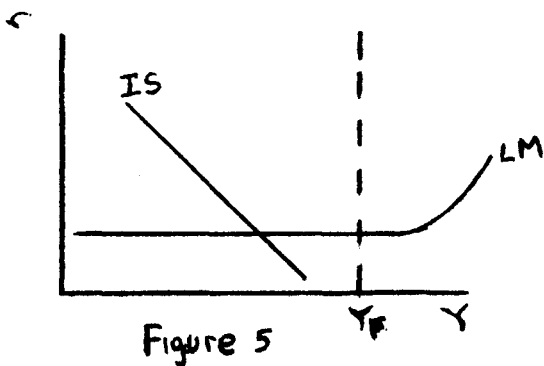Then the model works as follows.   The levels of output, employment, and
the real wage are determined by the supply and demand functions for
labor and the production function, as depicted by the bottom two diagrams
in Figure 2.   The equilibrium interest rate is determined by substituting

---

[3]The reader is invited to figure out how things work
when this specification is dropped.

the equilibrium level of output into the IS curve in order to find the interest rate that is required to call forth an aggregate demand that just equals the aggregate supply. The LM curve thus plays no role in determining the equilibrium interest rate. Instead, given the equilibrium interest rate and rate of output, the LM curve is used to solve for the equilibrium money wage. Thus, the money wage adjusts to such a level that the LM curve always goes through the (r,Y) combination determined in the aggregate supply and aggregate demand sectors of the model. Of course, given the equilibrium wage rate, it is very easy to solve for the price level, since real wages have already been determined in the labor market. Hence, it seems that there cannot be an under-employment equilibrium in the classical model, since flexible wages will assure us that the IS and LM curves intersect at the level of output determined on the supply side.

Notice, however, that the analysis above depends on the "good behavior" of the IS and LM curves. In particular, the analysis breaks down in the case in which $m_r = -\infty$ at some positive interest rate, since then changes in w are incapable of changing the interest rate that at a given level of income equilibrates the money market. The situation is depicted in Figure 5. There the level of income determined on the aggregate supply side, which we indicate by the vertical line at Y_F, is not the same as the level of income determined at the intersection of the IS and LM curves. No amount of wage deflation is capable of

Figure 5

driving the LM curve downward so that it intersects in IS curve at r_0,

the interest rate that is sufficiently low that it generates an aggregate demand equal to Y_F. Thus as we approach a regime in which the Keynesian liquidity trap is relevant, the flexibility of the interest rate becomes limited by the willingness of the public to hold virtually any quantity of money at the ruling interest rate. Injecting more real balances into the system, either through open market operations or money wage decreases, fails to drive the interest rate downward. Thus, the interest rate is prevented from playing its crucial role of equilibrating aggregate supply and demand. Formally, the rate of output Y is overdetermined in this system, one solution for the output rate being determined by the labor market and the production function, another by the intersection of the IS and LM curves. There is no device to reconcile the two solutions for output. The other side of the coin is that the money wage is under-determined since the model in no way restricts the level of w (the model can be written in a form in which w nowhere appears).

A response to this argument was advanced by Pigou.[4] Pigou argued that even in the face of a liquidity trap, full employment in the classical system would be guaranteed if the consumption function were amended to include the real net wealth of the public as an independent argument. In our model the public's real net wealth consists of the sum of the capital stock K and the real debt of the government, (M+B)/p. Then, following Pigou, suppose we replace (3) with the consumption function

(3a) $\qquad C = C(Y - T - \delta K - \pi \dfrac{M+B}{p}, \ r - \pi, \ \dfrac{M+B}{p} + K)$

_____

[4] "The Classical Stationary State," _Economic Journal_, 1943.

where C_3 is assumed to be positive, an increase in the public's net
wealth raising consumption at given levels of real perceived disposable
income and the real interest rate. If this consumption function re-
places (3), the differential of the IS curve, equation (8), must be
amended to

$$(8') \qquad (1 - C_1 - I'\frac{q_N}{Y_N} + C_1\pi \frac{M+B}{p} \frac{Y_{NN}}{Y_N^2} - C_3 \frac{M+B}{p} \frac{Y_{NN}}{Y_N^2}) \, dY =$$

$$- C_1 dT + dG + (C_2 + I'q_{r-\pi}) dr - (I'q_{r-\pi} + C_2 + C_1 \frac{M+B}{p}) \, d\pi$$

$$+ (C_1\pi \frac{M+B}{p} - C_3 \frac{M+B}{p}) \frac{dw}{w} .$$

As before, we assume that the coefficient on dY is positive. To simplify
things, we also assume that π equals zero, an assumption that is usually
made in discussions of the "Pigou effect." Then, we can determine the
effect of money wage changes on the IS curve by setting all differentials
except dr and dw equal to zero in (8'). We find that

$$\frac{dr}{dw} = \frac{C_3 \frac{M+B}{p}}{(C_2 + I'q_{r-\pi})}$$

an expression that is negative so long as M+B exceeds zero. Hence an
increase in the money wage shifts the IS curve downward. By causing
prices to rise, an increase in the money wage diminishes the real value
of the debt owed the public by the government, thus tending to reduce
the intended rate of consumption. Since we assumed that π is zero,
there can occur no countervailing effects on perceived disposable income
that could conceivably offset the effect of the changed level of wealth
on consumption. (On the other hand, if π and M+B were both positive,
an increase in money wages would tend to _increase_ consumption through
its effect on perceived real disposable income. This could offset part
or all of the "wealth effect").

In the context of the Keynesian dilemma depicted in Figure 5, the Pigou effect provides a means of salvaging the automatic nature of full employment in a regime of perfectly flexible money wages. Even if a drop in money wages fails to shift the LM curve downward via the "Keynes effect," it will drive the IS curve outward due to the wealth or "Pigou effect." Provided that this effect is sufficiently potent, eventually the IS curve will be driven far enough out that aggregate demand will be equal to the full employment output Y_F even at the rate of interest determined at the liquidity trap.

It is important to realize that this argument depends on several critical assumptions. Most important, there must occur no effects on perceived disposable income that could offset the wealth effect when wages change. If π is positive, these could conceivably offset the effects on consumption of real wealth changes stemming from jumps in the price level.

A Digression on Wealth, Saving, and the Rate of Interest in the Classical Model

In "Wealth, Saving, and the Rate of Interest,"[5] Lloyd Metzler argued that incorporating the Pigou effect into the classical macro-economic model meant that money ceased to be a "veil." Instead an increase in the money supply could effect the rate of interest and the rate of capital accumulation. Metzler pointed out that whether or not an increase in the money supply had this effect depended on whether it was engineered through an open market operation or whether it was dropped from the sky, like manna from heaven. While Metzler's version of the

_____

[5] _Journal of Political Economy_, 1951

classical model was somewhat different from the one dealt with here, his basic point can be established from an analysis of our classical model.

We assume that π equals zero, so that we can ignore perceived real capital gains on the public's net holdings of financial assets as a component of perceived disposable income. Then the Pigouvian consumption function (4a) becomes

$$C = C(Y - T - \delta K, \; r-\pi \;, \; \frac{M+B}{p} + K),$$

$$C_1 > 0, \; C_2 \leq 0, \; C_3 > 0.$$

In a classical model with such a consumption function, the equilibrium interest rate and price level are determined simultaneously by the aggregate demand--aggregate supply equality and the equilibrium condition for real money balances:

$$C(Y - T - \delta K, \; r-\pi \;, \; \frac{M+B}{p} + K) + I(q(N,K,r-\pi \;, \; \delta)-1)$$

$$+ G + \delta K = Y$$

$$\frac{M}{p} = M(r, \; Y)$$

Taking the total differentials of these two equations and setting $dY = dK = dN = d\delta = 0$, yields

$$(C_2 + I'q_{r-\pi} \;)dr - (C_2 + I'q_{r-\pi} \;)d\pi \; + C_3 \; (\frac{dM+dB}{p})$$

$$- C_3 \; \frac{M+B}{p} \; \frac{dp}{p} + dG - C_1 \; dT = 0$$

$$\frac{dp}{p} = \frac{dM}{M} - m_r \; \frac{p}{M} \; dr \; .$$

Substituting the second expression into the first and rearranging yields

(16) $\qquad (C_2 + I'q_{r-\pi} + C_3 \; \frac{M+B}{M} \; m_r) \; dr - (C_2 + I'q_{r-\pi} \;) \; d\pi$

$$+ C_3 \; (\frac{dM+dB}{p}) - C_3 \; \frac{M+B}{p} \; \frac{dM}{M} + dG - C_1 \; dT = 0.$$

an equation expressing dr in terms of the diffentials of the parameters π, M, B, T, and G. Now consider the effects on the interest rate of a change in the money supply engineered through an open market purchase of bonds, so that dM + dB = 0. Then dr/dM is given by

$$\frac{dr}{dM}\bigg|_{dM=-dB} = \frac{(C_3 \frac{M+B}{p} \cdot \frac{1}{M})}{(C_2 + I'q_{r-\pi} + C_3 \frac{M+B}{M} m_r)}$$

So long as M+B > 0, the above derivative is negative. An increase in the money supply accomplished through open market operations raises the price level, thereby reducing the real wealth of the public's net holdings of financial assets, reducing real wealth, stimulating saving, and lowering the rate of interest. Such a change in the money supply consequently has effects on the real variables of the economy.

On the other hand, consider a change in the money supply accomplished not through open market exchanges, but by simply engaging in a once-and-for-all give-away of money and bonds, the give-away being effected in such a fashion that the proportion of nominal money to nominal bonds remains unaltered. The increment in money and bonds can be thought of as once-and-for-all "bonus" that is dispensed by the government. Thus, we no longer retain the constraint on open-market operations, instead substituting the condition that

$$d \left(\frac{M+B}{M}\right) = 0,$$

which is simply the requirement that the gift of bonds and money is not to alter the proportion of money to bonds. Writing out the above differential, we have

$$\frac{dM + dB}{M} - \frac{dM}{M} \left(\frac{M+B}{M}\right) = 0$$

Multiplying both sides of this equation by M/p yields

$$\frac{dM + dB}{p} - \frac{dM}{M} \left(\frac{M+B}{p}\right) = 0,$$

an equality that holds for gift operations satisfying our stipulation that the relative quantities of bonds and money must not be altered by the gift. Substituting this equality into (13) yields

$$(C_2 + I'q_{r-\pi} + C_3 \frac{M+B}{p} m_r) \, dr - (C_1 - I') \, d\pi$$

$$+ \, dG - C_1 dT = 0$$

an equation that determines the differential of the interest rate and that does not involve dM. Thus, a change in the money supply accomplished by an operation that does not alter the ratio of money to bonds has no effect on the interest rate. The reason is that such a once-and-for-all increase in money and bonds has no effect on real wealth.

The conclusion of this analysis is that in a classical system incorporating a Pigouvian consumption function, the effects of a change in the nominal money supply depend on the nature of accompanying changes in the nominal stock of bonds. Only in the special case in which the ratio of money to bonds is constant does a change in the money supply produce no effect on the interest rate.

Keynesian Economics and Walras' Law

Walras' Law states that where excess demand functions have been constructed so that they obey the pertinent budget constraints, the dollar sum of all excess demands must be zero. This means that an excess supply in one market must be balanced by offsetting excess demands elsewhere.

It is frequently asserted that an important aspect of the
Keynesian model is that it violates Walras Law: there is an excess
supply of labor that is nowhere balanced by an offsetting excess demand.[6/]
Since Walras' Law is frequently misinterpreted in analyses of Keynesian
models, the subject bears brief attention.

It should just be noted that households face two sorts of con-
straints, the satisfaction of each of which will give rise to a form of
Walras' Law. First, households face the (stock) balance sheet constraint

$$\frac{B^D + V^D + M^D}{p} = \frac{B + V + M}{p} \quad (= W) \, ,$$

which says that they must allocate their entire wealth, but no more,
among bonds, equities, and money. Satisfaction of this constraint gave
rise to the cross-equation restrictions on the slopes of the asset
demand schedules $b_r = -m_r$, $b_Y = -m_Y$, $b_w + m_w = 1$. Furthermore, note that
the constraint can be rearranged to read

$$\frac{M^D}{p} - \frac{M}{p} = \frac{B - B^D}{p} + \frac{V - V^D}{p} \quad ,$$

which states that the excess demand for money equals the sum of the excess
supplies for bonds and equities. This is the form of Walras' Law for
stocks of paper assets in our model.

In addition to the balance sheet constraint, households face
the flow budget constraint

$$C + S = Y_D,$$

which says that they can allocate their disposable income between con-
sumption and saving. Let us ignore the term $\frac{M+B}{p} \pi$, say by assuming that π

_____

[6/]For example, see Robert Clower, "The Keynesian Counter-
revolution."

is zero. Stockholders receive as real dividends $Y - w/p \, N - \delta K$, where Y and N are actual quantities of output and employment, respectively. Here Y and N are determined via the aggregate supply curve and labor demand schedules, respectively, as functions of w/p. At a real wage of w/p, workers would like to supply an amount of labor $N^S(w/p)$. Consequently, as a function of w/p, "desired" disposable income is

$$Y_D^* = Y - \frac{w}{p} N - \delta K + \frac{w}{p} N^S(\frac{w}{p}) - T$$

$$= Y - \frac{w}{p} (N - N^S(\frac{w}{p})) - \delta K - T \, ,$$

where by "desired" we mean the amount of income consistent with firms and households being able to transact at whatever quantities they wish at the going real wage w/p. Suppose we equate Y_D^* to $C+S$, a step <u>not</u> taken the the Keynesian model. This gives

$$C = Y - \frac{w}{p} (N - N^S(\frac{w}{p})) - \delta K - T - S.$$

Aggregate demand Y_A is

$$Y_A = C + \dot{K} + G + \delta K = Y - \frac{w}{p} (N - N^S(\frac{w}{p})) - T - S + G + \dot{K}$$

which by virtue of the government budget constraint $(p(G - T) = \dot{M} + \dot{B})$ implies

$$(17) \qquad Y_A - Y = \frac{w}{p} [N^S(\frac{w}{p}) - N] + [\dot{K} + \frac{\dot{M}}{p} + \frac{\dot{B}}{p} - S],$$

which is Walras' Law for flows. It states that the excess aggregate demand for goods $(Y_A - Y)$ equals the sum of the excess supply of labor weighted by the real wage and the excess of the actual real rate of accumulation of assets $\dot{K} + \dot{M}/p + \dot{B}/p$ over the desired rate of accumulation S, i.e., the excess flow supply of real assets.

Such a Law does _not_ hold in the Keynesian model. In the Keynesian model, in equilibrium $N^S(w/p) > N$ while at the same time $Y_A = Y$ and $S = \dot{K} + (\dot{M} + \dot{B})/p$. The reason for this apparent violation of Walras' Law is that the Keynesian model does not impose a flow budget constraint of the form

$$C + S = Y_D^*,$$

the form needed to derive the above form of Walras' Law. Instead of constraining $C + S$ to equal _desired_ disposable income Y_D^*, the Keynesian model constrains $C + S$ to equal _actual_ disposable income Y_D, where actual disposable income is defined by taking the labor component of income to be $(w/p)N$ where N is the amount of actual employment, given by the labor demand schedule. Dividend income continues to be $Y - (w/p)N - \delta K$, so that actual disposable income is

$$Y_D = Y - \frac{w}{p} N - \delta K + \frac{w}{p} N - T = Y - \delta K - T.$$

Equating $C + S$ to Y_D gives

$$C = Y - \delta K - T - S,$$

so that aggregate demand must satisfy

$$Y_A = C + \dot{K} + G + \delta K = Y - T + G - S + \dot{K},$$

which by virtue of the government budget constraint implies

$$(18) \qquad Y_A - Y = [\dot{K} + \frac{\dot{M}}{p} + \frac{\dot{B}}{p} - S].$$

This states that excess aggregate demand for goods equals the excess flow supply of assets, a form of Walras' Law that deletes the labor market. The labor market's deletion is effected by using actual labor income and not desired (supplied) labor times the real wage in defining disposable income. Equation (18) is the form of Walras' Law implicit in

the Keynesian model. It is what lies behind the equivalence of "C + I + G + δK = Y" and "S = $\dot{K}$ + $\dot{M}$/p + $\dot{B}$/p" as descriptions of flow equilibrium in the Keynesian model.

According to some economists, the fact that for the Keynesian model, Walras' Law in the sense of equation (17) does not hold is important. The argument is that because the pertinent concept of labor income is less than workers desire it to be and this constrains workers' flow demands, excess supply in the labor market is prevented from generating offsetting, and presumably expansionary, excess demands in the flow market for goods or paper assets. But clearly it remains the rigidity of the money wage that causes the system to settle at an equilibrium with unemployment. Downward movements in the money wage rate cause outward shifts in the aggregate supply curve in the (p,Y) plane, causing movements downward along the demand schedule in the (p,Y) plane (the Keynes effect), thus causing output and employment to rise.

Notes on Tobin's "Dynamic Aggregative Model"

One of the most important features of the "classical" and "Keynesian" models presented above is that they incorporate an investment demand schedule, a schedule that owes its existence to firms' inability to hold the amount of capital they desire at any moment in time. In those models, firms were not permitted to adjust their capital stocks instantaneously, for it was assumed that there was no market in existing stocks of capital. Thus, a firm faced with a discrepancy between the marginal productivity of capital and the marginal cost of capital could respond only by investing or disinvesting at some finite rate per unit time.

In these notes, we investigate the model that emerges when we replace the "Keynesian" flow investment demand function with the assumption that there is a perfect market in existing capital, one in which firms can purchase or sell all the capital they want at a point in time. A major reason for performing this exercise is to highlight the important role played by the investment demand function in the standard "Keynesian" model. Another rationale for undertaking the exercise is that it forces us to face some difficult questions about the adequacy of the standard one-sector "Keynesian" model as a vehicle for representing the major doctrines in Keynes's General Theory.

The Firm's Optimization Problem

We now assume that the typical firm can purchase or sell all the capital it wants at any point in time at the price of the one good in the model p. Except where noted, all other characteristics of firms, individuals, and the government are as described in the previous notes. Thus among other things we continue to assume that firms issue no bonds, although this assumption could easily be relaxed.

The object of the firm is to maximize its present value, which is simply the discounted value of its net cash flow. The ith firm's net cash flow at time t is

$$p(t) \, Y_i \, (K_i(t), \, N_i(t)) - w(t) \, N_i(t) - p(t) \, (\dot{K}_i(t) + \delta \, K_i(t))$$

where i subscripts denote variables corresponding to the ith firm. The first term $p(t)\ Y_i$ is the firm's revenue, the second term $w(t)\ N_i$ is the firm's payroll, and the third term $p(t)\ (\dot{K}_i + \delta K_i)$ is the firm's current expenditures on capital goods. It pays out $p(t)\ \delta K_i$ to maintain its capital stock intact and $p(t)\ \dot{K}_i$ to add to its capital stock at the rate $\dot{K}_i$ per unit time.

We assume that bonds and equities are perfect substitutes from the viewpoint of wealth-holders, which implies that the nominal interest rate on bonds is the appropriate rate for discounting the firm's cash flow. Thus, the firm's present value is

(1) $V(K_i,\ N_i,\ \dot{K}_i,\ t) =$

$$\int_0^\infty e^{-rt}\ [p(t)\ Y_i\ (K_i(t),\ N_i(t)) - w(t)\ N_i(t) - p(t)\ (\dot{K}_i(t) + \delta K(t))]\ dt\ .$$

Following our practice with the previous models, we assume that the firm expects the price $p(t)$ and the wage $w(t)$ to follow the paths

$$p(t) = p\ e^{\pi t}$$
$$w(t) = w\ e^{\pi t}$$

where π is the anticipated rate of inflation. Then (1) becomes

(1') $V(K_i,\ N_i,\ \dot{K}_i,\ t) =$

$$\int_0^\infty e^{-(r-\pi)t}\ [p\ Y_i(K_i,\ N_i) - w\ N_i - p\ (\dot{K}_i + \delta K_i)]\ dt\ .$$

where K_i and N_i are understood as functions of time. Expression (1') is of the form

$$g(y,\ X,\ \dot{X},\ t) = \int_0^\infty f(g,\ X,\ \dot{X},\ t)\ dt$$

which obtains an extremum when the following "Euler equations" are satisfied:

$$\frac{\partial f}{\partial y} = 0$$

$$\frac{\partial f}{\partial X} - \frac{d}{dt}\ \frac{\partial f}{\partial \dot{X}} = 0\ .$$

For expression (1') to obtain a maximum we thus require

$$e^{-(r-\pi)t} \left(p \frac{\partial Y_i}{\partial N_i} - w\right) = 0$$

and

$$e^{-(r-\pi)t} \left(p \frac{\partial Y_i}{\partial K_i} - p(r + \delta - \pi)\right) = 0 .$$

These conditions simply state that the marginal product of each input must be equated to the appropriate real rental at each point in time:

(2) $\dfrac{\partial Y_i}{\partial N_i} = \dfrac{w}{p}$

(3) $\dfrac{\partial Y_i}{\partial K_i} = (r + \delta - \pi) .$

Expression (3) implies that $(r + \delta - \pi)$ ought to be regarded as the real cost of capital, just as w/p is the real wage of labor. Moreover, this is the cost of capital that ought to be used in defining the firm's profits. For then the firm's profits are

(4) $\Sigma_i = p(t) Y_i(K_i, N_i) - w(t) N_i - (r + \delta - \pi) p(t) K(t).$

Notice that maximizing profits at each moment in time is then equivalent to maximizing the firm's present value, that is it leads to the same marginal conditions, (2) and (3). The equivalence of the two approaches - profit maximization and present value maximization - is a consequence of using the appropriate cost of capital in calculating the firm's profits.

Next, notice that since w, p, and r are constants, (2) and (3) imply that present value is maximized where K_i, N_i, and Y_i are constant over time, so that $\dot{K}_i$ is zero for all time. Then (1') becomes

$$V = [p Y_i(K_i, N_i) - w N_i - \delta p K_i] \int_o^\infty e^{-(r-\pi)t} dt$$

$$= \frac{p \ Y_i \ (K_i, \ N_i) - w_i \ N_i - \delta \ p \ K_i}{r - \pi}$$

$$= \frac{p \ Y_i(K_i, \ N_i) - w_i \ N_i - (r + \delta - \pi) \ p \ K_i}{r - \pi} + \frac{r - \pi}{r - \pi} \ p \ K_i$$

$$= \frac{p \ [Y_i \ (K_i, \ N_i) - Y_N \ N_i - Y_K \cdot K_i]}{r - \pi} + p \ K_i \ .$$

But since Y_i is linearly homogeneous in N_i and K_i, the expression in brackets is zero, so that the value of the firm is the value of its capital stock:

$$V = p \ K_i.$$

Since the firm has no bonds outstanding, V equals the nominal value of the owners' equity in the firm. The nominal yield on equity, ρ, equals net cash flow (earnings) per value of equity plus the anticipated capital gain on equity:

$$\rho = \frac{p \ Y_i - w \ N_i - p \ \delta \ K_i}{p \ K_i} + \frac{\dot{V}}{pK_i}$$

$$= \frac{p \ Y_i - wN_i - (r + \delta - \pi) \ p \ K_i}{pK_i} + (r - \pi) + \frac{\dot{p} \ K_i + \dot{K}_i p}{pK_i} \ .$$

Setting $\dot{K}_i$ equal to zero, as we have above, and replacing $\dot{P}/P$ with the relevant anticipated magnitude π yields

$$\rho = r - \pi + \pi$$

$$= r$$

which confirms that the nominal yield on equities equals the bond rate.

We now assume that all firms in the economy maximize their present value, so that (3) holds at each moment for each firm. As a consequence, we assume that the marginal productivity condition for capital also holds for the entire economy, so that

(5) $Y_K = r + \delta - \pi$

at each instant. This equality is held to be enforced by firms' trading existing capital at each moment in time.

A version of Tobin's "Dynamic Aggregative Model"* is derived by replacing our "Keynesian" investment function with the marginal equality (5). The rest of the model we will deal with is identical with our "Keynesian" model. Thus, our model consists of the following six equations:

I $Y = Y(K, N)$

II $\frac{w}{p} = Y_N$

III $(r + \delta - \pi) = Y_K$

IV $\frac{M}{p} = m(r, Y)$

V $C = c(Y - \delta K - T - \frac{M+B}{p} \pi, r - \pi)$

VI $C + I + G + \delta K = Y$.

The variables of the model are Y, N, C, I, r, and p. The parameters are w, K, π, δ, G, T, and M. The constraints on the government and the behavior of holders of paper assets are all assumed to be the same as they are in the Keynesian model.

Studying the behavior of the model is facilitated by writing down the total differentials of equations I through VI:

(i) $dY = Y_N \, dN + Y_K \, dK$

(ii) $\frac{dw}{p} - \frac{w}{p} \frac{dp}{p} = Y_{NN} \, dN + Y_{NK} \, dK$

(iii) $dr - d\pi = Y_{KN} \, dN + Y_{KK} \, dK$

(iv) $\frac{dM}{p} - \frac{M}{p} \frac{dp}{p} = m_r \, dr + m_Y \, dY$

(v) $dC = c_1 \, dY - c_1 \, \delta \, dK - c_1 \, dT - c_1 \frac{M+B}{p} \, d\pi + c_1\pi \frac{M+B}{p} \frac{dp}{p} + C_2 \, dr - c_2 \, d\pi$

(vi) $dC + dI = dG + \delta \, dK = dY$.

The system can be displayed compactly in the following matrix equation.

---

*Journal of Political Economy, 1955.

$$
\begin{bmatrix}
1 & -Y_N & 0 & 0 & 0 & 0 \\
0 & -Y_{NN} & -\dfrac{w}{p^2} & 0 & 0 & 0 \\
0 & -Y_{KN} & 0 & 1 & 0 & 0 \\
m_y & 0 & \dfrac{M}{p^2} & m_r & 0 & 0 \\
-c_1 & 0 & -c_{1\pi}\dfrac{M+B}{p^2} & -c_2 & 1 & 0 \\
1 & 0 & 0 & 0 & -1 & -1
\end{bmatrix}
\begin{bmatrix}
dY \\ dN \\ dP \\ dr \\ dC \\ dI
\end{bmatrix}
=
\begin{bmatrix}
Y_K\,dK \\
-\dfrac{dw}{p} + Y_{NK}\,dK \\
d\pi + Y_{KK}\,dK \\
\dfrac{dM}{p} \\
-c_2\,d\pi - c_1\delta dK - c_1 dT \\
-c_1\dfrac{M+B}{p}\,d\pi \\
dG + \delta\,dK
\end{bmatrix}
$$

Inspection of the matrix on the left side of the equation reveals that the system is block recursive: there is a four-by-two matrix of zeroes in the upper right portion of that matrix, which implies that the first four equations of the model form an independent subset that determines Y, N, p, and r. Given those variables the fifth equation determines consumption (note the five-by-one matrix of zeroes in the upper right portion of our matrix, which means that the first five equations form an independent subset in Y, N, p, r, and C). Finally, the last equation determines investment.

To solve the system, our strategy will be to collapse the four equations (i) through (iv) into two equations in two unknowns. Tobin himself collapsed the system into two equations in p and N. For ease of comparison with our earlier models, we will collapse it into two equations in r and Y. The first equation is the LM curve, which is identical with the LM curve of the "Keynesian" model. Its total differential, derived by substituting (i) and (ii) into (iv), is

$$
(6) \quad dr = \frac{1}{m_r}\left(\frac{dM}{p} - \frac{M}{p}\frac{dw}{w} + \left(\frac{Y_{NN}}{Y_N^2}\frac{m}{p} - m_Y\right)dY\right).
$$

The LM curve depicts the combinations of the interest rate and output that equate the demand and supply of real balances.

The second schedule depicts the combinations of output and the interest rate

that satisfy the marginal productivity condition for capital. It is thus in the nature of a capital market equilibrium curve, which depicts the combinations of r and Y that make firms content to hold the existing stock of capital. The curve is derived from equations III and I. Its differential is derived by substituting (i) into (iii):

$$dr = d_\pi + Y_{KN} \, dN$$

$$(7) \qquad dr = d_\pi + \frac{Y_{KN}}{Y_N} \, dY$$

Since Y_{KN} is positive, the capital equilibrium curve is upward sloping in the r-Y plane. We assume, however, that its slope is less steep than is that of the LM curve, which will be seen to be a condition required for stability.*

The model is summarized graphically in figure 1, which depicts an LM curve intersecting an upward sloping "capital-equilibrium" curve, which we have labeled "KE." The intersection of those two curves determines an r and Y that solve our model.
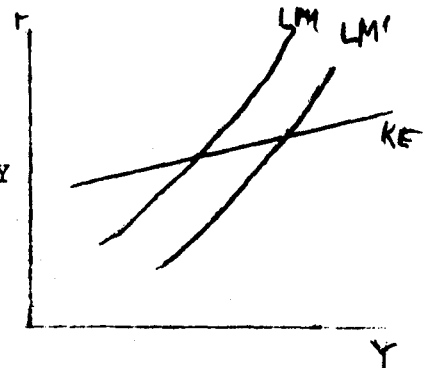

Figure 1

To illustrate the workings of the model, we will consider the response of output and the interest rate to changes in the money supply and the wage rate, which influence the position of the LM curve, and changes in the anticipated rate of inflation, which shift the KE curve. An increase in the money supply will shift the LM curve toward the right, just as in our Keynesian model. At the old equilibrium level of output, portfolio balance would require a decrease in the interest rate, as wealth holders would be content to hold more money and fewer bonds only at a lower interest rate. But such a decrease in the interest rate would disturb the equality between the net marginal product of capital and the real rate of interest, so that firms would want

---

*Tobin's article analyzes how the model might behave if the stability condition were not to be fulfilled in a certain region.

to acquire more capital. They would attempt to do so, but since the quantity of capital in existence is fixed, the result would be a rise in the pri-- of existing capital, which in our model is the price of the one good in the model. The rise in the price level would lower real wages, inducing employers to hire more employees and to produce more output. That would in turn increase the marginal productivity of capital, thus increasing the interest rate at which firms would be content to hold just the existing capital stock. However, the increase in the price level and the rate of output also increases the interest rate needed to achieve portfolio balance. If the rate at which the portfolio-balancing interest rate rises with output exceeds the rate at which the capital-balancing interest rate rises with output, as will be true if the LM curve is steeper than the KE curve, the system will be driven to equilibrium at a higher output and interest rate than characterized the initial equilibrium. The situation is depicted in figure 1.

A fall in the money wage works in very much the same way, since it also produces a shift of the LM curve to the right. The fall in wages induces a fall in prices, which reduces the interest rate required for portfolio balance at the initial income level. That in turn disturbs capital market equilibrium, causing firms to bid for capital, raising the price level, stimulating output, and finally reversing the tendency for the interest rate to fall through the effects of rising output on the demand for money and the marginal product of capital. Thus a fall in the money wage causes both output and the interest rate to rise.

The effect of an increase in the anticipated rate of inflation π is depicted in Figure 2. The effect is to shift the KE curve upward by the amount of the increase in π. Such a change disturbs the marginal productivity equality for capital at the old output level, causing firms to bid for more capital, which in turn causes the
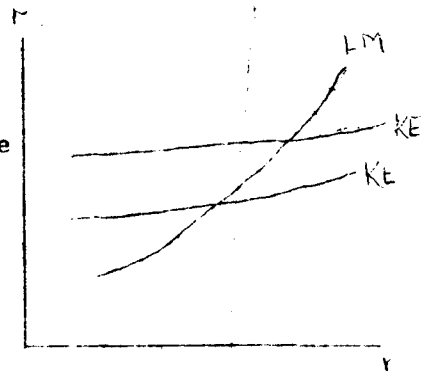

Figure 2

price level and output to rise. But that raises the marginal product of capital,
which means that the interest rate must eventually rise by more that the increase
in π in order to restore equilibrium.

The model can be solved analytically by equating (6) and (7), and then solving
for dY. This yields the differential of the reduced form for Y

$$(8) \quad [\frac{1}{m_r} (\frac{Y_{NN}}{Y_N^2} \frac{M}{P} - m_Y) - \frac{Y_{KN}}{Y_N}] \, dY = d_\pi - \frac{1}{m_r} \frac{dM}{P} + \frac{1}{m_r} \frac{M}{P} \frac{dw}{w}.$$

If the stability condition is satisfied, the coefficient on dY is positive. This
guarantees that Y varies directly with π and M and inversely with w.

As expression (8) reveals, fiscal policy has no impact on the level of output
at a point in time. Neither dG nor dT appears in (8). In addition, the consumption
function, and by implication also the Keynesian multiplier play no role in deter-
mining the rate of output. On the other hand, monetary policy is seen to play
an important role in determining output, one that depends not at all on its influ-
encing the desired flow of expenditures for consumption or investment, the only way
money acquired potency in our one-sector "Keynesian" model. Instead, money works
by disturbing equilibrium in the market for existing capital. That causes firms
to try to pruchase or sell capital, which in turn causes the price changes that are
the prerequisite for output changes.

Once output and the interest rate are determined by (6) and (7), employment
and the price level are easily calculated from equations (I) and (IV). Then
consumption is determined by equation V. Finally, investment is determined by
equation VI. Note that fiscal policy is important as a determinant of K, and hence
the economy's rate of growth overtime, even though it is incapable of influencing
the levels of output, employment, price, and the interest rate at a point in time.

Tobin's Model as A Special Case of A Two-Sector Model

How is it that an increase in government purchases or a decrease in taxes

fails to affect aggregate employment and output at a point in time? The best

way to explain the mechanism that denies any potency to such changes in fiscal

policy is to consider Tobin's model as a special case of a two-sector model in

which there are different consumption and investment goods. Production of the

investment good, Y_I, satisfies the linearly homogeneous production function,

$$Y_I = I(K_I, N_I), \quad I_K, I_N, I_{KN} > 0, \quad I_{NN}, I_{KK} < 0,$$

where K_I is capital employed in the investment goods industry, and N_I is

employment in the investment good industry. Output of consumption goods, Y_C,

satisfies the linearly homogenous production function

$$Y_C = C(K_C, N_C), \quad C_K, C_N, C_{KN}, > 0; \quad C_{NN}, C_{KK} < 0$$

where K_C and N_C are capital and employment in the consumption goods industry,

respectively.

The price of capital goods is $p_1 p$, while the price of consumption goods is p.

Firms in each industry can hire all the labor they want at the money wage w, can

purchase all the capital they want at the price $p_1 p$, and can sell all the output

they want at their respective product price. There are thus perfect markets in

capital and employment. Firms in each industry maximize profits with respect

to variations in both employment and capital. Profits in the capital goods industry

are

$$\Sigma_I = P_1 P I(K_I, N_I) - wN_i - (r + \delta - \pi) P_1 P K_I,$$

which implies the marginal conditions

$$I_K = (r + \delta - \pi)$$

$$I_N = \frac{1}{P_1} \frac{w}{P}$$

Profits in the consumer goods industry are

$$\Sigma_c = PC(K_c, N_c) - wN_c - (r + \delta - \pi)P_1 PK_c,$$

which implies the marginal conditions

$$C_K = P_1(r + \delta - \pi)$$

$$C_N = \frac{w}{P}$$

In addition to having the above marginal conditions hold, we require that the market in existing capital clear, i.e., $K_I + K_c = K$, where K is the amount of capital in existence.

Our assumptions about households' porfolio and saving behavior remain as before, where $Y = P_1 Y_I + Y_c$, GNP measured in consumption goods, appears in the portfolio equilibrium schedule and the consumption function.

Collecting the equations of our model, we have

9. $Y = P_1 Y_I + Y_c$

10. $Y_I = I(K_I, N_I)$

11. $Y_c = C(K_c, N_c)$

12. $I_K = (r + \delta - \pi)$

13. $I_N = \frac{1}{P_1}\frac{w}{P}$

14. $C_K = P_1(r + \delta - \pi)$

15. $C_N = \frac{w}{P}$

16. $K_I + K_c = K$

17. $Y_c = c(Y - \delta K - T - \frac{M+B}{P}\pi; \; r - \pi) + G$

18. $\frac{M}{P} = m(r, Y)$

which is a system of ten equations in the ten endogenous variables Y, Y_I, Y_c, N_I,

N_c, K_I, K_c, r, P_1, and P. The exogenous variables are w, L, T, G, K, B, and π.

Notice that we have assumed that the government purchases consumption goods at

the rate G.

We now consider the special case in which the production functions $I(\ )$

and $C(\ )$ are identical, so that

$$I(K,N) = C(K,N) = F(K,N)$$

Using (12) and (13), we have

$$(19) \quad \frac{F_K}{F_N} = \frac{I_K}{I_N} = \frac{P_1(r + \delta - \pi)}{\frac{w}{P}}$$

Using (14) and (15), we have

$$(20) \quad \frac{F_K}{F_N} = \frac{C_K}{C_N} = \frac{P_1(r + \delta - \pi)}{\frac{w}{P}}$$

By virtue of the linear homogeneity of $F(K,N)$, we have

$$F_K(K,N) = F_K(\frac{K}{N}, 1)$$

and

$$F_N(K,N) = F_K(\frac{K}{N}, 1),$$

so that F_K/F_N is a function of K/N with derivative

$$(21) \quad \frac{d}{d\frac{K}{N}} \left(\frac{F_K}{F_N}\right) = \frac{F_N F_{KK} - F_K F_{NK}}{F_N^2} < 0$$

An increase in the capital-labor ratio thus lowers F_K/F_N. Inequality (21)

implies that we can invert the above relation to obtain the capital-labor ratio

as a function of F_K/F_N,

$$(22) \quad \frac{K}{N} = \xi\left(\frac{F_K}{F_N}\right), \quad \xi' = \left(\frac{F_N^2}{F_N F_{KK} - F_K F_{NK}}\right) < 0.$$

Now by virtue of equations (19) and (20), we have

$$\frac{F_K(K_I, N_I)}{F_N(K_I, N_I)} = \frac{F_K(K_c, N_c)}{F_N(K_c, N_c)} = \frac{P_1(r + \delta - \pi)}{\frac{w}{P}}$$

In combination with equation (22), this implies

$$\frac{K_I}{N_I} = \frac{K_c}{N_c} = \xi\left(\frac{P_1(r + \delta - \pi)}{\frac{w}{P}}\right)$$

so that the capital-labor ratios are identical in the two industries. Taking the ratio of (15) to (13), we have

$$\frac{C_N}{I_N} = P_1$$

But Since $K_I/N_I = K_c/N_c$, we have $C_N = I_N$, so that the above equality implies that $P_1 = 1$.

Now notice that by virtue of Euler's theorem, we have

$$Y = P_1 Y_I + Y_c$$

$$= P_1 (F_K K_I + F_N N_I) + (F_K K_c + F_N N_c)$$

$$= (P_1 F_K K_I + F_K K_c) + (P_1 F_N N_I + F_N N_c)$$

$$= F_K (\frac{K}{N}, 1)K + F_N(\frac{K}{N}, 1)N$$

(23) $Y = F(K, N)$

where $N = N_c + N_I$. Moreover, notice that since

$$\frac{K_I}{N_I} = \frac{K_c}{N_c} = \frac{K_I + K_c}{N_I + N_c} ,$$

we have

(24) $F_K(K_c, N_c) = F_K(K_I, N_I) = F_K(K, N)$

(25) $F_N(K_c, N_c) = F_N(K_I, N_I) = F_N(K, N)$

Notice that since $Y = F(K,N)$ and $P_I = 1$, we have

$$Y_I + Y_c = F(K, N),$$

which is the "transformation curve" between Y_I and Y_c, a straight line with slope -1 and intercept $F(K, N)$. This curve gives the maximum Y_I (or Y_c) that the economy can produce for given values of K, N, and Y_c (or Y_I). To find the transformation curve, we can maximize $Y_I = F(K_I, N_I)$ subject to $\bar{Y}_c = F(K - K_I, N - N_I)$, which is accomplished by unconstrained maximization of

$$J = F(K_I, N_I) + \lambda(\bar{Y}_c - F(K - K_I, N - N_I)).$$

The first-order conditions are

$$\frac{\delta J}{\delta K_I} = F_{K_I} + \lambda F_{K_c} = 0$$

$$\frac{\delta J}{\delta N_I} = F_{N_I} + \lambda F_{N_c} = 0$$

$$Y_c = F(K - K_I, N - N_I)$$

Dividing the first marginal condition by the second gives

$$\frac{F_{K_I}}{F_{N_I}} = \frac{F_{K_c}}{F_{N_c}}$$

which by virtue of the linear homogeneity of $F(\ )$ implies

$$\frac{K_I}{N_I} = \frac{K_c}{N_c} = \frac{K}{N},$$

which, again by the linear homogeneity of F, implies that

$$F_{K_I}(K_I, N_I) = F_{K_c}(K_c, N_c) = F_K(K, N)$$

$$F_{N_I}(K_I, N_I) = F_{N_c}(K_c, N_c) = F_N(K, N).$$

Euler's theorem then implies that

$$Y_I + Y_c = F_{K_c}(K_c, N_c)K_c + F_{N_c}(K_c, N_c)N_c$$

$$+ F_{K_I}\left(K_I, N_I\right)K_I + F_{N_I}\left(K_I, N_I\right)N_I$$

$$= F_K(K, N)K + F_N(K, N)N$$

$$= F(K, N)$$

Since the transformation curve has a slope of -1, it follows that P_1 must equal unity so long as the marginal equalities for the factors of production are in force and both Y_I and Y_c are positive.

As a result of (23), (24), (25), and the fact that $P_1 = 1$, we have that for the special case in which the production functions in the two industries are identical, the model consisting of equations (9) through (18) can be collapsed to the following system:

$$Y = F(K, N)$$

$$\frac{w}{P} = F_N(K, N)$$

$$Y = Y_c + Y_I$$

$$Y_c = c(Y - T - K - \frac{M + B}{P} \pi, r - \pi) + G$$

$$\frac{M}{P} = m(r, Y) \qquad ,$$

which consists of six equations in the endogenous variables Y, N, r, p, Y_c, and Y_I and the exogenous variables, M, B, π, G, T, and K. This system is identical with Tobin's Dynamic Aggregative Model.

We are now in a position to analyze the forces that prevent changes in G from having any effects on the aggregate levels of output and employment. Suppose G increases by dG. The model predicts that Y_c will rise by dG while Y_I will fall by dG, so that holdings of capital and employment in the two industries will change according to

$$dN_c = \frac{N_c}{Y_c}dg = -\frac{N_I}{Y_I}dG = -dN_I$$

$$dK_c = \frac{K_c}{Y_c}dg = - \frac{K_I}{Y_I}dg = -dK_I$$

These changes occur at fixed values of P_1, P, w, and r, so that labor-output and capital-outputs remain unchanged. Notice that $N_c/Y_c = N_I/Y_I$, and $K_c/Y_c = K_I/Y_I$ by virtue of the assumption of identical production functions in the two industries.

What market forces prompt producers of capital goods to release just the amounts of capital and labor that producers of consumer goods need to meet the increased demand for their output? At the initial equilibrium values of Y, Y_I, and Y_c, the increased demand for consumer goods caused by the increase in G produces a slight tendency for P to rise, while P_1P remains constant. The rise in P induces a rise in the profits of consumer goods producers, while profits of investment goods producers remain unchanged:

$$\frac{d\Sigma_I}{dP}\bigg|_{d(P_1P) = 0} = 0$$

$$\frac{d\Sigma_c}{dp}\bigg|_{dP_1P = 0} = Y_c > 0$$

The rate of return on capital in the consumer goods industry thus rises, exceeding that in the investment goods industry. This causes firms in the consumer good industry to bid for capital, which firms in the investment goods industry are willing to sell or rent to them in order to obtain the highest possible rate of return in their capital. How much of a rise in P is required to bring about such changes in the distribution of capital and employment? The slightest upward movement in P (downward movement in P_1) is sufficient to bring about arbitrarily large required changes in the distribution of capital and employment across industries. That is, a drop in P_1 below unity causes the system to move along a transformation curve to the corner where all resources are devoted to

producing consumption goods. Consequently, in equilibrium p and P_1 remain at their initial levels, as do Y and N. For if P_1 were to fall below unity, all resources would be devoted to producing consumption goods, which would surely not be an equilibrium situation provided that the initial value of Y_I exceeds the new level of G.

Saving and Investment, Etc.

We notice that the economy's rate of net investment is determined as a "residual" in equation VI. By substituting the definition of disposable income into equation VI, it can be written in the equivalent form

$$I + G = S + T + \frac{M+B}{p} \pi \ .$$

Assume now that the government's budget is balanced, so that G equals T, and that the anticipated rate of inflation π is zero. Then we have

(26) $\quad I = S \ ,$

which is the familiar statement that "investment equals saving." After the publication of Keynes's General Theory there was a good deal of controversy about the meaning of this equality, which is of course present in the "classical" and "Keynesian" models as well as the model being studied here. In particular, some controversy centered about whether or not the investment rate appearing in (9) ought to be interpreted as "intended" or ex ante investment, and whether the equality should be interpreted as an equilibrium condition or as a statement of necessary relationships between ex post quantities. These old controversies are of some interest in the light of the model discussed in these notes, since the model contains no ex ante investment demand schedule. Consequently here (9) ought not to be interpreted as an equilibrium relationship between ex ante concepts. Instead, (9) states that saving intentions are realized and that these, together with the government's fiscal policies, govern the economy's rate of capital accumulation. In this sense, investment is determined entirely on the "supply side", that is by the flow of resources that consumers and the government are willing not to consume. Notice also that the model contains a device for reconciling firms to whatever rate of investment is determined by this process. The rate of investment determines subsequent levels of the capital stock. The model will generate time paths of prices, the interest rate, and output which guarantee that firms are content to hold just the existing capital stock at each moment in time. Thus, in this model no useful purpose is served by introducing a distinction between intended and realized

investment.*

The model is obviously quite "monetarist" in its predictions. Fiscal policy matters not at all while changes in the money supply produce effects on output at a point in time. Moreover, money acquires this potency not by its effects on desired flows of consumption and investment, as is true in the Keynesian model. Instead, money impringes on the economy by producing effects in the market for existing physical assets. It is from that market that pressures on the price level emerge, and it is changes in the price level that permit output to be affected at a point in time. In view of these properties of the model, it is not surprising that monetarists have on occasion emphasized the important role played by markets in stocks of physical assets in the process by which the effects of money supply changes are transmitted.**

The impotence of fiscal policy in this model may seem peculiar in the light of Keynes's doctrine of the marginal efficiency of capital, which pretty clearly demonstrates that the formal model of Keynes's General Theory contains a perfect market in which firms can always purchase and sell existing physical capital goods. The marginal efficiency of capital was defined as that rate of discount which equates the present value of the stream of returns associated with a given capital good to the purchase price of the good. Keynes asserted that firms would purchase capital

---

*In our view, the model under consideration more faithfully reflects Keynes's views of the determinants of capital accumulation than does our "Keynesian" model. As a consequence, it helps rationalize some of Keynes's views on the saving-equals-investment controversies just discussed.

**For example, see M. Friedman and D. Meiselman, "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958", in Stabilization Policies, Commission on Money Credit, (Prentice Hall) 1963. See especially pp. 217-22.

goods until the marginal efficiency of capital equals the rate of interest . As Keynes pointed out, this is equivalent to assuming that firms maximize their present value.* In terms of the one-sector model we are studying here note that the present cost of an additional unit of capital is simply the current price of the one good, p. The net return associated with an additional unit of capital at time t is $p(t) Y_K - p(t)\delta$. Thus the marginal efficiency of capital ξ is defined as that rate of discount which satisfies the equation

$$p = \int_0^\infty e^{-\xi t} [p(t)Y_K - \delta p(t)] \, dt$$

Assuming that p(t) is expected to follow the path

$$p(t) = pe^{\pi t} \ ,$$

the above equation simplifies to

$$p = (pY_K - p\delta) \int_0^\infty e^{-(\xi-\pi)t} \, dt$$

$$p = \frac{pY_K - p\delta}{\xi - \pi} \ .$$

Thus the marginal efficiency of capital is

$$\xi = Y_K + \pi - \delta \ .$$

Keynes asserted that firms will purchase or sell capital goods until ξ equals r:

$$r = \xi = Y_K + \pi - \delta$$

which can be written

$$Y_K = r + \delta - \pi \ ,$$

---

*Keynes pointed out that his marginal efficiency of capital was identical with Irving Fisher's "rate of return over cost". See The General Theory, chapter 11.

which is equivalent with the marginal condition for capital that springs from present-value maximization.

Now we have seen that a one-sector model that incorporates this marginal productivity condition for capital denies any potency to fiscal policy as a determinant of output and employment at a point in time. Moreover, in such a model, the consumption function plays no role in determining the current rate of output. Yet we know that Keynes believed both that the consumption function and fiscal policy actions were important in determining output and employment in the short run. How can we reconcile those beliefs with a model that, like Keynes's model, includes a perfect market in existing physical capital?

Miscellaneous Notes

1. The "Real Bills" Doctrine

Up to now we have taken the money supply as a parameter, one which is somehow determined by the government. It happens, however, that the government has not always viewed the money supply as a thing it ought to control, even when it has had the tools needed to control it. Instead, it has often been argued that the proper function of the monetary authorities is to set the interest rate at some reasonable level, allowing the money supply to be whatever it must be to ensure that the demand for money at that interest rate is satisfied. Such a rule was actually written into the original Federal Reserve Act that established the Federal Reserve System in the U.S. The rule was known as the "real bills" doctrine. It was alleged that the quantity of money would automatically be properly regulated if the monetary authorities ensured that banks always had enough reserves to meet the demand for loans intended to finance "real" (as opposed to "speculative") investments at an interest rate set "with a view of accommodating commerce and business."

Here we analyze the effect of such a rule in the context of the "classical" model discussed above. The effect of the rule is to make the interest rate a parameter, one determined by the monetary authorities, and the money supply a variable. The monetary authority simply conducts whatever open market operations required to make the interest rate what they want it to be. It offers to buy or sell whatever quantities of government bonds are offered at the announced interest rate.

We consider the version of the classical model discussed in "Notes on the Keynesian Model," which consists of the following seven equations:

(1) $\qquad Y = F(N, K)$

(2) $\qquad \dfrac{w}{p} = F_N(N, K)$

(3) $\qquad N = N^S(\dfrac{w}{p})$

(4) $\qquad C = c(Y-T-\delta K- \dfrac{M+B}{p} \pi, \; r-\pi)$

(5) $\qquad I = I(q(N, K, r-\pi) -1)$

(6) $\qquad C+I+G+\delta K = Y$

(7) $\qquad \dfrac{M}{p} = m(r, Y)$

The endogenous variables are Y, N, p, w, C, I, and M while the exogenous variables are G, T, K, π, and r. The monetary authority is assumed to peg r by permitting the money supply be whatever it must be to equal the demand for money at that interest rate.

The system is block recursive, equations (1), (2), and (3) determining Y, N, and $\dfrac{w}{p}$. Equations (4), (5), and (6) are then three equations that given the previously determined value of Y, can determine the three endogenous variables C, I, and p. Once these three variables have been determined, the last equation determines M, given the previously determined values of p and Y.

The workings of the model can be illustrated with Figure 1. The authority pegs the interest rate at $\bar{r}$. The first three equations determine output at its full employment level, which we denote by Y_F. The next three equations determine a level of aggregate demand which

equals Y_F at the pegged interest rate $\bar{r}$. Since $\bar{r}$ cannot adjust to make

aggregate demand equal aggregate supply, the entire burden falls on the

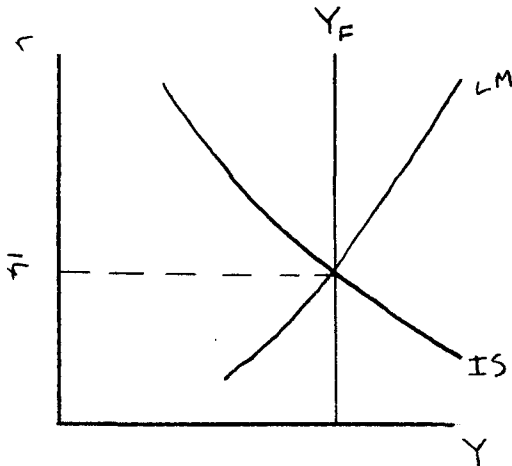price level, which influences consumption through its effect on perceived

real disposable income. The price level adjusts to make the IS curve pass through the intersection of $\bar{r}$ and Y_F. Finally, the portfolio balance equation determines the money supply. To peg the interest rate, the government must issue enough money to guarantee that the LM curve passes through the intersection of Y_F and $\bar{r}$.

Figure 1

To illustrate how the model works, suppose that there is an

increase in government expenditures. Since the interest rate and output

are fixed, the price level must adjust to diminish perceived disposable

income and consumption by just enough to offset the increase in G. In

particular, the change in p and the change in G must obey

$$dG = -C' (\frac{M+B}{p^2} \pi)\, dp$$

If M+B and π are both positive, prices must fall in order to increase

the real value of the government's debt to the public and the anticipated

real capital losses on that debt, thereby decreasing disposable income

and consumption. The fall in p in turn causes a fall in M through the

portfolio balance equation.

Notice that the "stability" of the system in response to a

change in G depends on the sign of $(\frac{M+B}{p})\pi$. For if G rises, at the

initial r, p, and Y, there is an excess demand for goods, which we usually posit leads to rising prices. But if $\frac{(M+B)}{p}\pi$ is positive, a fall in p is required to restore equilibrium, so that the system is driven away from equilibrium. On the other hand, if $\frac{M+B}{p}\pi$ is negative, the system is stable.

Suppose now that $\pi(M+B) = 0$, so that p no longer appears in the consumption function. Then the fourth, fifth, and sixth equations of our system are three equations in only two variables, C and I, since p no longer appears. These three equations in general "overdetermine" the two variables I and C. The investment equation can be viewed as determining I while the consumption function determines C. But in general, the values of C and I so determined will violate the national income equation, where Y has been determined by (1), (2), and (3).

While the model overdetermines C and I, p and M are "underdetermined". Each of these two variables appear only in the seventh equation, which is not capable of determining their values.
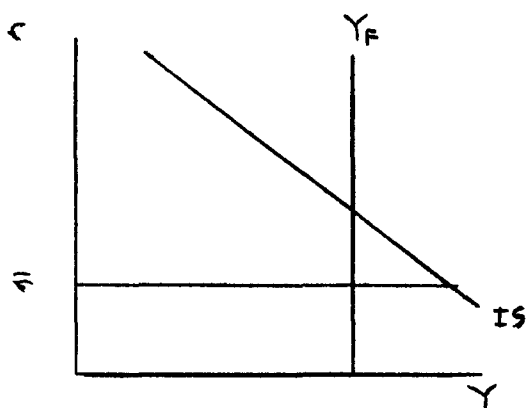
The problem with the model is illustrated by Figure 2. The first three equations determine Y_F, while the second three determine an IS curve. There is nothing to guarantee that the IS curve passes through the point $(Y_F, \bar{r})$. The model posits too many independent relationships between r and Y to enable equilibrium to be determined.

Figure 2

Nor can the problem be avoided by pegging r at the "correct" level, i.e., at the intersection of the Y_f line and the IS curve. It is

true that if r is pegged at that level, the solution for Y from the
labor market and production function agrees with that from the IS curve.
But it remains true that the portfolio balance curve only determines the
ratio $\frac{M}{p}$ and cannot determine the level of either M or p. So even if r
is pegged at the "correct" level, the price level and the money supply
are indeterminate.

This model illustrates the danger of naïve "equation counting"
as a technique supposedly capable of determining whether a model possesses
a unique equilibrium. Our model consists of seven linear equations in
seven variables. The equality between the number of equations and
number of variables does not suffice to guarantee that the equations
have a solution. The problem is in the last four equations, which given
Y, N, and $\frac{w}{p}$, form a system of four equations in the four variables C, I,
p, and M. But the system decomposes in an unfortunate way, the fourth,
fifth, and sixth equations forming an independent subset involving only
I and C while the seventh is an independent equation involving both p
and M.

The standard criticisms made of the "real bills" doctrine are
based on the analysis above. The argument is that any attempt to peg the
interest rate is destabilizing, to say the least. For example, Wicksell
argued that pegging the interest rate too low would set off increases in
the price level and money supply of indefinitely large magnitudes.
The model itself possesses no equilibrium when the monetary author ties
peg the interest rate at such a level, but "stability conditions" do
imply that the price level and money supply will rise, and rise by
indefinitely large amounts, if the monetary authorities pursue such a
policy.

It should be clear that the criticism of the "real bills" doctrine based on the classical analysis described above is formally correct only in a system that contains neither a Pigou effect, nor nonzero perceived real capital gains on the government's debt to the public. This is another illustration of Metzler's point that assigning a role to the Pigou effect fundamentally alters some important aspects of the classical macroeconomic system.

The model discussed above is important in the history of economic thought. Through analysis of such an overdetermined system, Wicksell performed some early "Keynesian" economic analysis.

2. "Inside" and "Outside" Money

Up to now we have assumed that firms have no bonds outstanding, issuing equities to finance all of their investment. We have also assumed that the government does not own any of the liabilities of firms. We now abandon these assumptions in order to provide a framework for understanding the distinctions some economists have made between "inside money" and "outside money."

We will work with a version of Tobin's "Dynamic Aggregative Model." We now assume that firms issue both variable-coupon bonds, whose nominal value is B_f, and equities, whose nominal value is V. Households regard firms' bonds and equities as perfect substitutes when their real yields are equal. They also regard firms' bonds as perfect substitutes for those of the government. These assumptions imply that government bonds, bonds issued by firms, and equities will all bear the same perceived real rate of return: $r - \pi$. The value of owners' equity equals the capitalized value of firms' net cash flow after paying off bond holders.

$$V(t) = \int_t^\infty [p(s)Y(s) - w(s)N(s) - \delta p(s)K(s) - r(s)B_f(s)] \ e^{-r(s-t)}ds$$

Assume that the variables Y, N, K, B_f and r are expected to remain at their current values, while w and p are expected to follow the paths

$$w(s) = w(t)e^{\pi(s-t)}$$

$$p(s) = p(t)e^{\pi(s-t)}$$

Then V(t) becomes

$$V(t) = [pY - wN - \delta pK] \int_t^\infty e^{-(r-\pi)(s-t)} ds - rB_f(t) \int_t^\infty e^{-r(s-t)} ds$$

$$= \frac{pY - wN - \delta pK}{r - \pi} - B_f(t).$$

From the linear homogeneity of Y and the profit-maximization conditions for K and N, this becomes,

$$V(t) = p(t)K(t) - B_f(t).$$

Thus, the value of owners' and creditors' claims on firms, $V(t) + B_f(t)$, equals the market value of firms' physical capital. Firms' consolidated balance sheet is as follows:

Assets	Liabilities
pK	B_f
	V

The household sectors' balance sheet is now:

Assets	Liabilities
$B_f(h)$	
B_G	
V	
M	Net worth

Here $B_f(h)$ denotes the nominal quantity of firms' bonds held by the household sector while B_G denotes the nominal quantity of government bonds, which are all held by households. Households continue to own all equities and all money in the system.

The government issues debt in the form of money and government bonds, the total quantity outstanding being B_G. The government also owns a quantity of firms' bonds denoted by $B_f(G)$. The government's balance sheet is thus:

Assets	Liabilities
$B_f(G)$	M
	B_G
	Net worth

Notice that $B_f = B_f(h) + B_f(G)$. The government conducts open market operations subject to $dM + dB_G = dB_f(G)$.

The economy is assumed to be described by the following "classical" version of Tobin's "Dynamic Aggregative Model":

(1) $\quad Y = Y(K, N)$

(2) $\quad \dfrac{w}{p} = Y_N$

(3) $\quad N = N(\dfrac{w}{p})$

(4) $\quad r = Y_K - \delta$

(5) $\quad \dfrac{M}{p} = m(r, Y)$

(6) $\quad C = C(Y - T - \delta K, r, W) \qquad C_3 > 0$

(7) $\quad C + I + G + \delta K = Y$

We have assumed that π equals zero. We have added a classical labor supply curve, equation (3), to Tobin's model. Households real wealth W is given by

$$W = \frac{M + V + B_G(h) + B_f(h)}{p}$$

Equation (6) is a Pigouvian consumption function, real consumption varying directly with households' real wealth.

As in the classical model, equations (1), (2), and (3) form an independent subset that determines Y, $\frac{w}{p}$, and N. Equation (4) then determines r. Equation (5) determines p. Equations (6) and (7) then determine the division of Y into C and I, and hence influence the economy's rate of growth over time. The determination of Y, N, $\frac{w}{p}$ and r is illustrated in Figure 3:
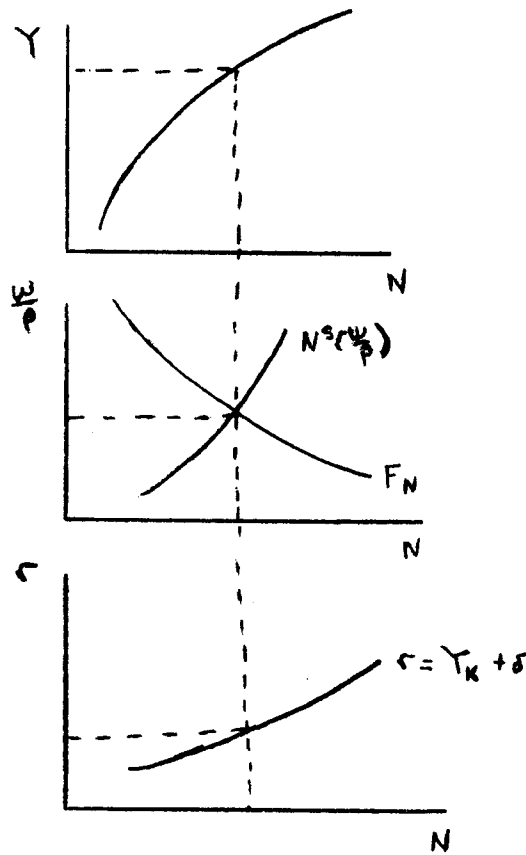

Figure 3

We will now examine the circumstances under which money is "neutral" in this model, having no effects on the real variables of the model. We already know that a point in time, Y, $\frac{w}{p}$, and N are determined by equations (1), (2), and (3) and hence cannot be influenced by changes in the government's balance sheet. However, it is possible that C and I, and hence the economy's growth rate, can be influenced by such changes.

Whether or not monetary policy affects C and I depends entirely on whether or not it is able to affect consumers' wealth, thereby influencing consumption.

We first consider a regime in which all money is "inside" money. In this regime, the government issues no bonds and holds a volume of firms' bonds equal to the value of the money supply:

$$M = B_f(G)$$

and

$$B_G = 0.$$

All money is matched or "backed" by government holdings of private debt. Such money is called "inside" money because it does not represent a net claim of the private sector, i.e., firms and households, against an outside sector. The claims that the money represents are just offset by government claims against firms. In such a regime, households' wealth is

$$W = \frac{M + V}{p} + \frac{B_f - B_f(G)}{p}$$

$$= \frac{M + V}{p} + \frac{B_f - M}{p}$$

$$= \frac{V + B_f}{p} = K.$$

Households' real wealth equals the real value of the capital stock. It follows that an open-market operation, i.e., a government exchange of money for firms' bonds, will have no influence on households' real wealth, and will not affect either C or I. The only affect will be to produce a proportional change in the price level. Thus, money is "neutral"; more precisely, under the inside money regime, the system "dichotomizes."

Now consider a regime in which all money is "outside" money. That is, the government owns no bonds issued by firms so that $B_f(G) = 0$. All money thus represents a net claim of households against an outside sector, the government. In this regime, households real wealth is

$$W = \frac{M + B_G + V + B_f}{p}$$

$$= \frac{M + B_G}{p} + K.$$

The constraint on open-market operations is now $dM = -dB_g$, since we assume $B_f(G)$ remains at zero. Recall that by virtue of equations (1) through (4) and (5), $dM/M = dp/p$. Now calculate the differential of W:

$$dW = \frac{p(dM + dB_G) - dp(M + B_G)}{p^2}$$

$$(8) \qquad = \frac{dM + dB_G}{p} - \frac{M + B_G}{p} \frac{dM}{M}$$

If the change in money is accomplished via open-market operations,

$$dW = -\frac{M + B_G}{p} \frac{dM}{M}$$

so that an increase in the money supply decreases the real wealth of households: it leaves $M + B_G$ unchanged, but causes an increase in the price level. The change in wealth will affect consumption through equation (6) and capital accumulation through equation (7). Thus in a regime of all "outside" money changes in the money supply accomplished through open-market operations are not "neutral," so that the system fails to dichotomize.

To establish the conditions under which money will be neutral in such a regime, we set dW equal to zero in (8) and rearrange to arrive at the condition:

$$\frac{dM}{M} = \frac{dB_G}{B_G}$$

This states that to be neutral, changes in the money supply must be accompanied by equiproportional changes in government bonds. In order to be neutral, changes in money must not alter the proportion of government bonds to money held by households.

Now consider a regime in which all money is "inside" money, so that it is always true that $M = B_f(G)$. Suppose that there are also government bonds outstanding, so that $B_G > 0$. The government only conducts open-market operations in firms' bonds, so that $dM = dB_f(G)$. Households' real wealth is then

$$W = \frac{M + V}{p} + \frac{B_G + B_f - M}{p}$$

$$= \frac{V + B_f}{p} + \frac{B_G}{p}$$

$$= K + \frac{B_G}{p} .$$

The differential of W is

$$dW = \frac{pdB_G - B_G dp}{p^2}$$

$$= \frac{dB_G}{p} - \frac{B_G}{p} \frac{dM}{M} .$$

This expression will equal zero only if

$$\frac{dB_G}{B_G} = \frac{dM}{M} ,$$

i.e., only if changes in money are accompanied by equiproportional changes in government bonds. Such a change couldn't be accomplished by open-market operations, but would have to occur as a result of a government

give-away program, for example, issuing a veterans' bonus. Under the
regime considered here, changes in money engineered via open-market
operations affect W and hence C and I and are not "neutral".

3. Unions and Real Wages

Here we present a "classical" analysis of the impact of unions on real wages and other aggregate variables. We assume that there are two kinds of jobs with skilled employment of N men and unskilled employment of n men. The total labor supply equals $N+n = \bar{N}$ men, which we assume is fixed. There is a union for skilled workers that sets a fixed money wage W_N for skilled workers. The union sets the wage for skilled workers so that it always exceeds the wage for unskilled labor, and so that there are always some skilled workers who can't find skilled jobs. If a person who is skilled can't find work as a skilled worker, he joins the unskilled labor force. There he can always find a job because the market for unskilled labor is competitive and has a market-determined money wage W_n.

There is one good in the model produced subject to the production function

$$(1) \qquad Y = F(N, n, K), \quad F_N, F_n, F_K > 0; \quad F_{NN}, F_{nn}, F_{KK} < 0,$$

$$F_{Nn} = F_{NK} = F_{nK} = 0.$$

The rest of the model is:

$$(2) \qquad \frac{W_N}{P} = F_N \qquad \qquad \text{marginal condition for skilled labor}$$

$$(3) \qquad \frac{W_n}{P} = F_n \qquad \qquad \text{marginal condition for unskilled labor}$$

$$(4) \qquad N+n = \bar{N} \qquad \qquad \text{full employment condition}$$

$$(5) \qquad C = c(Y-T), \quad 1 > C' > 0 \quad \text{consumption function}$$

$$(6) \qquad I = I(r-\pi), \quad I' < 0 \qquad \text{investment schedule}$$

(7) $C+I+G = Y$

(8) $M/P = m(r, Y)$, $m_r < 0 < m_Y$ portfolio balance schedule.

The model consists of eight equations in the eight endogenous variables
Y, N, n, C, I, r, p, and W_n. The exogenous variables are M, G, T, π,
K, W_N, and $\bar{N}$.

Equations (5), (6) and (7) define an IS curve in the r-Y plane.
Equations (1), (2), and (4) define an aggregate supply curve in the p-Y
plane, which can be combined with the portfolio balance schedule (8)
to form an LM curve.

To derive the differential of the aggregate supply curve,
compute the differential of (1) and (4):

$$dY = F_N dN + F_n dn + F_K dK$$

$$dN + dn = d\bar{N} = 0$$

or

$$dn = -dN.$$

So we have

(9) $dY = (F_N - F_n)dN + F_K dK.$

Since we assume that the union manages to keep W_N greater than W_n, it
follows that

$$F_N = \frac{W_N}{p} > \frac{W_n}{p} = F_n$$

or

$$F_N - F_n > 0,$$

so that from (9) output rises as skilled employment rises and unskilled employment falls. This occurs because of the greater productivity of skilled labor at the margin.

Computing the total differential of (2) gives

$$\frac{dp}{p} = \frac{-F_{NN}}{F_N} dN + \frac{dw_N}{W_N} \ ,$$

Substituting the above equation into (9) gives the total differential of the aggregate supply curve

$$(10) \qquad \frac{dp}{p} = \frac{-F_{NN}}{F_N} \frac{1}{(F_N - F_n)} dY + \frac{F_K F_{NN}}{F_N (F_N - F_n)} dK + \frac{dW_n}{W_n}$$

which has positive coefficients on dY and dW_N, and a negative coefficient on dK. The supply curve is upward sloping in the p-Y plane because in order for Y to increase, N must increase at the expense of n. This can occur if p rises, permitting W_N/p and hence F_N to fall. Substituting the differential of the aggregate supply curve (10) into the differential of the portfolio balance schedule gives

$$dr = \frac{1}{mr} \frac{dM}{p} - \frac{1}{mr} (m_Y - \frac{F_{NN}}{F_N} \frac{M}{p} \frac{1}{(F_N - F_n)}) dY - \frac{1}{mr} \frac{M}{p} \frac{F_{NN}}{F_N} \frac{F_K}{(F_N - F_n)} dK - \frac{1}{mr} \frac{M}{p} \frac{dW_N}{W_N}$$

which is the total differential of the LM curve. The LM curve is upward sloping in the r-Y plane, and shifts downward when M or K increases and when W_N decreases.

The equilibrium of the system occurs at the intersection of the IS and LM curves. To illustrate how the model works, consider the effects of an increase in government expenditures G, which we know shifts the IS curve outward and so drives the interest rate and output both up. The increase in aggregate demand causes the price level to rise, which with W_N fixed causes the real wage of skilled workers to

fall, creating incentives to hire more of them. Employment of unskilled labor falls because the money wage W_n rises more than proportionately with the price level. Notice that the model predicts that periods of high output are associated with a narrowing of the difference in wage rates between high skill and low skill workers.*

When the labor union imposes an increase in W_N, we have an example of "cost-push" inflation. The LM curve shifts back, driving output down and the interest rate up. Since output falls, we know that employment of unskilled labor rises. From the portfolio balance condition and the fact that r has risen while Y has fallen, we know that m(r, Y) has fallen; so M/p must also fall, meaning p must rise. This rise in p helps achieve the reduction in W_n/p required to stimulate employment of unskilled workers. From equation (3) we have that the differential of W_n obeys

$$\frac{dW_n}{p} = \frac{dp}{p} \frac{W_n}{p} + F_{nn} dn$$

Since dn > 0 and dp > 0, it is impossible to determine whether W_n rises or falls. Upward pressure on W_n is exerted by the increase in the price level p; but downward pressure is exerted by the increase in the number of workers seeking unskilled jobs because of the union's action in raising W_N.

In summary, then, the effect of the union's imposition of an increase in W_N is to raise the skilled real wage, but to decrease the

---

*Wachter has pointed to a tendency for interindustry wage differentials to narrow during periods of high output. See Michael L. Wachter, "Cyclical Variation in the Interindustry Wage Structure," American Economic Review, Vol. 60 (March 1970).

unskilled real wage. The effect on "average" real wages in the economy

is ambiguous. Similarly, the effect of the increase in W_N on some

average money wage level is also ambiguous.

There is a literature on the question of whether unions have

acted to affect the average level of real wages in the American economy.

The sort of model described here is perhaps potentially relevant for

this question, since in the American economy only about one-quarter of

the labor force is unionized.

4. Lange's Critique of the Classical Dichotomy

In a famous but confusing line of reasoning, Oscar Lange criticized classical models that dichotomize as being inconsistent with Walras' Law.* We ought immediately to be suspicious of such a criticism since we have seen above that the models that we have constructed obey versions of Walras' Law, even those models that dichotomize.

Lange characterized the real sector by the equations

(1) $q_i = d_i(p_1, \ldots, p_n) = d_i(\lambda p_1, \ldots, \lambda p_n)$ i=1, ..., n

(2) $q_i - q_i^o = e_i(p_1, \ldots, p_n) = e_i(\lambda p_1, \ldots, \lambda p_n)$, i=1, ..., n

$$\lambda > 0$$

Here p_i is the price of the ith good, measured in dollars per unit of the ith good; q_i is the rate of output of the ith good. Equations (1) are a set of n demand functions that are homogenous of degree zero in all prices, while equations (2) are a set of equilibrium conditions requiring that excess demands $q_i - q_i^o$, which are also homogeneous of degree zero in all prices, all equal zero.** Equations (1) and (2) are viewed as determining the rates of consumption q_i, i=1, ..., n, and the (n-1) relative prices p_i/p_1, i=2, ..., n. The demand and excess demand schedules build in budget constraints of the form $\sum_{i=1}^{n} p_i(q_i - q_i^o) = 0$, where q_i^o, i=1 ..., n, are quantities supplied. Then one of the excess demand equation is redundant (can be deduced from the others) so that the system (1) and (2) consists of 2n-1 independent equations for deter-

---

* Oscar Lange, <u>Price Flexibility and Full Employment</u>.

** For a description of general equilibrium systems see Henderson and Quandt, <u>Microeconomic Theory</u>.

mining the n quantities q_i, i=1, ..., n and the (n-1) relative prices p_i/p_1, i=2, ..., n. The absolute prices p_1 ..., p_n remain undetermined.

Now suppose that we add to this real system a single asset money, the stock of which is denoted M. Individuals are supposed to be satisfied with their holdings of this asset when the following quantity theory equation holds:

$$(3) \qquad \frac{M}{\sum\limits_{i=1}^{n} \delta_i p_i} = k \sum_{i=1}^{n} \sigma_i q_i$$

where the δ_i's and σ_i's are fixed (index number) weights; M is an exogenous variable and k is a parameter. Now notice that the system formed by (1), (2), and (3) is capable of determining absolute prices p_1, ..., p_n as well as the quantities q_1, ..., q_n. The role of equation (3) is only to determine the aggregate price index $\sum\limits_{i=1}^{n} \delta_i p_i$, which given the relative prices p_2/p_1, ..., p_n/p_1 determined by (1) and (2), pins down the absolute prices p_1, ..., p_n. So the system "dichotomizes."

Lange's criticism of that dichotomy was this. Suppose the system is initially in equilibrium, equations (1), (2) and (3) holding for a set of values $\bar{p}_1$, ..., $\bar{p}_n$, $\bar{q}_1$, ..., $\bar{q}_n$. Now increase the money supply. At the initial prices and quantites, there is an excess supply of money, as equation (3) is now violated. But notice that at the initial prices and quantities equations (1) and (2) continue to be satisfied. We thus encounter a situation in which there is an excess supply of money that is nowhere balanced by offsetting excess demands. Lange took this to mean that Walras' Law was violated, interpreting that as a flaw in the classical dichotomy. An excess supply of money seems

not to set up any pressures for the absolute prices of the n individual commodities to change.

There are two problems with this analysis. First, the q_i's are flows, while $M/\sum \delta_i p$ is a stock, so that the excess demands and supplies referred to in the above argument aren't comparable. Second, it is misleading simply to set down equations for excess demands and expect Walras' Law to hold unless care has been taken to insure that the demand functions are jointly determined so as to satisfy a budget constraint.

In the above model there is only one stock, money. There is therefore no portfolio problem facing individuals, there being no alternative assets for them to hold. So equation (3) is <u>not</u> a portfolio balance schedule, though it formally resembles some of the portfolio balance schedules that we encountered earlier. Instead, equation (3) must be interpreted as giving target wealth or a target wealth-income ratio. When actual wealth $M/\sum \delta_i p_i$ deviates from desired wealth $k\sum \sigma_i q_i$, it is reasonable to posit that individuals respond by accumulating or decumulating money at some rate per unit time. Thus suppose we posit the <u>ad hoc</u> accumulation schedule

$$\frac{\dot{M}}{\sum \delta_i p_i} = \phi \cdot [k\sum \sigma_i q_i - \frac{M}{\sum \delta_i p_i}], \quad \phi > 0.$$

At the same time we must also modify the budget constraint facing individuals to be

$$\sum_{i=1}^{n} p_i(q_i - q_i^0) + \dot{M} = 0$$

or

(4) $\quad \sum_{i=1}^{n} p_i(q_i - q_i^o) + \phi \sum_{i=1}^{n} \delta_i p_i [k \sum_{i=1}^{n} \sigma_i q_i - \dfrac{M}{\sum_{i=1}^{n} \delta_i p_i}] = 0.$

For a system of excess demand equations for the n goods constructed subject to the budget constraint (4), an "excess supply" of money clearly will always be accompanied by offsetting excess demands for goods as individuals attempt to decumulate money balances, making the demand for goods exceed available supplies. But notice that it is the stock excess supply of money _times_ the factor $\phi \sum \delta_i p_i$ that equals the excess flow demand for goods.

Demand and excess demand schedules constructed subject to (4) might be written in the form

(1') $\quad q_i = d_i(p_1, \ldots, p_n, \phi \sum \delta_i p_i [k \sum \sigma_i q_i - \dfrac{M}{\sum \delta_i p_i}])$

$\quad\quad = d_i(\lambda p_1, \ldots, \lambda p_n, \lambda \phi \sum \delta_i p_i [k \sum \sigma_i q_i - \dfrac{M}{\sum \delta_i p_i}]), \lambda > 0$

(2') $\quad q_i - q_i^o = e_i(p_1, \ldots, p_n, \phi \sum \delta_i p_i [k \sum \sigma_i q_i - \dfrac{M}{\sum \delta_i p_i}])$

$\quad\quad = e_i(\lambda p_1, \ldots, \lambda p_n, \lambda \phi \sum \delta_i p_i [k \sum \sigma_i q_i - \dfrac{M}{\sum \delta_i p_i}], \lambda > 0,$

which are homogeneous of degree zero in all arguments (not just prices alone). Equations (1'), (2') and (3) now form a system capable of determining $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$. Walras' Law obtains for this system.

Notice that the comparative statics of this system is formally identical with that of the system formed by equations (1), (2), and (3). That is, the system dichotomizes. Substituting equation (3) into (1') and (2') gives

$\quad\quad q_i = d_i(p_1, \ldots, p_n, 0)$

$$q_i - q_i^o = e_i(p_1, \ldots, p_n, 0),$$

which determine the q_i's and $(n-1)$ relative prices. Equation (3) then determines the absolute price level.

The following experiment shows the forces that propel the system formed by (1'), (2') and (3) from one equilibrium postion to another. Suppose that the system is initially in equilibrium. Then a jump occurs in the money supply (a gift from heaven—for there can't be any open market operations in a world with only one asset). At the initial prices and quantities, that upsets equations (1'), (2'), and (3) since individuals have larger wealth than they desire. They attempt to dispose of some of their money at a finite rate per unit time by purchasing additional goods at some rate per unit time. This causes prices to be bid up, which continues until the price level $\sum \delta_i p_i$ has increased proportionately with the money supply, all relative prices and quantities being the same as in the initial equilibrium. So the effect of a jump in the money supply is an instantaneous jump in all absolute prices with all real variables being left unaltered.

The subject of this section produced a large and tortuous literature, often characterized by careless treatment of distinctions between stocks and flows. Most of that literature concerned a model with only one asset, money, a model that therefore requires no theory of portfolio balance. That model is therefore quite different from the models that we have described in earlier sections.

5. The "Loanable Funds" Equation*

Consider a version of our classical model in which π is zero, in order to justify failing to carry around capital gains nuisance terms. The classical theory of interest asserts that (in the absence of a Pigou effect) the IS curve determines the interest rate. (In the presence of a Pigou effect or a nonzero π, the IS curve and portfolio balance condition mutually determine interest and the price level.) The classical theory of interest is thus often summarized by the statement that the interest rate adjusts to equate saving plus taxes to investment plus government expenditures. Disposable income Y_D satisfies

$$C + S = Y_D = Y - T - \delta K.$$

Substituting for Y from the national income identity gives

$$C + S = C + I + G + \delta K - T - \delta K$$

or

(1) $S + T = I + G.$

We can write the saving function as

$$S = Y - T - \delta K - C(Y-T- K, \ r-\pi)$$

(2) $S \equiv s(Y-T- K, \ r-\pi), \ 1 > s_1 > 0, \ s_2 > 0_y$

which follows from the definition of disposable income and the consumption function. The investment schedule as before is

(3) $I = I(q(N, \ K, \ r-\pi).$

Equations (1), (2), and (3) determine I, S, and r. Substituting (2) and (3) into (1) we obtain the IS curve

(4) $s(Y-T- K, \ r-\pi) + T = I(q(N, \ K, \ r-\pi)) + G,$

---

*Dale Henderson explained to me the material in this section.

which since Y and N are predetermined by the labor market and production function, is one equation in the endogenous variable r. This is the usual presentation of the classical theory of interest, and it is logically consistent, if somewhat simplified by the special assumptions ruling out Pigou effects and complicating movements in capital gains on the government debt.

The famous "loanable funds" equation is often used in discussing the classical theory of interest. While the loanable funds equation is itself valid, these expositions of the classical theory are usually erroneous.

To derive the loanable funds equation, we note that in the classical model, portfolio balance obtains continuously. It follows that

$$(5) \qquad \frac{\dot{M}}{P} = \frac{\dot{M}^D}{P}$$

where the dots denote right-hand time derivatives and the D superscript denotes demand. Adding (1) and (5) we obtain the famous "loanable funds" equation

$$(6) \qquad S + \frac{\dot{M}}{P} = I + (G-T) + \frac{\dot{M}^D}{P}$$

which says that real saving plus money creation equals investment plus "hoarding" plus the government deficit. This equation in effect says that the flow supply of bonds and equities equals the flow demand for bonds and equities. To see this, notice that saving satisfies

$$S = \frac{\dot{M}^D}{P} + \frac{(\dot{B} + \dot{V})^D}{P}$$

where $(\dot{B} + \dot{V})^D$ is households' total flow demand for firms' equities and

the government's interest bearing debt. Substituting the above equation

and the government's budget constraint into (6) gives

$$\frac{\dot{M}^D}{p} + \frac{(\dot{B} + \dot{V})}{p} + \frac{\dot{M}}{p} = I + \frac{\dot{M} + \dot{B}}{p} + \frac{\dot{M}^D}{p}$$

or

(7) $$\frac{(\dot{B} + \dot{V})^D}{p} = I + \frac{\dot{B}}{p} ,$$

which states the equality of the flow demand for bonds and equities with

the flow supply (firms are issuing equities at the real rate I per unit

time). Thus equation (7) states a condition that is equivalent with the

"loanable funds" equation (6).

Loanable funds theorists claim that (6), in conjunction with

(2), (3) and supplementary equations for $\dot{M}$ and $\dot{M}^D$, determines the interest

rate. But this cannot be in a classical model, since (1), (2), and (3)

are sufficient to determine S, I, and r. In the standard classical

model, the portfolio balance condition then determines p. The loanable

funds equation (6) holds, but it plays no independent role itself,

being merely the sum of (1) and (5). It only serves to verify that the

"flow" supply of bonds and equities equals the flow demand, and just

shows that the accounting has been done correctly.

After the publication of Keynes's General Theory, a large and

terribly tortuous literature developed about the relationship of the

loanable funds theory of interest to Keynes's liquidity preference

theory. The literature is full of stock-flow errors and failures to

recognize that especially in Keynes's model, but also in the classical

model, the theory of interest is macroeconomic and not of the character

of partial equilibrium.

Dynamic Analysis of a Keynesian Model

These notes describe the dynamics of the Keynesian macroeconomic model under two alternative assumptions about the formation of expectations of inflation. One assumption is that expectations of inflation are formed "adaptively" so that π is governed by the differential equation

$$D\pi = \beta\left(\frac{Dp}{p} - \pi\right), \quad \beta > 0$$

where D is the right-hand time derivative operator. The solution of the above differential equation is

$$\pi(t) = \pi(t_0)e^{-\beta(t-t_0)} + \beta\int_{t_0}^{t} e^{-\beta(t-s)}\frac{\dot{p}}{p}(s)ds,$$

so that $\pi(t)$ is formed as a geometric "distributed lag" of past actual rates of inflation. The other assumption under which the model will be analyzed is that of perfect foresight, so that $\pi = Dp/p$. Changing from the first to the second assumption about expectations will be seen to convert the particular model that we analyze from a Keynesian one to a classical one, even though the "structural" equations of the model remain the same.

The Model with Adaptive Expectations

The model is identical with the Keynesian model described above, except that it is augmented with a Phillips curve. We will take advantage of the linear homogeneity of the production function and write it in the intensive form

$$\frac{Y}{K} = F(1, \frac{N}{K}) = f(\frac{N}{K})$$

or

(1) $y = f(\lambda), \; f'(\lambda) > 0, \; f''(\lambda) < 0,$

where $y = \frac{Y}{K}$ and $\lambda = N/K$.

The marginal product condition for employment can be written.

(2) $\frac{w}{p} = f'(\lambda) \; (= \frac{\partial}{\partial N} K f(\frac{N}{K}) = \frac{\partial}{\partial N} F(K, N)).$

The Keynesian investment schedule will be written in the intensive form

$$\frac{\dot{K}}{K} = I(F_K - (r + \delta - \pi)), \; I' > 0$$

or

(3) $i = \frac{I}{K} = I(f(\lambda) - \lambda f'(\lambda) - (r + \delta - \pi)).$

Notice that

$$F_K(K, N) = \frac{\partial}{\partial K}(K f(\frac{N}{K})) = f(\frac{N}{K}) + K f'(\frac{N}{K})[-\frac{N}{K^2}] = f(\lambda) - f'(\lambda).$$

The consumption function is assumed to be linear and "proportional," which permits us to write it in the capital intensive form

$$\frac{C}{K} = z \cdot [\frac{Y}{K} - \frac{T}{K} - \frac{\delta K}{K}], \; 0 < z < 1$$

or

(4) $c = z(y - \tilde{t} - \delta)$

where $\tilde{t} = T/K$. The parameter z is the marginal propensity to consume.

We write the national income identity as

(5) $y = c + i + g + \delta$,

where $g = G/K$ and $c = c/K$.

The portfolio equilibrium condition is assumed to take the form

(6) $\dfrac{M}{pK} = m(r, y)$,

which can be rationalized by assuming the demand for money $m(r, Y)$ is homogeneous of degree one in output.

We posit that the evolution of the money wage is governed by the Phillips curve

(7) $\dfrac{Dw}{w} = h(\dfrac{N}{N^S}) + \pi$, $h' > 0$, $h(1) = 0$.

where N^S is the labor supply. Given π, equation (7) depicts a trade-off between the rate of employment relative to the labor supply and the rate of wage inflation. An increase in π shifts the Phillips curve upward by the amount of that increase.

The labor supply is exogenous and is governed by

(8) $N^S(t) = N^S(t_0)e^{n(t - t_0)}$

where n is the proportionate rate of growth of the labor supply.

The model is completed by specifying that expectations of inflation obey the adaptive scheme

(9) $\qquad \pi(t) = \pi(t_0)e^{-\beta(t-t_0)} + \beta \int_{t_0}^{t} e^{-\beta(t-s)} \frac{Dp(s)}{p(s)} \, ds.$

Collecting equations, the complete model is:

(1) $\qquad y = f(\lambda)$

(2) $\qquad \frac{w}{p} = f'(\lambda)$

(3) $\qquad i = I(f(\lambda) - f'(\lambda) - (r+\delta-\pi)) = \frac{\dot{K}}{K}$

(4) $\qquad c = z(y-\tilde{t}-\delta)$

(5) $\qquad y = c+i+g$

(6) $\qquad \frac{M}{pK} = m(r, \ y)$

(7) $\qquad \frac{Dw}{w} = h(\frac{\lambda K}{N^S})+\pi$

(8) $\qquad N^S(t) = N^S(t_0)e^{n(t-t_0)}$

(9) $\qquad \pi(t) = \pi(t_0)e^{-\beta(t-t_0)} + \beta \int_{t_0}^{t} e^{-\beta(t-s)} \frac{Dp(s)}{p(s)} \, ds.$

$\qquad K(t) = K(t_0)+\int_{t_0}^{t} i(s)K(s)ds$

Given the initial conditions $w(t_0)$, $\pi(t_0)$, $K(t_0)$, and given time paths for the exogenous variables M, g, and $\tilde{t}$ for $t \geq t_0$, the model will generate time paths of the endogenous variables y, λ, K, c, w, p, r and π. Notice that even though w, π, and K are fixed or "exogenous" at a point in time, being inherited from the past according to (3), (7), and (9), they are endogenous variables from the point of view of our dynamic analysis. The analysis determines their evolution over time.

The momentary equilibrium of our system can be determined by solving equations (1)-(6) for IS and LM curves. The IS curve gives the

combinations of r and y that make the demand for output equal to the supply. It is derived by substituting (3) and (4) into (5):

(10) $y = z(y-\tilde{t}-\delta)+I(f(\lambda)-\lambda f(\lambda)-(r+\delta-\pi))+g+\delta$.

Since $f'(\lambda) > 0$, we can invert (1) and obtain

$$\lambda = \lambda(y), \ \lambda'(y) = \frac{1}{f'(\lambda)} > 0, \ \lambda''(y) = \frac{-f''(\lambda)\lambda'(y)}{f'(\lambda)^2} > 0.$$

Substituting this into (10) yields the IS curve:

$$y = z(y-\tilde{t}-\delta)+I(y- \frac{\lambda(y)}{\lambda'(y)} -(r+\delta-\pi))+g+\delta.$$

The slope of the IS curve in the r-y plane is given by

$$\frac{dy}{dr}\bigg|_{IS} = \frac{-I'}{1-z-I'\frac{\lambda(y)\lambda''(y)}{\lambda'(y)^2}} ,$$

which is of ambiguous sign since $\lambda''(y) > 0$. The denominator of the above expression is Hicks's "supermultiplier," the term $I'\lambda\lambda''/\lambda'^2$ being the marginal propensity to invest out of income. We will assume that this term is less than the marginal propensity to save, so that the IS curve is downward sloping. The position of the IS curve depends on the parameters g, $\tilde{t}$, and π in the usual way. An increase in π shifts the IS curve upward by the amount of that increase.

We can write the marginal productivity condition for labor as

$p = w\lambda'(y)$.

Substituting the expression for p into (6) yields the LM curve:

$M = w\lambda'(y)Km(r, y)$,

the slope of which is easily verified to be positive in the r-y plane. The LM curve shows the combinations of r and y that guarantee portfolio balance. Its position depends on M, w, and K, all of which are parameters at a point in time.

The momentary equilibrium of the system is determined at the intersection of the IS and LM curves. That equilibrium will in general be a nonstationary one, the interest rate, the real wage, and the capital-labor ratio possibly changing over time. However, given fixed values of g, $\tilde{t}$, and $\dot{M}/M$, the system may over time approach a steady state in which the interest rate, real wage, and employment-capital ratio are fixed, while prices and wages change at a rate equal to $\dot{M}/M$ minus n. We will use two curves to characterize the steady-state growth path in the r-y plane. The first is simply a vertical line at the steady-state output-capital ratio, which we denote by y*. From (5), the rate of growth of capital is

$$i = y - z(y - \tilde{t} - \delta) - g - \delta.$$

Since $D(K/N)/(K/N) = DK/K - DN/N = i - n$,

we have

$$\frac{D(K/N)}{K/N} = y - z(y - \tilde{t} - \delta) - g - \delta - n.$$

Setting $D(K/N)$ to zero and solving for y yields the value of y*:

$$y* = \frac{n + g + \delta(1-z) - zt}{1-z} .$$

This is the value of the output-capital ratio at which the capital-labor ratio is stationary, i.e., unchanging through time. We show y* as a

vertical line in Figure 1. On our assumptions, the steady-state value
of y is independent of the interest rate.

If firms are to be content to increase the capital stock at
the steady-state rate n, so that i-n equals zero, we require

(11) $I(y - \frac{\lambda(y)}{\lambda'(y)} - (r+\delta-\pi)) - n = 0$

which is implicitly an equation that tells us what r+δ-π must be if the
system is to be in a steady-state equilibrium at a given y. Taking the
total differential of the above equation and rearranging gives

$$\frac{d(r+\delta-\pi)}{dy} = \frac{\lambda''\lambda}{\lambda'^2} > 0$$

as the slope of the locus of points in the r-y plane along which (11) is
satisfied. The slope is positive, reflecting the direct dependence of
the marginal product of capital on the output-capital ratio. We call
(11) the capital-market equilibrium curve and label it KE. Note that an
increase in π causes the KE curve to shift upward by the full amount of
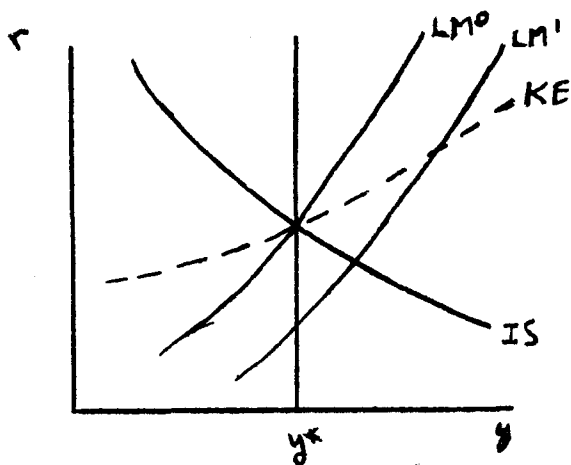the increase.

Figure 1

The determination of momentary and steady-state equilibrium can be illustrated with Figure 1. Notice that the IS curve has been drawn so that it intersects the KE curve at y*, the steady-state output-capital ratio. That this must be so can be verified as follows. Along the KE curve, equation (11) is satisfied. Substituting for I( ) from (11) into the IS curve gives

$$y = z(y-\tilde{t}-\delta)+n+g+\delta$$

or

$$y = \frac{n+g+\delta(1-z)-z\tilde{t}}{1-z} \text{ ,}$$

which is identical with our expression for y*. A steady state is determined at the intersection of the KE and IS curves. Momentary equilibrium is determined at the intersection of the IS and LM curves. If the IS and LM curve intersect at an r-y combination below the KE curve, capital is growing more rapidly than n at that moment. The model possesse mechanisms propelling over time the intersection of the IS and LM curve toward the intersection of the IS and KE curves. The dynamics of capital and the money wage are the key elements in the mechanism.

To illustrate how the model works, suppose that the system is initially in a full, steady-state equilibrium, the IS, LM, and KE curves all intersecting at y*, as in Figure 1. Suppose that $\dot{M}/M = n$, so that the equilibrium rate of inflation is zero. We know this because in the steady-state r and y and therefore m(r, y) are constant through time. Therefore, M/pK must be constant through time, so that $\dot{M}/M-\dot{p}/p-\dot{K}/K = 0$ or $\dot{M}/M-\dot{p}/p-n = 0$. The system is in a steady state so that the LM, IS, and KE curves are unchanging through time. Since in that steady state $\dot{w}/w = \dot{p}/p = \pi = \dot{M}/M-n = 0$, we know from the Phillips curve that N/N^s must

equal unity. Now suppose that at some point in time there occurs a once-and-for-all jump in M, engineered via an open-market operation that leaves M/M unaltered. To simplify matters, we will suppose that remains fixed at zero, its steady-state value, during the movement to a new steady-state. So we temporarily suspend (9) and substitute $\pi(t) = 0$. We also assume that g and $\tilde{t}$ are constant over time. The immediate effect of the jump in the money supply is to shift the LM curve to the right, say to LM, in Figure 1. The result is an instantaneous jump in employment and the output-capital ratio. Employment now exceeds the labor supply, causing the money wage to adjust upward over time, as described by the Phillips curve (7). In addition, the nominal interest rate has fallen, creating a larger discrepancy between the marginal product of capital, which has risen, and the real rate of interest. Firms respond by adding to the capital stock at a rate exceeding n—this occurs at each moment the momentary solution is at an r-y combination below the KE curve. Since capital is growing faster than the money supply, and since money wages are rising over time, the LM curve shifts upward over time, from LM, toward LM_0 in Figure 1. To show this, notice that, for fixed values of y, logarithmic differentiation of the LM curve with respect to time yields

$$(12) \qquad \dot{r}\Big|_{LM} = \frac{m(r,\ y)}{m_r}[\frac{\dot{M}}{M} - \frac{\dot{K}}{K} - \frac{\dot{w}}{w}].$$

If the expression in brackets is negative, then at each value of y, the r that maintains portfolio balance is increasing over time. Now when enough time has passed to move the LM curve back to LM^0, so that y and r are back at their initial values, the adjustment is not yet complete. When the LM curve has shifted back to LM^0, the value of y and λ are

at their initial values. But since $\dot{K}/K$ exceeded n all during the inter-
vening period, we know that N/N^S, which had the initial value of unity,
since $\dot{w}/w = 0$ initially, now exceeds unity. For $N/N^S = \lambda K/N^S$; λ has
returned to its initial value, but K has grown faster than n at each
intervening moment, so that N/N^S must now exceed unity. The Phillips
curve therefore implies that $\dot{w}/w > 0$, implying that the LM curve continues
to shift up over time (see (12)) since $\dot{M}/M-\dot{K}/K-\dot{w}/w$ is still less than
zero even though $\dot{K}/K$ has returned to its steady-state value once the LM
curve has come back to its initial position. The system must therefore
"overshoot," having y fall below y* as wages rise and the LM curve moves
to an intersection with the IS curve above the KE curve. During moments
when such an intersection occurs, DK/K < n, which in itself helps to
increase $\dot{M}/M-\dot{K}/K-\dot{w}/w$, and thus tends to reverse the direction of move-
ment of the LM curve over time. Given stability, eventually this effect
dominates the tendency of rising wages to shift the LM curve to the left,
so that the LM curve starts moving back toward the right. Furthermore,
as λ falls and the ratio K/N^S falls with the passage of time (notice
that K is growing more slowly than N^S so long as the LM curve intersects
the IS curve above the KE curve), (N/N^S) and hence Dw/w fall over time.
Depending on the particular parameter values, it is possible, though not
necessary, that N/N^S actually falls below unity, in which case Dw/w
becomes negative. This means that w has "overshot" its new steady-state
value and must rise again, which requires another "boom" period during
which $N/N^S > 1$. If w is falling (and $N/N^S < 1$) when the LM curve
again shifts back to its original, pre-open-market-operation position,
it follows that the LM curve continues shifting toward the right, inducing
a y above y* again. Thus, it is possible that the approach back to the

steady state is an oscillatory one, characterized by alternating periods of boom and recession. The model thus has implicit in it, depending on the particular parameter values, a theory of the business cycle.

The alternative possibility is that following the initial jump in y, y overshoots the steady-state value y* only once. As the LM curve shifts leftward and passes an intersection with the IS curve at y*, it can happen that w rises continuously toward its new steady-state value and fails to overshoot it, so that $N/N^S \geq 1$ throughout the adjustment process. In this lucky state of affairs, the "boom" is followed by no bust. Notice that in this case the "stabilizing" force is provided by the fact that DK/K < n when the LM curve intersects the IS curve above the KE curve, tending to drive the LM curve rightward, toward the steady state.

The adjustment process in response to the once-and-for-all jump in M would be even more complex if we were to permit π to respond to the occurrence of actual inflation, say by restoring our equation (9). For then the IS curve and the KE curve would shift upward during the early part of the transition as π increases in response to the emergence of inflation. It is easily verified that when π changes, the IS curve continues to intersect the KE curve at y*. The result of allowing π to depend on past values of $\dot{p}/p$ is to accentuate the "over-shooting" or cyclical phenomenon. Following the original jump in M, the LM curve will now be shifting toward an intersection of y* with an IS curve associated with a positive rate of expected inflation, since inflation has occurred during the transition up to that moment. Since π is positive and, as we verified earlier, $N/N^S > 1$ at the moment the LM curve has shifted back enough to intersect an IS curve at y*, wages are

rising even faster at that moment than they would have been had π been zero throughout the transition. Assuming that the system is dynamically stable, the final resting place for all variables will be the same as if π had remained at its steady-state value of zero throughout the adjustment process; but the path to steady-state equilibrium may be much different.

The final effect of the once-and-for-all jump in M, once the system has returned to its steady state, is to leave all real variables unaltered and to increase the price level and money wage proportionately with the money supply. The variables r and y have steady-state values determined at the intersection of the IS and KE curves, which aren't affected by the jump in M. In the new steady-state $\dot{w}/w$ must be zero, which means that N/N^S must be unity, which means since λ is at its initial steady-state value, that K/N^S must be at its initial steady-state value. This means that once enough time has elapsed after the open market operation occurred, K(t) is what it would have been had no open market operation occurred. In other words, the open market operation redistributes investment over time, but does not affect the capital stock in moments far into the future. Since $M/pK = m(r, y)$ is not moved from its steady-state value, and K eventually returns to the path it would have followed without the open market operation, p must eventually have increased proportionately with the initial jump in M.

Though this model is clearly Keynesian in its momentary or point-in-time behavior, its steady-state or long-run properties are "classical" in the sense that real variables are unaffected by the money supply. The real variables are determined at the intersection of the KE and IS curves which are determined by the propensities to save and invest and the government's fiscal policies. In the steady-state, the price level must adjust so that the LM curve passes through the intersection of the IS and KE curves.

Perfect Foresight ($\pi = Dp/p$)

We now abandon equation (9) and for it substitute the assumption of perfect foresight or "rationality":

(9') $\pi(t) = Dp(t)/p(t)$,

where we continue to interpret D as the right-hand time derivative operator. Equation (9') asserts that people accurately perceive the right-hand time derivative of the log of the price level, the rate at which inflation is proceeding.

The dynamics of the model in response to shocks is very much different when (9') replaces (9). To solve the model, we begin by substituting (9') into (7) to obtain.

(13) $\dfrac{Dw}{w} = h\left(\dfrac{\lambda K}{N^s}\right) + \dfrac{Dp}{p}$.

Differentiating (2) logarithmically with respect to time gives

(14) $\dfrac{Dw}{w} = \dfrac{f''(\lambda)}{f'(\lambda)} D\lambda + \dfrac{Dp}{p}$.

Equating (13) with (14) gives

(15) $h\left(\dfrac{\lambda K}{N^s}\right) = \dfrac{f''(\lambda)}{f'(\lambda)} D\lambda$, where $\dfrac{f''(\lambda)}{f'(\lambda)} < 0$.

Now (15) is a differential equation in the employment capital ratio λ, which may be solved for λ in terms of past values of K and N^s. To illustrate, suppose that $f(\lambda)$ is Cobb-Douglas, so that

$$y = f(\lambda) = A\lambda^{(1-\alpha)}.$$

Then we have

$$f'(\lambda) = A(1-\alpha)\lambda^{-\alpha}$$

$$f''(\lambda) = -\alpha(1-\alpha)A\lambda^{-\alpha-1}$$

$$\frac{f''(\lambda)}{f'(\lambda)} = \frac{-\alpha}{\lambda} \ .$$

Also suppose that $h(\lambda K/N^S)$ takes the form

$$h(\frac{\lambda K}{N^S}) = \gamma \cdot \log\frac{N}{N^S}$$

$$= \gamma\log N - \gamma\log N^S.$$

where log denotes the natural logarithm. Then (15) becomes

$$(16) \qquad \gamma\log N - \gamma\log N^S = -\alpha\,\frac{\dot{\lambda}}{\lambda} = -\alpha D\log N + \alpha D\log K.$$

Rearranging, we have

$$(\gamma+\alpha D)\log N = \gamma\log N^S + \alpha D\log K$$

or

$$(\frac{\gamma}{\alpha}+D)\log N = \frac{\gamma}{\alpha}\log N^S + D\log K.$$

This is a linear, first-order differential equation in log N. To find

its solution, divide through by $\frac{\gamma}{\alpha}+D$ to get

$$\log N = \frac{1}{D+\frac{\gamma}{\alpha}}\,[\frac{\gamma}{\alpha}\log N^S + D\log K].$$

Notice that

$$\frac{1}{D+\frac{\gamma}{\alpha}} = \frac{e^{(s-t)(\frac{\gamma}{\alpha}+D)}}{\frac{\gamma}{\alpha}+D}\,\Bigg|_{s=-\infty}^{t}$$

$$= \int_{-\infty}^{t} e^{(s-t)(\frac{Y}{\alpha}+D)} ds.$$

$$= \int_{-\infty}^{t} e^{(s-t)\frac{Y}{\alpha}e^{(s-t)D}} ds.$$

Recalling that $e^{(s-t)D}x(t) = x(s)$, we have

$$\log N(t) = \int_{-\infty}^{t} e^{(s-t)\frac{Y}{\alpha}e^{(s-t)D}}[\frac{Y}{\alpha} \log N^s(t)+D\log K]ds$$

$$(17) \qquad \log N(t) = \frac{Y}{\alpha}\int_{-\infty}^{t} e^{(s-t)\frac{Y}{\alpha}}\log N^s(s)ds+\int_{-\infty}^{t} e^{(s-t)\frac{Y}{\alpha}} \frac{DK(s)}{K(s)} ds.$$

Equation (17) is the solution to equation (16), and expresses log N at t as distributed lags of past values of the labor supply and capital stock. Since these are predetermined at time t, we immediately know that employment and hence output will not respond at t to the imposition of shocks to the system at t.

Given the value of N at t determined from some version of (17), and given the quantity of K inherited from the past, output is determined by equation (1), the real wage by (2), and c by (4). Given c and y, (5) then determines i. Given i and λ, equation (4) determines r−π at t. Equation (7) determines Dw/w−Dp/p.

Notice that given entire time paths from now until forever for the fiscal variables g and $\tilde{t}$, the model determines entire time paths of the real variables without using the portfolio balance equation. In effect, then, this model dichotomizes.

All real variables have now been determined and it remains only to determine the values of p and Dp/p at instant t. They are determined by the portfolio balance condition in the manner indicated in our notes on the classical model. To illustrate, suppose the portfolio balance equation (6) assumes the special form

(18) $\dfrac{M}{pK} = e^{\beta r}y. \quad \beta < 0$

We know that in this system, r is determined by (5) which we express by inverting (3) and writing

$$r = f(\lambda) - \lambda f'(\lambda) - \delta + \pi + \xi(i), \quad \xi' < 0.$$

Substituting this into (18) gives

$$\dfrac{M}{pK} = y \exp\beta(f(\lambda) - \lambda f'(\lambda) - \delta + \pi + \xi(i)).$$

or

$$\log M - \log p - \log K = \log y + \beta[f(\lambda) - \lambda f'(\lambda) - \delta + \pi + \xi(i)].$$

Substituting Dp/p for π and rearranging gives

$$[\dfrac{1}{\beta} + D]\log p = \dfrac{1}{\beta}[\log M - \log K - \log y - \beta[f(\lambda) - \lambda f'(\lambda) - \delta + \xi(i)]].$$

Pursuing a line of reasoning identical to that in our Notes on the Classical Model, it is verified that the solution of this differential equation is

(19) $\log_e p(t) = -\dfrac{1}{\beta}\displaystyle\int_t^{\infty} e^{(s-t)/\beta}[\log M(s) - \log K(s) - \log y(s) - \beta[f(\lambda(s)) -$

$$\lambda(s)f'(\lambda(s)) - \delta + \xi(i(s))]]ds.$$

(We are again imposing a terminal condition that suppresses the term $ce^{-t/\beta}$ that should be added to the above solution.) Equation (19) expresses the current price level as a function of the entire <u>future</u> paths of the money supply, the capital stock, the employment-capital ratio λ, and the rate of investment. The value of π is also determined by (19), and can be obtained by differentiating the path given by (19)

with respect to time from the right. Notice that in this model, the entire time paths of the variables appearing on the right side of (19) can all be determined before the current price level is determined. That is, we showed how (17) determines the values of N and λ at t, and how this readily enables us to calculate the rate of growth of capital, i. This enables us to update the capital stock, and so to determine subsequent values of N. Proceeding in this way, given the exogenous fiscal policy variables, the entire time paths of all the real variables can be determined before determining the price level at any moment.

Since all of the structural equations remain the same, the model continues to be characterized by the IS, LM, and KE curves. The steady-state is determined at the intersection of the IS and KE curves. But now the instantaneous value of y is determined by substituting the value of λ determined by the solution to (15), e.g., (17), into the production function. The momentary equilibrium value of y so determined, call it $\hat{y}$ depicted in Figure 2, when substituted into the IS curve gives the momentary value of $r-\pi$. The LM curve must pass through the intersection of the IS curve and the vertical line at $\hat{y}$; jumps in the price level occur to assure this.
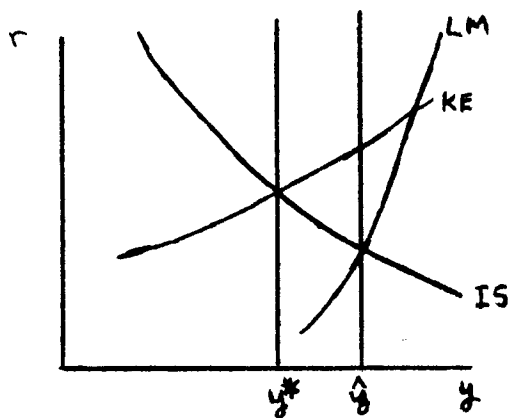

Figure 2

If we return to the experiment performed earlier under adaptive expectations, we shall see how radically the substitution of (9') for (9) alters the adjustment dynamics. We assume that the system is initially in a steady state and that $\dot{M}/M = n$ so that $\dot{w}/w = \dot{p}/p = \pi = 0$. At the current moment there occurs an <u>unexpected</u> jump in M that leaves the right-hand derivative DM/M unchanged.* The only result is an instantaneous jump in p and w proportional to the jump in M with all real variables unaltered. Since the solution to (15), e.g., (17), implies that y will not jump in response to the jump in M, the price level must jump to keep the LM curve passing through the intersection of the KE and IS curves. To show this explicitly for the sample portfolio balance equation that leads to (19) as the solution for the price level, suppose that the initial path of the log of the money supply was expected to be

$\log M(s)$, $s \geq t$.

After the once-and-for-all jump in the money supply that leaves DM/M unaltered, the new path of the log of the money supply is expected to be

$\emptyset + \log M(s)$, $s \geq t$, $\emptyset > 0$.

The new price level p'( ) differs from the initial price level given by (19) according to

$$\log p'(t) - \log p(t) = -\frac{1}{\beta} \int_t e^{(s-t)/\beta} [\log M(s) + \emptyset] ds$$

$$+ \frac{1}{\beta} \int_t^\infty e^{(s-t)/\beta} \log M(s) ds$$

$$= -\frac{1}{\beta} \int_t^\infty \emptyset e^{(s-t)/\beta} ds$$

$$= \emptyset.$$

So the price level jumps by the same multiplicative factor $\emptyset$ as does the money supply. Notice that in the above calculations we use the fact that the paths of the variables $K(s)$, $\lambda(s)$, and $i(s)$ will be unaltered when the money supply jumps.

Now consider the response of the system to a jump in M which had previously been anticipated. Suppose that the money supply follows the path

$$\log M(s) = \log M(t_0) + n(s-t_0), \quad t_0+\theta \geq s,$$

$$\log M(s) = \log M(t_0) + \emptyset + n(s-t_0), \quad t_0+\theta \leq s,$$

so that the money supply is expected to and does jump at $t_0+\theta$. Using (19) and ignoring the terms in the real variables, we compute the price path for $t < t_0 + \theta$ as

$$\log p(t) = -\frac{1}{\beta} \int_t^{t_0+\theta} e^{(s-t)/\beta} n(s-t_0)\,ds - \frac{1}{\beta} \int_t^{t_0+\theta} e^{(s-t)/\beta}(n(s-t_0)+\emptyset)\,ds$$

$$+ \text{ other terms, } \quad t_0 \leq t \leq t_0 + \theta$$

$$= -\frac{1}{\beta} \int_t^{\infty} e^{(s-t)/\beta} n(s-t_0)\,ds - \frac{\emptyset}{\beta} \int_t^{t_0+\theta} e^{(s-t)/\beta}\,ds + \text{ other terms}$$

$$= -\frac{1}{\beta} \int_t^{\infty} e^{(s-t)/\beta} n(s-t_0)\,ds + \emptyset e^{(t_0+\theta-t)/\beta} + \text{ other terms.}$$

$$t_0 \leq t \leq t_0 +$$

For $t \geq t_0 + \theta$, the solution for the price level is

$$\log p(t) = -\frac{1}{\beta} \int_t^{\infty} e^{(s-t)/\beta} n(s-t_0)\,ds + \emptyset + \text{ other terms.}$$

The jump in the money supply at $t_0+\theta$ is reflected in earlier values of the price level. The above calculations show that the price level is continuous at $t_0+\theta$, so that no jump occurs in p at the moment when M jumps. However, the expected rate of inflation π does jump at $t_0+\theta$. For notice that

$$\frac{d}{dt}\, \emptyset e^{(t_0+\theta-t)} = -\emptyset\beta e^{t_0+\theta-t} \, .$$

At $t_0+\theta$, therefore, the left-hand derivative of the log of the price level exceeds the right-hand derivative by $-\emptyset\beta > 0$. So there is a sudden fall in π at $t_0+\theta$. It is this downward jump in π, leading to a downward jump in r, that stimulates the demand for money enough at t_0+ to guarantee that portfolio balance is maintained in the face of the jump in M that occurs at that moment.

The two experiments just performed show how essential it is to distinguish between jumps in policy variables that are anticipated and unanticipated in the model with perfect foresight.

The effect of substituting the perfect foresight assumption (9') for the adaptive expectations mechanism (9) has been to convert the model from one with Keynesian momentary behavior to one with classical momentary behavior. In the system with perfect foresight, (9'), money is a "veil," momentarily as well as in the long run. Jumps in the money supply don't cause any real movements of the sort that they do in the system with adaptive expectations.

The distinction between the models under (9) and (9') is mathematically subtle. Under the adaptive expectations scheme (9), the model must be manipulated under the "Keynesian" assumption that the money wage does not jump at a point in time, so that the Phillips curve

(7) gives the time derivative of the wage (= the right-hand time derivative = the left-hand time derivative). Essentially, this is because at any moment t, equations (8) and (9) make $N^S(t)$ and $\pi(t)$ predetermined from past variables. Of course $K(t)$ is also inherited from the past. Equations (1)-(7) then form a system of seven equations in the seven endogenous variables $y(t)$, $\lambda(t)$, $i(t)$, $c(t)$, $p(t)$, $r(t)$, and $Dw(t)/w(t)$. The model is incapable of restricting any additional variables, in particular $w(t)$, at the moment t. So $w(t)$ must be regarded as fixed and inherited from the past at each point in time.

However, in the system with $\pi=Dp/p$, it is employment that is predetermined at any moment in time by the differential equation (15). Since employment is predetermined at t, say by (17), y, λ, and w/p are also predetermined and constrained to change continuously as functions of time. They cannot jump at a point in time. But if w/p can't jump, and neither can K or y, then if M jumps at a point in time we know that p and w must jump in order to satisfy the portfolio balance equation at each moment.

Exercises

1. Under both adaptive expectations and perfect foresight, perform a dynamic analysis of Tobin's Dynamic Aggregative Model, formed by substituting for equation (3) the following equation

$$r + \delta - \pi = f(\lambda) - f'(\lambda).$$

Now instantaneous equilibrium occurs at the intersection of the KE and LM curves, while steady-state equilibrium occurs at the intersection of the KE curve and y* line.

2. For both adaptive expectations and perfect foresight, analyze the dynamics of the model when the Phillips curve is modified to assume the form.

$$\frac{\dot{w}}{w} = h\left(\frac{N}{N^s}\right) + \alpha\pi, \quad h' > 0, \quad h(1) = 0, \quad 0 < \alpha < 1.$$

3. For both adaptive expectations and perfect foresight, analyze the dynamics of the model where the monetary authority pegs the nominal interest rate r at each instant, letting the money supply be whatever it must to guarantee portfolio balance.

Footnote

If the jump in M that occurs at some moment t had previously been expected to occur, it would not cause a jump in p at t. This follows from equation (19), which implies that p(s) will be a continuous function of time at t, even if there is a discontinuity in M(s) at t or anywhere else. However, $\pi(t)$ will jump at t in response to a previously anticipated jump in M. By contrast, the experiment that we are performing here is one in which the monetary authority at t suddenly and unexpectedly moves from a previously planned money supply path of M(s), $s \geq t$, to a new planned path of $e^{\emptyset}M(s)$ for $s \geq t$.

Notes on the Investment Schedule

As a comparison of the "Keynesian" model with Tobin's Dynamic
Aggregative Model reveals, whether or not it is assumed that there
exists a market in stocks of capital at each moment has drastic theoretical
implications, particularly about the potency of fiscal policy as a
device for inducing short-run movements in output and employment. The
structure of the Keynesian model depends sensitively on ruling out a
market in existing stocks of capital and instead positing a demand
schedule on the part of firms for a finite rate of addition per unit
time to their capital stocks. That element of the Keynesian model is
perhaps its most essential piece, ruling out as it does the movements of
capital across firms and industries that thwart fiscal policy in Tobin's
model; at the same time, the investment schedule is the weakest part of
the Keynesian model from a theoretical viewpoint, being defended (at
least until recently) on a very ad hoc basis.

These notes describe the most successful attempt to rationalize
the Keynesian investment schedule, a line of work due to Eisner and
Strotz, Lucas, Gould, and Treadway.[*] The key to the theory is the
assumption that there are costs associated with adjusting the capital
stock at a rapid rate per unit of time, and that these costs increase
rapidly with the absolute rate of investment, so rapidly that the firm

---

[*] R. Eisner and R. H. Strotz. "Determinants of Business Investment,"
Research Study Two in Impacts of Monetary Policy, (Englewood Cliffs,
N. J.: Prentice Hall, 1963); R. E. Lucas, "Adjustment Costs and the
Theory of Supply," Journal of Political Economy, 75 (1967), 321-334;
J. P. Gould, "Adjustment Costs in the Theory of Investment of the
Firm," Review of Economic Studies, 35 (1969) 47-56; A. B. Treadway,
"On Rational Entrepreneurial Behavior and the Demand for Investment,"
Review of Economic Studies, 36 (1969), 227-240.

never attempts to achieve a jump in its capital stock at any moment. These adjustment costs occur at a rate per unit of time (measured in capital goods per unit of time) described by the twice differentiable function $C(\dot{K})$, which obeys

$$C'(\dot{K}) \gtrless 0 \text{ as } \dot{K} \gtrless 0,$$

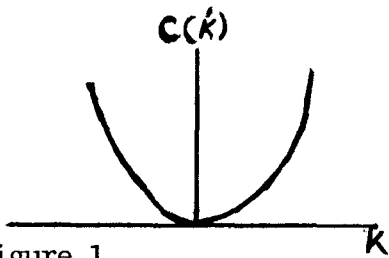$$C''(\dot{K}) > 0, \; C(0) = 0.$$

$C(\dot{K})$

Figure 1

Costs of adjusting the capital stock are nonnegative and increase at an increasing rate with the absolute value of invesmtent.

The firm's discounted net cash flow net of costs of adjustment is defined to be

$$f(N(t),K(t),\dot{K}(t),t) = e^{-rt}[p(t)F(K(t),N(t)) - w(t)N(t) - J(t)\delta K(t)$$

$$- J(t)\dot{K}(t) - J(t)C(\dot{K})]$$

where $J(t)$ is the price of capital goods at time t, and r is the instantaneous interest rate, assumed constant over $[0,T]$. We continue to assume that $F(K,N)$ is linearly homogeneous in K and N. The firm chooses paths of N and K over time to maximize its present value over the time interval $[0,T]$, which is

$$PV = \int_0^T f(N(t),K(t),\dot{K}(t),t) \; dt + S(T)K(T)e^{-rt}$$

where $S(T)K(T)$ is the scrap value, if any, of the capital stock at time T. We think of T as being in the very distant future. The firm operates in competitive markets for output and labor, being able to rent all the

labor it desires at the wage w, and to sell all the output it wants at the price p. The firm starts out with capital stock K(0).

Among the necessary conditions for an extremum for present value are the "Euler equations"

$$\frac{\partial f}{\partial N} = 0 \qquad\qquad t \epsilon [0,T]$$

$$\frac{\partial f}{\partial K} - \frac{d}{dt} \frac{\partial f}{\partial \dot{K}} = 0 \qquad\qquad t \epsilon [0,T].$$

(A heuristic explanation of these Euler equations is contained in the appendix to these notes.) Evaluating these derivatives, we have

$$\frac{\partial f}{\partial N} = e^{-rt}[p(t)\frac{\partial F}{\partial N} - w(t)]$$

$$\frac{\partial f}{\partial K} = e^{-rt}[p(t)\frac{\partial F}{\partial K} - J(t)\delta]$$

$$\frac{\partial f}{\partial \dot{K}} = -e^{-rt}[J(t) + J(t)C'(\dot{K})]$$

$$\frac{d}{dt} \frac{\partial f}{\partial \dot{K}} = -e^{-rt}[\dot{J}(t) + \dot{J}(t)C'(\dot{K}) + J(t)C''(\dot{K})\ddot{K}]$$
$$\qquad\qquad + re^{-rt}[J(t) + J(t)C'(\dot{K})].$$

So for our problem the Euler equations become

(1) $\qquad p(t)\frac{\partial F}{\partial N} - w(t) = 0$ or $\frac{\partial F}{\partial N} = \frac{w(t)}{p(t)}$

and

(2) $\qquad p\frac{\partial F}{\partial K} - J\delta - rJ + \dot{J} - (rJ-\dot{J})C'(\dot{K}) + JC''(\dot{K})\ddot{K} = 0$

Equation (1) requires that the marginal product of labor equal the real wage at each moment, an equation that determines the labor-capital ratio at each moment (since F(N,K) is linearly homogeneous). Equation (2) is

a differential equation that determines the (finite) rate of growth of the capital stock at each moment. To simplify the problem, we now assume that firms expect the prices $J(t)$, $p(t)$, and $w(t)$ to grow over time at the same constant rate per unit time π, over the entire horizon of our problem. This makes $\dot{J}/J = \dot{p}/p = \dot{w}/w = \pi$ for all t, and leaves relative prices and wages constant over time. Furthermore, assume for simplicity that the cost-of-change function is quadratic, so that

$$C(\dot{K}) = \frac{\gamma}{2} \dot{K}^2 \qquad \gamma > 0.$$

On this assumption, equation (2) becomes

$$p\frac{\partial F}{\partial K} - J\delta - rJ + \dot{J} - (rJ-\dot{J})\gamma\dot{K} + J\gamma\ddot{K} = 0.$$

Dividing by J and solving for $\ddot{K}$ gives

(3) $\ddot{K} = \frac{1}{\gamma}[r + \delta - \frac{\dot{J}}{J} - \frac{p}{J}F_K] + (r - \frac{\dot{J}}{J})\dot{K}.$

On our assumptions, since all relative prices are constant over time, and since r and $\dot{J}/J$ are constant over time, the above equation is a fixed coefficient, linear differential equation in $\dot{K}$:

(4) $\frac{d\dot{K}}{dt} = A + B\dot{K}$

where $A = \frac{1}{\gamma}[r + \delta - \pi - \frac{p}{J}F_K]$

$B = (r-\pi) > 0.$

Since w/p is constant over time, so is N/K, making F_K and therefore A constant over time. The differential equation (4) has the solution

(5) $\dot{K} = \alpha e^{Bt} - \frac{A}{B}$

where α is a constant chosen to insure that an initial condition or terminal condition is satisfied. If $\alpha = 0$, then $\dot{K} = - A/B$ for all t.

But if $\alpha \neq 0$, $\dot{K}$ follows an exponential path in which (after a time) the absolute value of investment increases at an exponential rate. (See Figure 2)
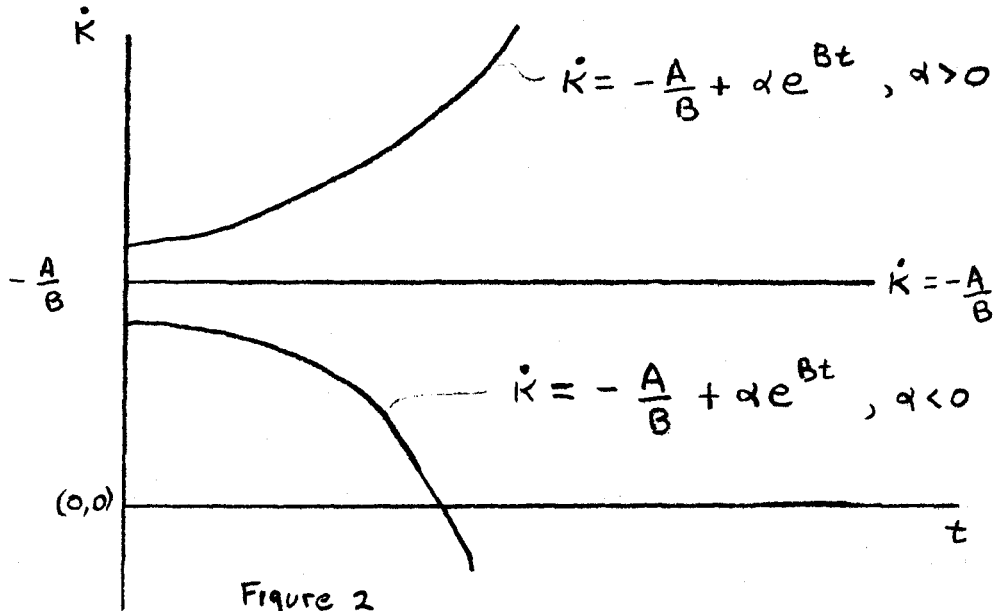

$$\dot{K} = -\frac{A}{B} + \alpha e^{Bt}, \quad \alpha > 0$$

$$\dot{K} = -\frac{A}{B}$$

$$\dot{K} = -\frac{A}{B} + \alpha e^{Bt}, \quad \alpha < 0$$

Figure 2

Given that the time path of $\dot{K}$ is described by equation (5), the time path of K itself is described by

$$(6) \qquad K(t) = \frac{\alpha}{B} e^{Bt} - \frac{A}{B}t + K(0) - \frac{\alpha}{B}.$$

Given expressions (5) and (6), we can compute the firm's discounted net cash flow $f(\ )$ as

$$f(N(t), K(t), \dot{K}(t), t) = e^{-rt}[p\frac{F}{K} - w\frac{N}{K} - J\delta] \cdot K(t)$$

$$-e^{-rt}J\dot{K} - e^{-rt}J\frac{\gamma}{2}\dot{K}^2$$

$$= e^{-(r-\pi)t}D[-\frac{A}{B}t + \frac{\alpha}{B}e^{Bt} + K(0) - \frac{\alpha}{B}]$$

$$- e^{-(r-\pi)t}J(0)(\alpha e^{Bt} - \frac{A}{B}) - e^{-(r-\pi)t}(\alpha e^{Bt} - \frac{A}{B})^2 \cdot \frac{\gamma}{2}J(0)$$

where $D = [p(0)F/K - w(0)N/K - J(0)\delta]$ and $p(t) = e^{\pi t}p(0)$,

$w(t) = e^{\pi t}w(0)$, and $J(t) = e^{\pi t}J(0)$. Noting that $B = (r-\pi)$, we have

$$f(\ ) = e^{-Bt}D[-\frac{A}{B}t + \frac{\alpha}{B}e^{Bt} + K(0) - \frac{\alpha}{B}] - e^{-Bt}J(0)(\alpha e^{Bt} - \frac{A}{B})$$

$$- e^{-Bt}(\alpha^2 e^{2Bt} - 2\alpha e^{Bt} + (\frac{A}{B})^2)\frac{\gamma}{2}J(0)$$

$$f(N(t),K(t),\dot{K}(t),t) = D\{-\frac{A}{B}te^{-Bt} + \frac{\alpha}{B} + K(0)e^{-Bt}\}$$

$$- J(0)\{\alpha - \frac{A}{B}e^{-Bt}\}$$

$$- \frac{\gamma}{2}\{\alpha e^{Bt} - 2\alpha + (\frac{A}{B})^2 e^{-Bt}\}J(0).$$

As $t \to \infty$, the first term in braces approaches α/B; the second term in braces approaches α, while the third term in braces approaches ∞ unless $\alpha = 0$, in which case it approaches zero. These calculations imply that as t becomes large, the discounted net cash flow $f(\ )$ becomes a larger and larger negative number (since the last term in braces in multiplied by $-\gamma/2$), unless $\alpha = 0$. This occurs because the rate of investment is increasing approximately exponentially, causing costs of adjustment to rise at an even faster exponential rate. These costs of adjustment become so large eventually that they swamp the firm's revenue, and lead to large negative net returns. For large enough T, that will make present value a very large negative number. Clearly, such paths are not optimal ones for the firm to follow, even though they satisfy the Euler equations. To rule out such paths, the condition $\alpha = 0$ must be met, implying that the pertinent solution of our differential equation (4) is

$$\dot{K} = -\frac{A}{B}$$

or

$$(7) \qquad \dot{K} = \frac{1}{\gamma}\left[\frac{\frac{p}{J}F_K - (r+\delta-\pi)}{r - \pi}\right],$$

Notice that in the context of a one-sector model, p=J, so that (7) is a version of our Keynesian investment schedule

$$\dot{K} = I(q-1) \qquad I' > 0$$

where $q = (F_K - (r+\delta-\pi))/(r-\pi) + 1$.

Derivations of the Keynesian investment schedule along the lines sketched above seem to provide the most satisfactory theoretical foundations yet laid down for that schedule. The theory obviously depends critically on the assumption that costs of adjustment increase at an increasing rate with the absolute value of investment $(C'' > 0)$. The arbitrary nature of that critical assumption explains why some economists are uneasy with the notion of the investment schedule.

An Alternative Derivation

It is interesting that the Keynesian investment schedule can be derived in a slightly different way in the context of the preceding setup. The real present value of the firm, measured in consumption goods, is

$$RPV(0) = \int_0^T e^{-rt}[F(N,K) - \frac{w}{p}N - \delta\frac{J}{p}K - \frac{J}{p}C(\dot{K}) - \frac{J}{p}\dot{K}]dt + \frac{S(T)K(T)e^{-rT}}{p(T)}$$

which for $0 < \epsilon < T$ can be written as

$$RPV(0) = \int_0^\epsilon e^{-rt}[F(N,K) - \frac{w}{p}N - \delta\frac{J}{p}K - \frac{J}{p}\dot{K} - \frac{J}{p}C(\dot{K})]dt + e^{-r\epsilon}RPV(\epsilon).$$

For small ϵ, the above equality can be well approximated as

$$(8) \qquad RPV(0) = \epsilon[F(N,K) - \frac{w}{p}N - \delta\frac{J}{p}K - \frac{J}{p}\dot{K} - \frac{J}{p}C(\dot{K})] + e^{-r\epsilon}RPV(\epsilon),$$

where the term in brackets is evaluated at some point $0 \leq t < \epsilon$, say t=0.

We interpret RPV as the (consumption goods) value of the firm in the stock market. In making its investment plans, the firm is assumed to attempt to maximize the (discounted) future value of present stock-holders' equity in the company. For moment $\varepsilon > 0$, the discounted future value of the company is $e^{-r\varepsilon}RPV(\varepsilon)$. But part of this total will be owned by stockholders who finance whatever new capital the firm buys between time 0 and ε by purchasing equity newly issued by the firm. Between time 0 and ε the firm will purchase capital goods in the (approximate) amount $\varepsilon I(0)$, and will have to issue new equities valued (in consumption goods) at

$$\varepsilon[\frac{J}{p}I(0)+\frac{J}{p}C(I(0))]$$

in order to invest at this rate. (Remember that the cost of investing at the rate $\dot{K}$ is $J\dot{K}+JC(\dot{K})$.) So the discounted future value of the equity of present owners of the firm is given by

$$e^{-r\varepsilon}RPV(\varepsilon) - \varepsilon[\frac{J}{p}I(0)+\frac{J}{p}C(I(0))],$$

which it is convenient to write as

(9) $\qquad e^{-r\varepsilon}\frac{RPV(\varepsilon)}{K(\varepsilon)} [K(0)+\varepsilon I(0)] - \varepsilon[\frac{J}{p}I(0)+\frac{J}{p}C(I(0))]$

We view $RPV(\varepsilon)/K(\varepsilon)$ as a constant that is given to the firm, and that is independent of $I(0)$. The firm maximizes (9) with respect to $I(0)$, which requires the first-order condition

$$e^{-r\varepsilon}\frac{RPV(\varepsilon)}{K(\varepsilon)} \varepsilon - \varepsilon[\frac{J}{p}+\frac{J}{p}C'(I(0))] = 0$$

or

(10) $\qquad e^{-r\epsilon}\dfrac{RPV(\epsilon)}{K(\epsilon)} = \dfrac{J}{p}(1+C'(I(0)))$.

Equation (10) says that investment is pushed to the point at which the cost of additional investment measured in consumption goods equals the discounted value (in consumption goods) that the stock market will place on the firm's capital. Substituting (10) into (8) gives

(11) $\qquad RPV(0) = \epsilon[F(N,K) - \dfrac{w}{p}N - \delta\dfrac{J}{p}K - \dfrac{\dot{J}}{p}K - \dfrac{J}{p}C(\dot{K})] + \dfrac{J}{p}(1+C'(I(0)))(K(\epsilon))$

Taking the limit of (11) as $\epsilon \to 0$ gives

$\qquad RPV(0) = \dfrac{J}{p}(1+C'(I(0)))K(0)$

or

$\qquad \dfrac{\dfrac{p}{J}RPV(0)}{K(0)} - 1 = C'(I(0))$

Since $C'' > 0$, the above equation can be inverted to yield

(12) $\qquad I(0) = g(\dfrac{\dfrac{p}{J}RPV(0)}{K(0)} - 1) \qquad g' > 0,$

which is a version of the Keynesian investment schedule. If p=J, (12) becomes the one-sector Keynesian investment schedule

$\qquad I(0) = g(q-1)$

where q=RPV/K. More generally, it is straightforward to show that on our earlier assumption about the time paths of J, p, w, and r,

$\qquad \dfrac{p}{J} \dfrac{RPV}{K} = \dfrac{\dfrac{p}{J}F_K - (r+\delta-\pi)}{r-\pi} + 1,$

so that (12) is equivalent with (7).

For this setup, then, we have shown that by maximizing the (discounted) expected future value of the equity of current stockholders in the company, the firm in effect maximizes the present value of the firm.

Appendix: Heuristic Explanation of the Euler Equations

We are interested in choosing time paths of $x(t)$ and $y(t)$, $t \in [0,T]$ to obtain the extremum of the functional

(1) $J(x,y,\dot{y},t) = \int_0^T f(x(t), y(t), \dot{y}(t))dt.$

Among the necessary conditions for J to obtain an extremum are the Euler equations:

$$\frac{\partial f}{\partial x(t)} = 0 \qquad\qquad t \in [0,T]$$

$$\frac{\partial f}{\partial y} - \frac{d}{dt} \frac{\partial f}{\partial \dot{y}} = 0 \qquad\qquad t \in [0,T].$$

To motivate these equations, we consider the following discrete time approximation to (1):

(2) $\tilde{J} = \sum_{t=0}^T \varepsilon f(x(t), y(t), \frac{y(t+\varepsilon)-y(t)}{\varepsilon})$

where $t = 0, \varepsilon, 2\varepsilon, \ldots, T-\varepsilon, T$.   Notice that the limit of the sum in equation (2) as ε approaches zero is the integral in equation (1).   We propose to study the first-order conditions for obtaining an extremum of $\tilde{J}$ as ε approaches zero.   It is illuminating to write out several terms of $\tilde{J}$ explicitly:

$$\tilde{J} = \varepsilon f(x(0),y(0), \frac{y(\varepsilon)-y(0)}{\varepsilon}) + \varepsilon f(x(\varepsilon), y(\varepsilon), \frac{y(2\varepsilon)-y(\varepsilon)}{\varepsilon})$$

$$+ \ldots + \varepsilon f(x(n), y(n), \frac{y(n+\varepsilon)-y(n)}{\varepsilon}) + \ldots.$$

Differentiating $\tilde{J}$ with respect to $x(t)$, $t = 0, \varepsilon, \ldots, T$, and setting the derivatives to zero we obtain

(3) $\frac{\partial f}{\partial x(t)} = 0, \ t = 0, \varepsilon, \ldots, T.$

Setting the partial derivative of $\tilde{J}$ with respect to $y(t+\varepsilon)$ equal to zero for $t = 0, \ldots, T-\varepsilon$, we have

$$\frac{\partial \tilde{J}}{\partial y(t+\varepsilon)} = \varepsilon \frac{\partial f(x(t),y(t),\frac{y(t+\varepsilon)-y(t)}{\varepsilon})}{\partial (\frac{y(t+\varepsilon)-y(t)}{\varepsilon})} \cdot \frac{1}{\varepsilon}$$

$$+ \varepsilon \frac{\partial f(x(t+\varepsilon),y(t+\varepsilon),\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon})}{\partial y(t+\varepsilon)}$$

$$- \varepsilon \frac{\partial f(x(t+\varepsilon),y(t+\varepsilon),\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon})}{\partial (\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon})} \cdot \frac{1}{\varepsilon} = 0$$

Dividing by ε and rearranging, we obtain

$$\frac{\partial f(x(t+\varepsilon),y(t+\varepsilon),\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon})}{\partial y(t+\varepsilon)}$$

$$- \frac{1}{\varepsilon}\left\{ \frac{\partial f(x(t+\varepsilon),y(t+\varepsilon),\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon}}{\partial (\frac{y(t+2\varepsilon)-y(t+\varepsilon)}{\varepsilon})} \right.$$

$$\left. - \frac{\partial f(x(t),y(t),\frac{y(t+\varepsilon)-y(t)}{\varepsilon})}{\partial (\frac{y(t+\varepsilon)-y(t)}{\varepsilon})} \right\} = 0.$$

Taking limits as ε goes to zero we have

$$(4) \qquad \frac{\partial f}{\partial y(t)} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}(t)} = 0 \qquad \text{for } t\varepsilon[0,\bar{r}].$$

Equations (3) and (4) are the "Euler equations" associated with our continuous-time extremum problem. It is not surprising that the limits of the discrete time marginal conditions coincide with continuous time Euler equations, since the limit of our discrete time $\tilde{J}$ equals the continuous time J.

Notes on the Consumption Function

The literature on the consumption function is primarily addressed
to explaining three empirical findings that emerged from early attempts
to fit to actual data the simple linear Keynesian consumption function.

(1) $C = a + bY$.

For <u>cross-sections</u> where the data on C and Y correspond to n observations
on the consumption and income of n households over some short period of
time, estimates of (1) typically are characterized by a >0, so that the
average propensity to consume (APC) exceeds the marginal propensity to
consume. Similarly, for aggregate <u>time series</u> regressions, where the
data on C and Y are economy-wide total consumption and income over a
year, estimates of (1) reveal a > 0 and an APC $>$ MPC. For example, for
annual data for the U.S. for the period 1929-1941, where C is consumption
expenditures and Y is disposable income, Ackley[1]/ (p. 225) reports the
estimated consumption function

$$C_t = 26.5 + .75Y_t.$$

As against the above findings, however, data assembled by Kuznets
that extended over the period 1869-1938 and that consisted of (overlapping)
ten-year averages of data on aggregate consumption and aggregate disposable
income, gave estimates of (1) with b of about .9 and a of about zero.
These data were interpreted as indicating that in the very long run,
APC = MPC, and consumption is proportional to income.

The tasks of the literature on the consumption function have mainly been:

a) to reconcile the disparity between the time series regressions fitted
over short periods, which have APC $>$ MPC, with the proportional (APC = MPC)

consumption schedules estimated using Kuznets' data over very long periods of time; and

b) to reconcile the difference between the cross-section regressions that portray APC > MPC with the implications of Kuznets' data.

These notes describe aspects of Milton Friedman's celebrated explanation of these empirical paradoxes.[2/] The treatment here is compatible with Friedman's work, but at some points deviates from being a simple reproduction of it.

The foundation of Friedman's theory is the hypothesis that essentially consumption is proportional to income, measured appropriately. Whether or not the proportionality of consumption and income in Kuznets' data is evidence for that hypothesis is something we shall discuss presently. (Actually, though, Kuznets' data have often been interpreted as lending support to the hypothesis that the true long-run relationship between consumption and income is a proportional one.)

Friedman began with Irving Fisher's theory about consumers' saving. Following Fisher, he posited that the representative household seeks to maximize utility U, where

$$U = U(C_0, C_1, \ldots, C_n)$$

and U( ) satisfies $U_i > 0$, and is strictly concave; C_i is the household's consumption in period i. The household is assumed to be able to borrow or lend all it desires for i periods at the i-period market determined interest rate R_i. The household is then supposed to maximize U( ) subject to the constraint

$$C_0 + \sum_{i=1}^{n} \frac{C_i}{(1+R_i)^i} = Y_0 + \sum_{i=1}^{n} \frac{Y_i}{(1+R_i)^i}$$

where Y_i is the household's income in period i; the constraint thus states that the present value of the household's consumption program must equal the present value of its income stream, i.e., its wealth.

From the assumption that utility is homothetic in consumption at different points in time, Friedman deduced that current consumption is proportional to wealth, the factor of proportionality k depending on the interest rate, among other things.

$$(2) \qquad C = k(\quad)W$$

where $W = Y_0 + \sum_{i=1}^{n} \frac{Y_i}{(1+R_i)^i}$. For several good reasons, Friedman chose to develop the model by at this point introducing the concept of permanent income, which can be defined as the average rate of income that the consumer expects to receive over the rest of his life. Like wealth or present value, permanent income is thus a concept that collapses a stream over time of prospective income into a single summary measure. Permanent income then takes the place of wealth in (2), which is modified to become

$$(3) \qquad C = \beta(\quad)Y_p$$

To make (3) tractable for the purposes of empirical implementation, the dependence of β on the rate of interest and its other determinants is ignored, at least for analyzing the questions to be discussed here, though for other questions the dependence of β on various variables played an important part in Friedman's analysis.

The Cross Section Data

For cross sections, Friedman proposed the model

(4) $\qquad C_i = \beta Y_{pi} + u_i$

(5) $\qquad Y_i = Y_{pi} + Y_{Ti} \qquad i=1,\ldots,n$

Here C_i is measured consumption of th i-th household, Y_i is measured income of the i-th household, while Y_{pi} is the permanent income and Y_{Ti} the transitory income of the i-th household; u_i is the nonsystematic or transitory part of the i-th household's consumption. Friedman assumed that u_i and Y_{Ti} are statistically independent random variables, both with means of zero, and with variances var u and var Y_T, respectively. In addition, u_i and Y_{Ti} are both assumed to be distributed independently of Y_{pi}, so that $E[u_i|Y_{pi}] = E[Y_{Ti}|Y_{pi}] = 0$. To ease the presentation, we will here assume that u_i, Y_{Ti}, and Y_{pi} follow a trivariate normal distribution. The role of this assumption is to assure the linearity of certain regressions that will be described below.

Solving (5) for Y_{pi} and substituting into (4) we find the following relationship between C_i and Y_i:

(6) $\qquad C_i = \beta Y_i + u_i - \beta Y_{Ti}.$

The composite error term $u_i - \beta Y_i$ is correlated with Y_i, since

$$\operatorname{cov}[Y_i, u_i - \beta Y_{Ti}] = E[Y_i \cdot (u_i - \beta Y_{Ti})] = E[(Y_{pi} + Y_{Ti}) \cdot (u_i - \beta Y_{Ti})]$$

$$= EY_{pi}u_i + EY_{Ti}u_i + \beta EY_{Ti}Y_{pi} - \beta EY_{Ti}^2$$

$$= -\beta EY_T^2 = -\beta \operatorname{var} Y_T.$$

The correlation between the composite error term $u_i - \beta Y_{Ti}$ and Y_i means that the assumption required to guarantee the unbiasedness of the least squares estimate of β (and also the intercept) is violated. Moreover, the population regression of C_i on Y_i (i.e., $E(C_i|Y_i)$) does not equal βY_i, since taking expectations in (4) and (6) conditional on Y_i shows

$$(7) \qquad E(C_i|Y_i) = \beta EY_{pi}|Y_i = \beta Y_i - \beta EY_{Ti}|Y_i.$$

The assumption that Y_{pi}, Y_{Ti}, and u_i are distributed according to a tri-variate normal distribution guarantees that the above regressions are linear. We need to evaluate $EY_{pi}|Y_i$. Since Y_{pi}, Y_{Ti} follow a bivariate normal distribution, we know

$$(8) \qquad E(Y_{pi}-EY_p|Y_i-EY) = \theta(Y_i-EY)$$

where
$$\theta = \frac{\text{cov}(Y_{pi}-EY_p, Y_i-EY)}{\text{var } Y_i}$$

But we have

$$\text{cov}(Y_{pi}-EY_p, Y_i-EY) = E((Y_{pi}-EY_p)(Y_i-EY_i))$$

$$= E(Y_{pi}-EY_p)(Y_{Ti}+(Y_{pi}-EY_p))$$

$$= E(Y_p-EY_p)^2 = \text{var } Y_p.$$

So we have

$$\theta = \frac{\text{var } Y_p}{\text{var } Y},$$

where from (5) we have that var $Y = $ var Y_p + var Y_T.

Rearranging (8), noting since $EY_T = 0$ that $EY = EY_p$, we have

(9) $\qquad E(Y_{pi}|Y_i) = \dfrac{\text{var } Y_p}{\text{var } Y} \cdot Y_i + (1 - \dfrac{\text{var } Y_p}{\text{var } Y}) EY.$

Notice that $1 - \dfrac{\text{var } Y_p}{\text{var } Y} = \dfrac{\text{var } Y_T}{\text{var } Y}.$

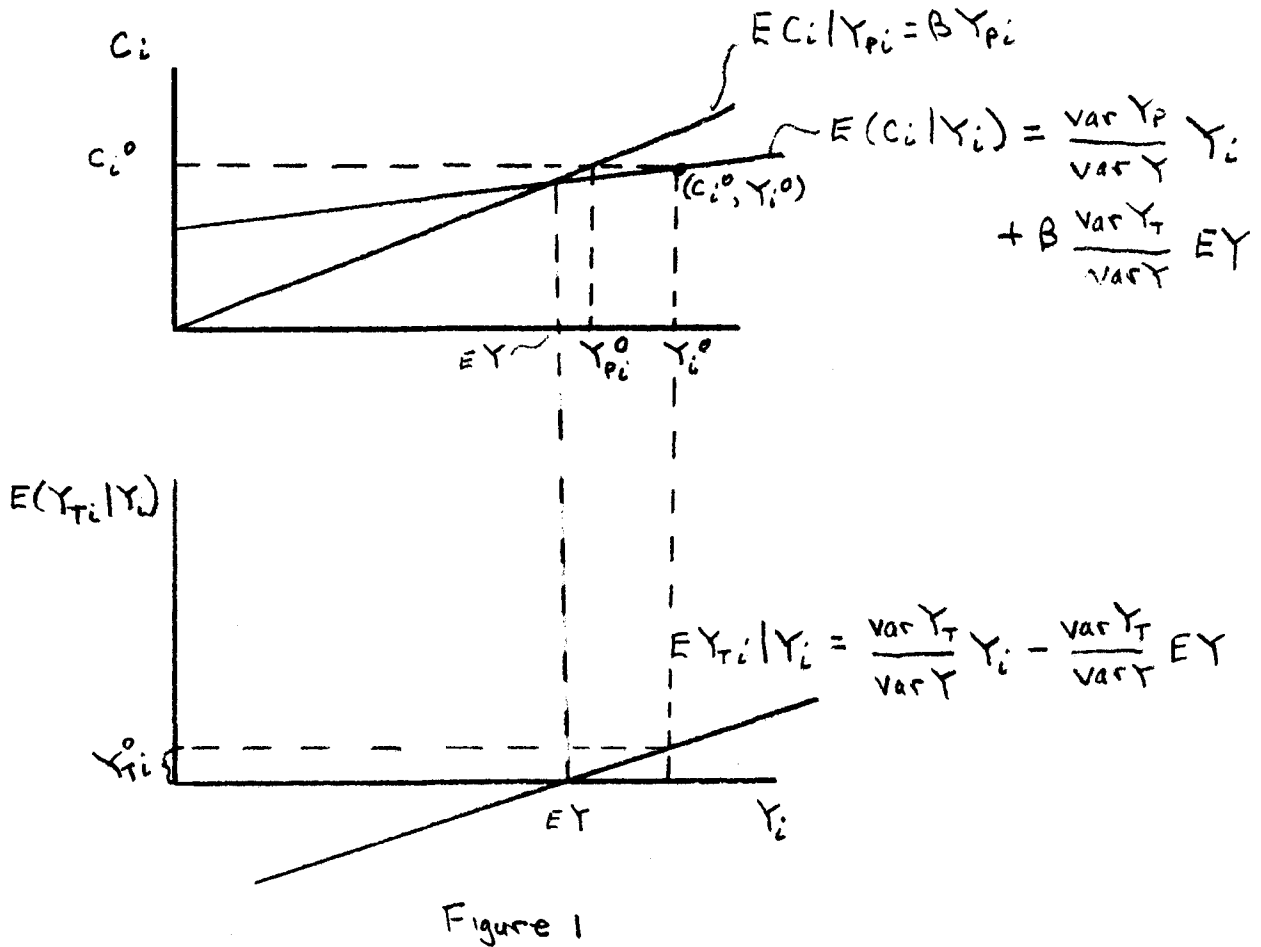Now subsituting equation (9) into (7) we have

(10) $\qquad E(C_i|Y_i) = \beta \dfrac{\text{var } Y_p}{\text{var } Y} Y_i + \dfrac{\text{var } Y_T}{\text{var } Y} \beta EY.$

The regression of C_i on Y_i is linear with slope $\beta \dfrac{\text{var } Y_p}{\text{var } Y}$ and intercept $\dfrac{\text{var } Y_T \beta E(Y)}{\text{var } Y}$. So long as the variance of transitory income is positive, the regression of C_i on Y_i has a positive intercept and a less steep slope than does the regression of C_i on Y_{pi}. The regression of C_i on Y_i is thus not the same as that of C_i on Y_{pi}, the regression $E(C_i|Y_i)$ possessing a lower apparent marginal propensity to consume.

The situation is depicted in figure 1. A consumer with measured income Y_i will on average have transitory income

$$EY_{Ti}|Y_i = \dfrac{\text{var } Y_T}{\text{var } Y} Y_i - \dfrac{\text{var } Y_T}{\text{var } Y} EY.$$

On average, then, such a consumer's permanent income will be $Y_i - EY_{Ti}|Y_i$. The consumer will on average consume at the rate $\beta(Y_i - EY_{Ti}|Y_i)$. The bottom panel depicts $EY_{Ti}|Y_i$ as a linear function going through the point $Y_i = EY$, $Y_{Ti} = 0$, since means of both variables lie on the regression line. The slope of the regression line is positive, since measured income increases when transitory income does. On average, people with higher than average measured income have higher than average transitory income, a consequence of transitory income being uncorrelated with permanent income. Consumers with measured income $Y_i^0 > E(Y)$ on average have transitory income $Y_{Ti}^0 > 0$, so that they consume $C_i^0 = \beta(Y_i^0 - Y_{Ti}^0) = \beta Y_{pi}^0$.

$$E C_i | Y_{pi} = \beta Y_{pi}$$

$$E(C_i | Y_i) = \frac{var\, Y_P}{var\, Y} Y_i + \beta \frac{var\, Y_T}{var\, Y} EY$$

$$E Y_{Ti} | Y_i = \frac{var\, Y_T}{var\, Y} Y_i - \frac{var\, Y_T}{var\, Y} EY$$

Figure I

Consequently the observation (c_i^0, Y_i^0) lies below the "true" consumption function that relates consumption to permanent income. On the other hand, if $Y_i < EY$, then on average transitory income is negative, meaning that measured income on average understates permanent income. The result is that for observations with $EY > Y_i$, observations on consumption and measured income will on average lie <u>above</u> the consumption function connecting the permanent magnitudes. The result, then, will be to flatten out the consumption function relating consumption to measured income.

So far, our entire discussion has been about the population parameters characterizing $E(C_i | Y_i)$ and $E(C_i | Y_{pi})$, rather than about the lease squares

estimates of those parameters on the basis of particular samples. For samples of the very large size used in cross-section studies, however, the probability is very high that the relations connecting population parameters will provide a very good description of the least-squares parameter estimates. There is nothing to be gained by going into the details of least squares estimation at this point.

The Time Series

For the time series, Friedman posits

$$(11) \qquad C_t = \beta Y_{pt} + u_t \qquad t=1,\ldots,T$$

where Y_{pt} is permanent income at time t, β is the marginal propensity to consume out of permanent income, and u_t is a random disturbance term distributed independently of past, present, and future Y_{pt}, and having mean zero and a finite variance. To make (11) operational, since Y_{pt} is not observable, requires a statistical model linking permanent income to measured income. A model that provides this link while at the same time providing a fairly deep rationalization for the concept of "permanent" income is the following one due to John F. Muth:[3]

$$(12) \qquad Y_{pt} = Y_{pt-1} + \varepsilon_t + k$$

$$(13) \qquad Y_t = Y_{pt} + Y_{Tt},$$

Here Y_t is measured income, while Y_{Tt} is transitory income. The variable ε_t is a serially uncorrelated, identically distributed random variable with mean zero and finite variance, k is a constant representing the trend rate of growth of permanent income, and transitory income, Y_{Tt} is a serially uncorrelated, identically distributed random variable with mean zero and finite variance; Y_{Tt} and ε_t are independently distributed. We again assume that all random variables follow normal distributions. According to (12), permanent income Y_{pt} follows a "random walk" with trend k. Adding Y_{Tt} to both sides of (12) gives

$$(Y_{pt}+Y_{Tt}) = (Y_{pt-1}+Y_{Tt-1}) + (Y_{Tt}-Y_{Tt-1}) + \varepsilon_t$$

(14) $Y_t = Y_{t-1} + (Y_{Tt} - Y_{Tt-1}) + \varepsilon_t + k.$

The random term in (14), which is $Y_{Tt} - Y_{Tt-1} + \varepsilon_t$ has variance

$$E[Y_{Tt} - Y_{Tt-1} + \varepsilon_t]^2 = 2 \text{ var } Y_T + \text{var } \varepsilon ;$$

the "autocovariances" of this random term are

$$E[Y_{Tt} - Y_{Tt-1} + \varepsilon_t, \ Y_{Tt-i} - Y_{Tt-i-1} + \varepsilon_{t-i}]$$

$$= \begin{cases} - \text{var } Y_T & \text{for } i = \pm 1, \\ \\ 0 & \text{for } |i| > 1. \end{cases}$$

The random term thus displays negative first-order serial correlation.
The distribution of the composite random variable $(Y_{Tt} - Y_{Tt-1}) + \varepsilon_t$ can be completely characterized by its covariogram, which is simply the covariance of the random term with itself lagged i times, where $i = -\infty, \ldots, -2, -1, 0, 1, 2, \ldots.$
We have established that the covariogram of the composite error is

$$\text{cov}(Y_{Tt} - Y_{Tt-1} + \varepsilon_t, \ Y_{Tt-i} - Y_{Tt-i-1} + \varepsilon_{t-i})$$

$$= \begin{cases} 2 \text{ var } Y_T + \text{var } \varepsilon, & i = 0 \\ \\ - \text{var } Y_T & , \ i = 1, -1 \\ \\ 0 & , \ |i| \neq 1, 0. \end{cases}$$
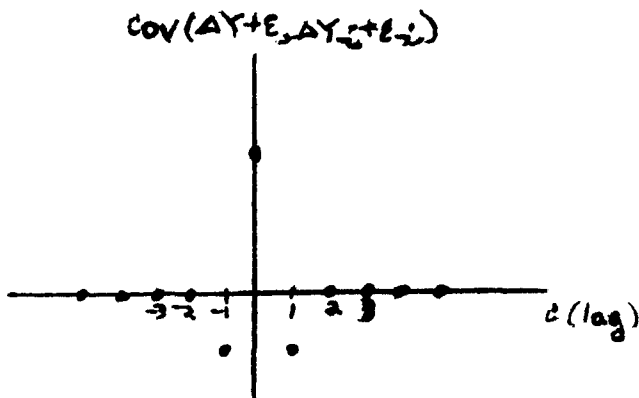
figure 2

The covariogram is depicted in figure 2.

It is now convenient to replace the composite error term $Y_{Tt} - Y_{Tt-1} + \varepsilon_t$ by a single, equivalent random variable that equals and has the same probability distribution as the composite error. To have the same probability distribution, it must have mean zero and a covariogram identical with that of the composite error. We propose to take

$$(15) \qquad W_t - aW_{t-1} = Y_{Tt} - Y_{Tt-1} + \varepsilon_t,$$

where W_t is a serially uncorrelated random variable with mean zero and variance var W. The variance of $W_t - aW_{t-1}$ is var $W(1+a^2)$; the covariance between $W_t - aW_{t-1}$ and $W_{t-1} - aW_{t-2}$ is $-a$ var W. All other lagged covariances are zero. So that the covariograms of the composite error and $W_t - aW_{t-1}$ are equal, we thus require

$$2 \text{ var } Y_T + \text{var } \varepsilon = \text{var } W(1+a^2)$$

$$- \text{var } Y_T = - a \text{ var } W.$$

These are two equations that can be used to solve for a and var W as functions of var ε and var Y_T. The solution for a turns out to be

$$(16) \qquad a = 1 + \frac{1}{2}\left(\frac{\text{var } \varepsilon}{\text{var } Y_T}\right) - \sqrt{\frac{\text{var } \varepsilon}{\text{var } Y_T}\left(1 + \frac{1}{4}\frac{\text{var } \varepsilon}{\text{var } Y_T}\right)} \quad,$$

where a will obey $0 \leq a \leq 1$.

Rewriting (14) in terms of W_t, we have

$$(17) \qquad Y_t = Y_{t-1} + W_t - aW_{t-1} + k.$$

The model can be rewritten in a revealing way by calculating the geometric

sum $\sum_{i=0}^{\infty} a^i Y_{t-i}$:

$$\sum_{i=0}^{\infty} a^i Y_{t-i} = \sum_{i=0}^{\infty} a^i Y_{t-i-1} + \sum_{i=0}^{\infty} a^i W_{t-i} - a \sum_{i=0}^{\infty} a^i W_{t-i-1}.$$

This can be rewritten as

$$Y_t + \sum_{i=1}^{\infty} a^i Y_{t-i} = \sum_{i=1}^{\infty} a^{i-1} Y_{t-i} + W_t + \sum_{i=1}^{\infty} a^i W_{t-i} - \sum_{i=1}^{\infty} a^i W_{t-i} + \frac{k}{1-a}$$

which in turn can be written

$$Y_t = (1-a) \sum_{i=1}^{\infty} a^{i-1} Y_{t-i} + W_t + \frac{k}{1-a}$$

To get an expression for permanent income, we simply substract transitory

income Y_{Tt} from both sides to obtain

$$Y_{pt} = (1-a) \sum_{i=1}^{\infty} a^{i-1} Y_{t-1} + W_t - Y_{Tt} + \frac{k}{1-a} .$$

By virtue of equation (13), we also have

$$(18) \qquad Y_{pt-1} = (1-a) \sum_{i=1}^{\infty} a^{i-1} Y_{t-i} + W_t - Y_{Tt} - \varepsilon_t + \frac{ka}{1-a}$$

To obtain the best estimate of permanent income Y_{pt-1} on the basis of

variables observed up through time t-1, we replace the random variables

W_t, Y_{Tt}, and ε_t by their (conditional) means of zero to obtain

$$\hat{Y}_{pt-1} = (1-a) \sum_{i=1}^{\infty} a^{i-1} Y_{t-i} + \frac{ka}{1-a} .$$

Thus, the best estimate of permanent income is a constant plus a "dis-

tributed lag" of current and past measured incomes, the weights in the

distributed lag declining geometrically with decay parameter a. As expression (16) above shows, the parameter a depends on the ratio of the variance of transitory income Y_T to the variance of ε, which is the innovation or new part of past income.

Substituting (18) into (11) we now obtain the following schedule relating measured income and consumption,

$$C_t = \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{t-i} + \beta W_{t+1} - \beta Y_{Tt+1} - \beta\varepsilon_{t+1} + \beta\frac{k}{(1-a)} + u_t$$

or

$$(19) \qquad C_t = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{t-i} + u_t^1$$

where $u_t^1 = \beta W_{t+1} - \beta Y_{Tt+1} - \beta\varepsilon_{t+1} + u_t$. It is easily verified that u_t^1 is a serially uncorrelated random variable with mean zero and a finite variance.

An alternative and useful way of writing (19) is found by first writing down a C_{t-1},

$$aC_{t-1} = \frac{a\beta k}{(1-a)} + \beta(1-a) \sum_{i=1}^{\infty} a^i Y_{t-i} + au_{t-1}^1 .$$

Subtracting the above expression from (19) then gives

$$C_t - aC_{t-1} = \beta k + \beta(1-a)Y_t + u_t^1 - au_{t-1}^1$$

or

$$(20) \qquad C_t = \beta k + aC_{t-1} + \beta(1-a)Y_t + u_t^1 - au_{t-1}^1 .$$

Equations (19) and (20) are alternative but equivalent ways of expressing

the relationship between consumption and measured income.

Now consider two alternative paths of measured income, $\{Y_t\}$,

$t = -\infty, \ldots, -1, 0, 1, 2, \ldots, \infty$ and

$$\{\tilde{Y}_t\} = \begin{cases} Y_t & \text{for } t < 0 \\ Y_t + \Delta & \text{for } t \geq 0. \end{cases}$$

Suppose that u_t^1 were to remain the same under these two alternative paths

for measured income. For the path $\{Y_t\}$, the path of consumption $\{C_t\}$ is

given by equation (19). The path of consumption corresponding to $\{\tilde{Y}_t\}$,

denoted $\{\tilde{C}_t\}$, found by substituting the $\{\tilde{Y}_t\}$ path into (19), is given by

$$\tilde{C}_t = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{t-i} + u_t^1, \quad t < 0.$$

$$\tilde{C}_0 = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{0-i} + u_0^1 + \beta(1-a)\Delta$$

$$\tilde{C}_1 = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{1-i} + u_1^1 + \beta(1-a)\Delta + \beta(1-a)a\Delta$$

.

.

.

$$\tilde{C}_j = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{j-i} + u_j^1 + \beta(1-a)\Delta(1+a+a^2+\ldots+a^j),$$

$$\text{for } j \geq 0.$$

Subtracting the path $\{\tilde{C}_t\}$ from the path $\{C_t\}$ gives

$$C_t - \tilde{C}_t = \begin{cases} 0 & \text{for } t < 0 \\ \beta(1-a)(1+a+a^2+\ldots+a^t)\Delta & \text{for } t \geq 0. \end{cases}$$

We have in effect calculated the time pattern of the response of C to a once-and-for-all jump in measured income by Δ. In the period that the once-and-for-all jump in income occurs, consumption rises by $\beta(1-a)\Delta$, while in the subsequent period consumption jumps an additional $\beta(1-a)a\Delta$, while in the next period consumption jumps an additional $\beta(1-a)a^2\Delta$, and so on. The eventual response of consumption to the jump in income is

$$\beta(1-a)(1+a+a^2+\ldots+a^n+\ldots)$$

$$= \beta(1-a)\cdot\frac{1}{1-a}$$

$$= \beta,$$

so that β is the long-run marginal propensity to consume out of income, or as before, the marginal propensity to consume out of permanent income. The "short-run marginal propensity to consume" is $\beta(1-a)$, which is less than β since $a < 1$.

The fact that the long-run marginal propensity to consume exceeds the short-run marginal propensity to consume in (19) does not in itself reconcile the difference between time series regressions of C_t on Y_t over the short and long periods of time; and neither does it imply that the regressions of C_t on Y_t over very long periods of time will tend to capture the long-run marginal propensity to consume as the coefficient on Y_t. In fact, the regression of C_t on Y_t that emerges in the data, given a consumption function like Friedman's (19), depends sensitively on how income Y_t is moving over time. If the nature of the evolution over time of Y_t changes, then the regression of C_t on Y_t will change,

even if we assume that (19) remains valid. For example, if income follows a trend

$$Y_t = \alpha t,$$

then it can be shown that the population regression of C_t on Y_t is

$$E[C_t|Y_t] = \xi + \beta Y_t, \text{ where } \xi \text{ is a constant.}$$

To show this, we note that $Y_t = \alpha t$ implies that

$$Y_{t-i} = Y_t - i\alpha.$$

Substituting this expression for Y_t into (19) gives

$$C_t = \beta\frac{k}{(1-a)} + \beta(1-a) \sum_{i=0}^{\infty} a^i (Y_t - i\alpha) + u_t^1$$

$$= \beta\frac{k}{(1-a)} + \beta Y_t + u_t^1 - \qquad - \beta(1-a)\alpha \sum_{i=0}^{\infty} a^i i.$$

Here the regression of C_t on Y_t has β as the coefficient on Y_t. On the other hand, if Y_t is assumed to follow the process

$$Y_t = A \cos (2\pi f t)$$

where A is a constant and f is frequency (cycles per unit time), then it is possible to show that where consumption is governed by (19), the regression of C_t on Y_t will have a slope less than β.

Notice, however, that we are really not free to assume any arbitrary process for Y_t. For in deriving (19) we have made certain explicit assumptions about the way Y_t is evolving and the way people are assumed to be forecasting Y_t in the future. In particular, we assumed that income is evolving according to (17) and that in forecasting Y_t (i.e., in estimating permanent income) consumers are using their

knowledge of the parameters of (17) to construct optimal forecasts. Our rationalization for the geometric distributed lag in (19) thus depends on (17) being the actual process for Y. Presumably, if some other process instead described Y, consumers would not form their estimate of permanent income in the fashion assumed in (19), since that would not yield the best forecasts of subsequent income. So if income no longer were described by (17), it is unlikely that consumption would long be described by (19), since (19) then incorporates a suboptimal formula for extracting forecasts of income.[4/]

The upshot of all of this is that the most interesting assumption about the evolution of Y_t to couple with (19) is our equation (17). Here we propose to derive the regression between C_t and Y_t that prevails when both (17) and (19) hold. To do this in the easiest way, we begin by differencing equation (19) to arrive at

$$\Delta C_t = \beta(1-a) \sum_{i=0}^{\infty} a^i \Delta Y_{t-i} + \Delta u_t^1$$

or

$$(21) \qquad \Delta C_t = \beta(1-a)\Delta Y_t + \beta(1-a) \sum_{i=1}^{\infty} \Delta a^i Y_{t-i} + \Delta u_t^1$$

where $\Delta C_t = C_t - C_{t-1}$, etc. In addition we have from (17),

$$(22) \qquad \Delta Y_t = k + W_t - aW_{t-1}.$$

The parameters linking the differences of C and Y in (21) are obviously the same as the parameters of the equation linking their levels. Now suppose that instead of estimating equation (21), we were to estimate the simple Keynesian consumption function by regressing ΔC_t against ΔY_t,

i.e., compute γ and δ in the regression

$$E[\Delta C_t | \Delta Y_t] = \gamma \Delta Y_t + \delta$$

Then if (21) is really the correct consumption function, it is possible to show[5] that the population parameters γ and δ obey

$$\gamma = \beta(1-a) + \xi$$

$$\delta = \psi$$

where ξ and ψ are the population regression coefficients in the regression of $\beta(1-a) \sum\limits_{i=1}^{\infty} a^i \Delta Y_{t-i}$ against ΔY_t and one:

$$E[\beta(1-a) \sum_{i=1}^{\infty} a^i \Delta Y_{t-i} | \Delta Y_t] = \psi + \hat{\xi} Y_t.$$

Now by virtue of (22), we know that $\Delta Y_{t-2}, \Delta Y_{t-3}, \ldots$ are statistically independent of ΔY_t, since W_t is serially independent. Then

$$E[\beta(1-a) \sum_{i=2}^{\infty} a^i \Delta Y_{t-i} | \Delta Y_t] = \beta(1-a)\frac{ka}{1-a}.$$

We also have that

$$E[\Delta Y_{t-1} | \Delta Y_t] = \theta \Delta Y_t + (1-\theta)k$$

where

$$\theta = \frac{\text{cov}(\Delta Y_t, \Delta Y_{t-1})}{\text{var} \Delta Y_{t-1}}$$

$$= \frac{-a \text{ var } W_t}{(1+a^2)\text{var } W_t} = \frac{-a}{1+a^2}.$$

Then the regression of $\beta(1-a) \sum_{i=1}^{\infty} a^i \Delta Y_{t-i}$ against ΔY_t is

$$E[\beta(1-a) \sum_{i=1}^{\infty} a^i Y_{t-i} | \Delta Y_t]$$

$$= \beta(1-a) E[\sum_{i=2}^{\infty} a^i Y_{t-i} | \Delta Y_t] + \beta(1-a) a E[\Delta Y_{t-1} | \Delta Y_t, 1]$$

$$= \frac{\beta(1-a)k}{1-a} + (1 + \frac{a}{1+a^2}) k \beta(1-a) a - \frac{\beta(1-a)a^2}{1+a^2} \Delta Y_t.$$

Therefore, we have that

$$\gamma = \beta(1-a) - \frac{\beta(1-a)a^2}{1+a^2}$$

$$= \frac{\beta(1-a)}{1+a^2}$$

$$\delta = \beta k[1 + (\frac{1+a+a^2}{1+a^2})(1-a)].$$

So long as $0 < a < 1$, the usual presumption, then the population value of γ, the slope of the population regression of ΔC_t on ΔY_t (or of C_t on Y_t) is _less_ than both the long-run MPC, β, and the short-run MPC, $\beta(1-a)$. Furthermore, the intercept in the regression of ΔC on ΔY is positive, so long as $k > 0$, meaning that the relationship between the levels of C_t and Y_t has the form[6/]

$$E[C_t | Y_t, t] = \alpha + \delta t + \gamma Y_t,$$

so that the regression of C_t on Y_t shifts up over time at the rate δ.

Recall, however, that Kuznets' data were ten-year averages of consumption and income. The relationship between such averages will of

course not be the same as the relationship between the annual (or quarterly), one-period data that we have just derived. To show the effects of averaging data over time on the relationship between C and Y, let us consider the effects of taking five-period averages of the data where C_t and Y_t are, as before, related by

$$(19) \qquad C_t = \frac{\beta k}{1-a} + \beta(1-a) \sum_{i=0}^{\infty} a^i Y_{t-i} + u_t^1$$

$$(17) \qquad Y_t = Y_{t-1} + W_t - aW_{t-1} + k.$$

We analyze five-period rather than ten-period averages to shorten the computations without sacrificing the essential point.

Given the model formed by equations (19) and (17), we seek the relationships between the five-period sums

$$\overline{C}_t = \sum_{j=0}^{4} C_{t-j}$$

$$\overline{Y}_t = \sum_{j=0}^{4} Y_{t-j} \; .$$

Taking sums on both sides of equations (19) and (17), we have

$$(23) \qquad \sum_{j=0}^{4} C_{t-j} = 5\frac{\beta k}{1-a} + \beta(1-a) \sum_{i=0}^{\infty} a^i \sum_{j=0}^{4} Y_{t-i-j} + \sum_{j=0}^{4} u_{t-j}^1$$

$$(24) \qquad \sum_{j=0}^{4} Y_{t-j} = \sum_{j=0}^{4} Y_{t-j-1} + \sum_{j=0}^{4} W_{t-j} - a\sum_{j=0}^{4} W_{t-j-1} + 5k.$$

Notice that

$$\sum_{j=0}^{4} W_{t-j} - a \sum_{j=0}^{4} W_{t-j-1}$$

$$= W_t + (1-a)W_{t-1} + (1-a)W_{t-2} + (1-a)W_{t-3} + (1-a)W_{t-4} - aW_{t-5}.$$

Then we can write equations (23) and (24) as

$$(25) \qquad \overline{C}_t = \alpha + \beta(1-a) \sum_{i=0}^{\infty} a^i \overline{Y}_{t-i} + \overline{u}_t^1$$

$$(26) \qquad \overline{Y}_t = 5k + \overline{Y}_{t-1} + W_t + (1-a)W_{t-1} + (1-a)W_{t-2}$$

$$+ (1-a)W_{t-3} + (1-a)W_{t-4} - aW_{t-5},$$

where $\alpha = \dfrac{5\beta k}{(1-a)}$ and $\overline{u}_t^1 = \sum_{j=0}^{4} u_{t-j}^1$.

We seek the nature of the simple regression of $\overline{C}_t$ against $\overline{Y}_t$, a constant, and possibly a trend. As before, it is computationally convenient to calculate the regression of $\Delta\overline{C}_t$ on $\Delta\overline{Y}_t$ and a constant, where of course the population value of the coeffcient on $\Delta\overline{Y}_t$ equals the population value of the coefficient on $\overline{Y}_t$ in the regression of $\overline{C}_t$ on $\overline{Y}_t$, a constant, and trend.

First differencing equations (25) and (26), we have

$$\Delta\overline{C}_t = \beta(1-a) \sum_{i=0}^{\infty} a^i \Delta\overline{Y}_{t-i} + \Delta\overline{u}_t^1$$

$$\Delta\overline{Y}_t = 5k + W_t + (1-a)W_{t-1} + (1-a)W_{t-2} + (1-a)W_{t-3}$$

$$+ (1-a)W_{t-4} - aW_{t-5}$$

where $\Delta\overline{C}_t = \overline{C}_t - \overline{C}_{t-1}$, etc. We seek the regression $E[\Delta\overline{C}_t | \Delta\overline{Y}_t]$, which is given by

$$(27) \qquad E[\Delta\overline{C}_t | \Delta\overline{Y}_t] = \left\{ \beta(1-a) \sum_{i=0}^{\infty} a^i \frac{\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-i})}{\text{var } \Delta\overline{Y}_t} \right\} \Delta\overline{Y}_t + \phi$$

where $\phi = \beta(1-a) \sum_{i=0}^{\infty} a^i (1 - \frac{\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-i})}{\text{var } \Delta\overline{Y}_t})$.

We calculate the following covariances

$$\text{var } \Delta\overline{Y}_t = \text{var } W[1+4(1-a)^2+a^2]$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-1}) = [(1-a)+3(1-a)^2-a(1-a)] \text{var } W$$

$$= 4(1-a)^2 \text{var } W$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-2}) = 3(1-a)^2 \text{var } W$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-3}) = 2(1-a)^2 \text{var } W$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-4}) = (1-a)^2 \text{var } W$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta\overline{Y}_{t-5}) = -a \text{ var } W$$

$$\text{cov}(\Delta\overline{Y}_t, \Delta Y_{t-i}) = 0, \qquad i \geq 6.$$

Substituting the formulas for these covariances into (27), we obtain

$$E[\Delta\overline{C}_t | \Delta\overline{Y}_t] = \phi + \Delta\overline{Y}_t \cdot \beta(1-a)[1+4(1-a)^2+a^2+4(1-a)^2 a$$

$$+3a^2(1-a)^2+2a^3(1-a)^2+(1-a)^2 a^4-a^6]/(1+4(1-a)^2+a^2)$$

$$= \phi + \Delta\overline{Y}_t \ (1-a) \cdot \frac{[1+a^2-a^6+(1-a)^2[4+4a+3a^2+2a^3+a^4]]}{1 + 4(1-a)^2 + a^2}$$

Therefore, the regression of $\overline{C}_t$ against $\overline{Y}_t$ and time is

$$(28) \qquad E[\overline{C}_t | \overline{Y}_t, t] = \delta + \phi t + \beta(1-a)\xi\overline{Y}_t,$$

where $\xi = \dfrac{1 + a^2 - a^6 + (1-a)^2[4+4a+3a^2+2a^3+a^4]}{1 + 4(1-a)^2 + a^2}$

and where δ is a parameter.

Equation (28) is the population regression of the five-period sum $\overline{C}_t$ on time and the five-period sum $\overline{Y}_t$. "Good" statistical estimators of and the slope $(\beta(1-a)\xi)$ will have the property of consistency, which means that for a big enough sample size, the probability that the estimate lies within any given interval of the population parameter becomes arbitrarily close to unity.

For the population regression coefficient of $\overline{C}_t$ on $\overline{Y}_t$ to equal the marginal propensity to consume out of disposable income requires that $\beta(1-a)\xi = \beta$, or $\xi = \dfrac{1}{(1-a)}$, an equality that in general does not hold. The value of ξ for various values of a is recorded in Table 1, which also records the associated value of $(1-a)\xi$, which is the bias factor associated with accepting the regression coefficient of $\overline{C}_t$ on $\overline{Y}_t$ as an estimate of the marginal propensity to consume out of permanent income. Obviously there is no presumption that the regression of $\overline{C}_t$ on $\overline{Y}_t$ and time bears a coefficient on $\overline{Y}_t$ that is a good estimate of β.

Table 1

a	$\xi(a)$	$(1-a)\xi(a)$
0	1.0	1.0
.05	1.04	.99
.10	1.08	.97
.20	1.17	.93
.30	1.25	.87
.40	1.31	.78
.50	1.33	.67
.60	1.30	.52
.70	1.19	.36
.80	1.00	.20
.90	.76	.08
1.0	.50	0.00

Table 2 contains the figures pertinent for determining the population regression of ten-period averages of consumption, $\overline{C}_{10,t}$ against ten-period averages of income, $\overline{Y}_{10,t}$. Proceeding in a manner analogous to our treatment of five-period averages, it is possible to show that where one-period income and consumption are again governed by (19) and (17), the regression of $\overline{C}_{10,t}$ on $\overline{Y}_{10,t}$, a constant, and time is

$$E(\overline{C}_{10,t} | \overline{Y}_{10,t}, t) = \delta^1 + \phi^1 t + \beta(1-a)\psi\overline{Y}_{10,t},$$

where δ^1 and ϕ^1 are constants and where ψ is given by

$$\psi(a) = \frac{1 + a^2 - a^{11} + (1-a)^2(\sum_{t=1}^{9} t a^{10-t}+9)}{1 + a^2 + 9(1-a)^2}$$

As Table 2 shows, if a happens to be close to zero, $\psi(a)\beta(1-a)$ is close to β. Friedman estimated that for annual data a $\approx$ .3; for an a of .3,

$\psi(a)(1-a)$ equals .90. That is a substantial bias factor, considering

that the estimates of the MPC using Kuznet's data approximate .9.

However, those regressions included no time trend, which makes comparisons

with our calculations difficult.

Table 2

$\underline{a}$	$\underline{\psi(a)}$	$\underline{(1-a)\,\psi(a)}$
.00	1.00	1.00
.05	1.04	.99
.10	1.09	.98
.20	1.18	.95
.30	1.29	.90
.40	1.40	.84
.50	1.50	.75
.60	1.56	.62
.70	1.52	.46
.80	1.33	.27
.90	.97	.10
1.00	.50	.00

In summary, we have derived the implicit relationship between

the ten-year averaged data that would obtain if Friedman's model were

correct, and if income were evolving in a fashion for which Friedman's

distributed lag consumption function reflected optimal forecasting of

subsequent income by households. While we have been able to show that

the slope of the ten-year average regression of consumption on income

will exceed the slope for the one-year data, there is no presumption

that the regression for the ten-year average data provides a good estimate

of the marginal propensity to consume out of permanent income. Despite

widely held beliefs to the contrary, it seems that the task of reconciling

the empirical paradoxes about the consumption function is not yet completed.

[1] Gordon Ackley, <u>Macroeconmic Theory</u>, MacMillan, 1960.

[2] Milton Friedman, <u>A Theory of the Consumption Function</u>, NBER, 1956.

[3] John F. Muth, "Optimal Properties of Exponentially Weighted Forecasts," <u>Journal of the American Statistical Association</u>, 55, 229-306 (June, 1960).

[4] This is Muth's doctrine of "rational expectations." See John F. Muth, "Rational Expectations and the Theory of Price Movements," <u>Econometrica</u>, July, 1961, pp. 315-335.

[5] This is Theil's specification theorem. See Henri Theil, <u>Principles of Econometrics</u> (Wiley, 1971), pp. 548-49.

[6] Suppose we have a regression of the form
$$\hat{Z}_t = \hat{a} + \hat{b}t + \hat{c} X_t. \qquad t = 1,2,\ldots,T$$
Subtracting $\hat{Z}_{t-1}$ from the above equations gives
$$(\hat{Z}_t - \hat{Z}_{t-1}) = b(t) - b(t-1) + c(X_t - X_{t-1})$$
$$= b + c(X_t - X_{t-1}).$$

Thus, the constant in the regression on first differences equals the coefficient on time in the regression on "levels."

Notes on Behavior Under Uncertainty

We assume that the individual's preferences can be described by a utility function that makes utility depend on the amount of the one good that the individual consumes. So we write

(1)   Utility = U(C)

where C is the amount of the good consumed. We assume that utility is increasing in consumption, $U'(C) > 0$; that the marginal utility of consumption, though positive, decreases with increases in C, $U''(C) < 0$; and that U(C) is bounded for $C \varepsilon [0, \infty)$.

We assume that the individual is making plans for the future, which we collapse to a single date in the future. To incorporate the existence of uncertainty, we assume that there are n mutually exclusive states of the world, indexed by $\theta = 1, 2, \ldots, n$. The individual has a set of subjective probabilities $\pi(\theta) \geq 0$ giving the probability that he assigns to the occurrence of state θ.

An individual's claim to future consumption goods will in general depend on the state of nature that happens to occur. We let $C(\theta)$ denote his consumption if state θ occurs. The individual is assumed to maximize his expected utility

$$v = \sum_{\theta=1}^{n} \pi(\theta)U(C(\theta)),$$

subject to certain constraints.

Suppose that there are only two states of the world, so that n = 2. The individual's expected utility is then

$$v = \pi(1)U(C(1)) + \pi(2)U(C(2)).$$

Along lines of constant expected utility (indifference curves) we have

$$dv = 0 = \pi(1)U'(C(1))dC(1) + \pi(2)U'(C(2))dC(2)$$

which implies that

$$(2) \qquad \left.\frac{dC(2)}{dC(1)}\right|_{dv = 0} = -\frac{\pi(1)U'(C(1))}{\pi(2)U'(C(2))} .$$

Expression (2) gives the slope of the indifference curve. The concavity of the indifference curve is found by differentiating (2):

$$\frac{d^2C(2)}{dC(1)^2} = -\left[\frac{\pi(1)U''(C(1))}{\pi(2)U'(C(2))}\right] + \frac{\pi(1)U'(C(1))U''(C(2))\frac{dC(2)}{dC(1)}}{\pi(2)U'(C(2))^2} > 0$$

The slope of the indifference curves increases as C(1)/C(2) increases, implying that they are convex.

Next notice that for C(1) = C(2), we have

$$(3) \qquad \left.\frac{dC(2)}{dC(1)}\right|_{\substack{dv = 0 \\ C(1) = C(2)}} = -\frac{\pi(1)}{\pi(2)} .$$

Bundles of C(1), C(2) for which C(1) = C(2) correspond to certain claims, since regardless of the state that occurs, the individual is able to consume the same amount. A 45° line through the origin in C(1), C(2) plane thus contains all certain bundles (see Figure 1), so that it is appropriately called the certainty line. Along the certainty line, the
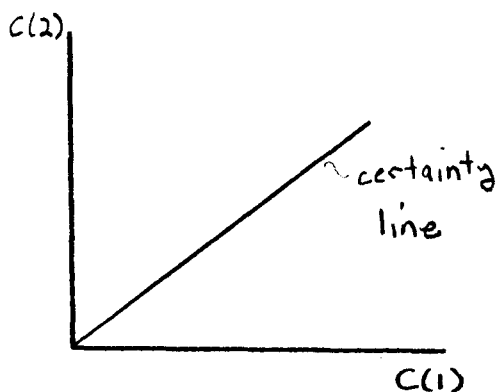
Figure 1

slope of the indifference curves equals $-\pi(1)/\pi(2)$, independently of the form of the utility function-- so long as the form of the utility function does not itself depend on the state of nature that occurs. The certainty line is thus an "expansion path."

Consider an individual whose initial endowment consists of a certain claim on Y_0 units of consumption goods. The individual is then confronted with a bet which he can undertake in any amount α so long as $\alpha \leq Y_0$. If α units of the bet are taken, the individual will receive an additional $\alpha X(1)$ units of the consumption good if state 1 occurs, but must sacrifice α units of output if state 2 occurs. Thus the payoff, cost $(dC(1), dC(2))$ associated with taking α goods worth of the bet is $(\alpha X(1), -\alpha)$. The bet is said to be "favorable" if its expected value in terms of goods is positive. The expected value of the bet's payoff stream is

$$\pi(1)\alpha X(1) - \pi(2)\alpha = \pi(1)\alpha X(1)-(1-\pi(1))\alpha$$

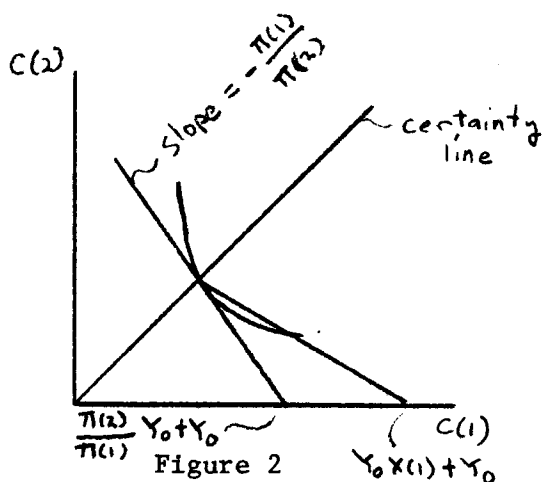$$= \pi(1)\alpha(X(1) + 1) - \alpha.$$

The bet is then favorable if

$$\pi(1)\alpha(X(1) + 1) - \alpha > 0$$

or

(4) $$\pi(1) > \frac{1}{X(1) + 1} .$$

The bet is said to be "fair" if the above inequality is replaced by an equality.

Figure 2

If the individual undertakes α units of the bet, his claims to consumption across states of nature become

$$C(1) = Y_0 + \alpha X(1)$$
$$C(2) = Y_0 - \alpha,$$

from which we can deduce that by varying the amount of the bet taken, α, the individual can substitute $C(1)$ for $C(2)$ at the (constant) rate

$$(5) \qquad \frac{dC(2)}{dC(1)} = -\frac{1}{X(1)} .$$

So $-1/X(1)$ is the slope of the "budget line" through (Y_0, Y_0) along which the individual can trade claims to consumption in state 1 for claims to consumption in state 2. If $\alpha = 0$, the individual's claims remain (Y_0, Y_0). If $\alpha = Y_0$, the individual's claims become $(Y_0 X(1) + Y_0, 0)$. The straight line connecting these two points is the individual's budget line (see Figure 2).

As long as the slope of the budget line exceeds the slope of the individual's indifference curves at Y_0, Y_0, the individual can increase his expected utility by undertaking at least a small part of the bet. This requires that

$$-\frac{1}{X(1)} > -\frac{\pi(1)}{\pi(2)} = -\frac{\pi(1)}{1-\pi(1)} ,$$

from equations (3) and (5). The above inequality can be rearranged to read

$$\pi(1) > \frac{1}{X(1) + 1} ,$$

which is identical with inequality, (4), the condition that the bet be favorable.   For our special case, we have thus proved Arrow's proposition that an individual will always take at least a small part of a favorable bet.[1]

Within this framework, we now consider securities that entitle the individual to alternative patterns of consumption across our two states of nature.   Consider a security, one unit of which entitles the owner to receive $X(1)$ units of the consumption good if state 1 occurs and $X(2)$ units if state 2 occurs.   If the individual buys α units of the security, he is entitled to receive a pattern of returns $(\alpha X(1), \alpha X(2))$ across states of nature.   In Figure 3, one unit of the security gives
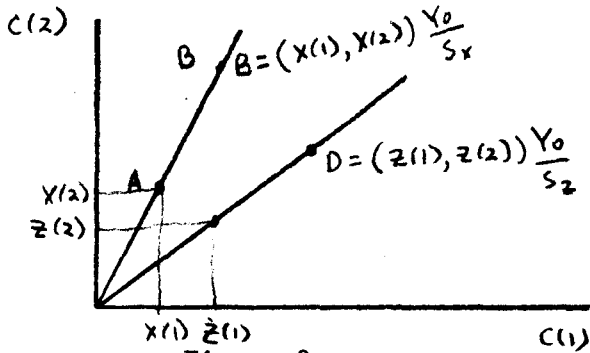
the returns labeled by point A. Suppose that the security costs the individual S_X units of current output per unit of security. If the individual has an investment

Figure 3

portfolio worth Y_0 units of current output, he could then buy Y_0/S_X units of the security and obtain a pattern of returns $(Y_0 X(1)/S_X, Y_0 X(2)/S_X)$ across states of nature.   Point B in Figure 3 depicts such a pattern of returns.

Now suppose that there is a second security, one unit of which gives a pattern of returns $(Z(1), Z(2))$ across states of nature, where $Z(1)$ and $Z(2)$ are both denominated in consumption goods.   If one unit of the security costs S_Z, the individual could purchase Y_0/S_Z units of the security if he put his whole portfolio of Y_0 current goods into that

security. Then his pattern of returns across states of nature would be $(Y_0 Z(1)/S_Z, Y_0 Z(2)/S_Z)$. Such a pattern of returns across states is labeled D in Figure 3.

Now suppose that the individual considers putting a percentage λ of his portfolio into security X, and $1-\lambda$ into security Z. He would then purchase $\lambda Y_0/S_X$ units of security X, and $(1-\lambda)Y_0/S_Z$ units of security Z. His pattern of returns across states of nature would then be

$$(6) \qquad (C(1),C(2)) = Y_0 (\lambda \frac{X(1)}{S_X} + (1-\lambda)\frac{Z(1)}{S_Z}, \lambda\frac{X(2)}{S_X} + (1-\lambda)\frac{Z(2)}{S_Z}).$$

Such points are linear combinations of $(X(1),X(2))Y_0/S_X$ and $(Z(1),Z(2))Y_0/S_Z$, and so lie on the straight line connecting points D and B in Figure 3. A change in λ brings changes in the consumption stream across states according to

$$dC(1) = \left[\frac{X(1)}{S_X} - \frac{Z(1)}{S_Z} \right] Y_0 d\lambda$$

$$dC(2) = \left[\frac{X(2)}{S_X} - \frac{Z(2)}{S_Z} \right] Y_0 d\lambda,$$

so that the "budget line" along which the consumer can alter the pattern of claims to the consumption good across states of nature has slope

$$\frac{dC(2)}{dC(1)} = \frac{\dfrac{X(2)}{S_X} - \dfrac{Z(2)}{S_Z}}{\dfrac{X(1)}{S_X} - \dfrac{Z(1)}{S_Z}} .$$

For this slope to be negative, the numerator and denominator must be of opposite sign, which means that one security must not dominate another.

That is, one unit of current output's worth of security X must offer more consumption in state 1 if it offers less in state 2 than does one unit of current output's worth of security 2. In effect, the ratio

$$\frac{dC(2)}{dC(1)} = \frac{\dfrac{X(2)}{S_X} - \dfrac{Z(2)}{S_Z}}{\dfrac{X(1)}{S_X} - \dfrac{Z(1)}{S_Z}}$$

measures the relative price at which the individual can exchange C(1) for C(2) by trading security X for security Z.

By suitably choosing λ (which need not be between 0 and 1), the individual is able to obtain any combination of C(1) and C(2) in the nonnegative quandrant satisfying equation (6). A negative λ or one exceeding unity indicates that one security or the other is being sold short or being issued by the individual (see F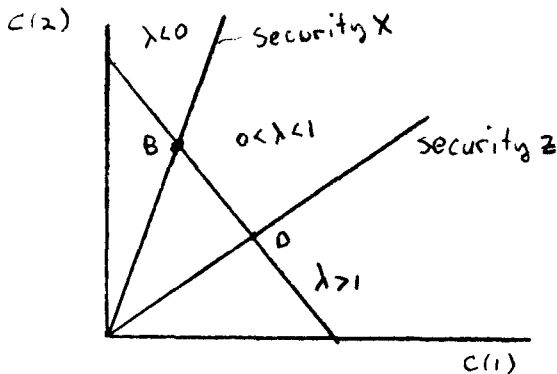igure 4). Notice that by choosing his portfolio suitably, the individual can set C(1) equal to C(2), so that he need bear no risk, if that is his desire.

The individual chooses his portfolio so as to maximize

Figure 4

his expected utility subject to the budget constraint (6). Usually, this involves choosing λ so that it corresponds to a point of tangency between an indifference curve and the budget line. As always, however, corner solutions are possible.

Suppose now that a third security, security y, is added to our setup. The security has returns across states $(y(1)/S_y, y(2)/S_y)$ measured in consumption goods in states 1 and 2, respectively, per unit of current consumption good; S_y is the price of one unit of the security in terms

of current <u>consumption goods</u>. If the individual initially uses his

entire investment portfolio to purchase X or y or Z he will obtain the

C(1), C(2) combination denoted by point X, y, or Z in Figure 5. Now

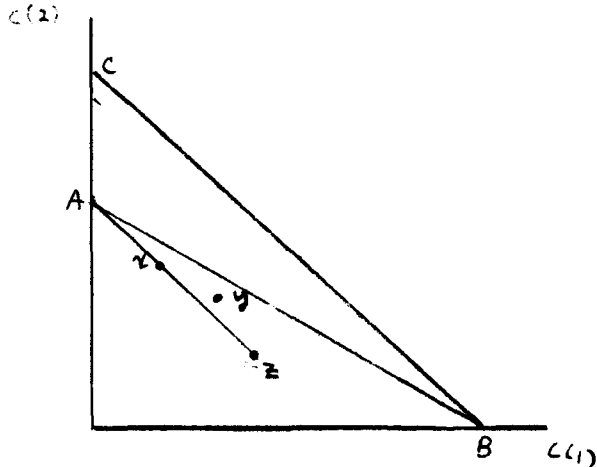since there are three securities, and since individuals are assumed to

be perfect competitors in buying,

selling and issuing these securities,

there are three ratios that

summarize how individuals can

exchange C(1) for C(2). One

ratio refers to exchanging C(1)

for C(2) by trading X for Z; a

second by trading X for y; a

third by trading y for Z. These

three ratios are

Figure 5

$$\frac{dC(2)}{dC(1)} = \frac{\frac{X(2)}{S_X} - \frac{Z(2)}{S_Z}}{\frac{X(1)}{S_X} - \frac{Z(1)}{S_Z}} \qquad \text{X for Z}$$

$$\frac{dC(2)}{dC(1)} = \frac{\frac{X(2)}{S_X} - \frac{Y(2)}{S_Y}}{\frac{X(1)}{S_X} - \frac{Y(1)}{S_Y}} \qquad \text{X for Y}$$

$$\frac{dC(2)}{dC(1)} = \frac{\frac{Y(2)}{S_Y} - \frac{Z(2)}{S_Z}}{\frac{Y(1)}{S_Y} - \frac{Z(1)}{S_Z}} \qquad \text{Y for Z}$$

Unless these three ratios are equal, arbitrage possibilities are present.

By this we mean that by capitalizing on the discrepancies among the

ratios, individuals can execute a series of transactions that provide a

certain return of indefinite magnitude. For example, consider the situation depicted in Figure 5, where the three ratios above are not equal since points X, y, and Z fail to lie along a straight line. Suppose the individual initially puts his entire investment portfolio into security X, and then issues security Z and buys security X with the proceeds so as to move from point X to point A in Figure 5. Next suppose that he moves along line AB to point B by selling security X and purchasing security y. Once at point B he can move to point C by selling security Z and buying more y. Once at C, he can become even better off by issuing more X and buying more y, and so on. There is thus no limit to the $(C(1), C(2))$ combination that the individual can command by executing such transactions. Notice that the individual if he chooses can obtain an indefinitely large but perfectly certain $(C(1) = C(2))$ consumption stream by exploiting these opportunities. This illustrates how pure arbitrage transactions involve no risk.

The upshot of course is that the three exchange ratios $dC(2)/dC(1)$ listed above must be equal. That is, all three points X, y, and Z in Figure 5 must lie along the same straight line. In the example above, presumably investors' attempts to exploit the arbitrage possibilities would cause the prices S_X, S_Z, and S_y to change so that X, y, and Z would lie along a straight line. Thus arbitrage requires that where there are two states of the world there be at most two securities whose returns across states are linearly independent. Similar reasoning implies that where there are n states of the world, the returns on at most n securities can be linearly independent (i.e., the rank of the matrix of returns on securities across states is at most n).

There are really only two goods in the preceding example, C(1) and C(2). The securities X, y, and Z have value only because they represent claims to C(1), C(2) bundles. In effect, having three securities

with linearly independent returns implies that three prices obtain for
those two goods, raising the arbitrage possibilities discussed above.
To exhibit the logic of this and other problems, it is sometimes analytically
convenient to work with "pure" securities that pay off one unit of
consumption in state i and nothing in any other state. Such securities
were introduced by Arrow and are known as Arrow-Debreu contingent securities. The return vector for such a security lies along one of the axes. Thus, in our 2-state example, one unit of a state 1 contingent security offers a return vector $(1,0)$, while one unit of a state 2 security offers a return vector $(0,1)$ (see Figure 6).
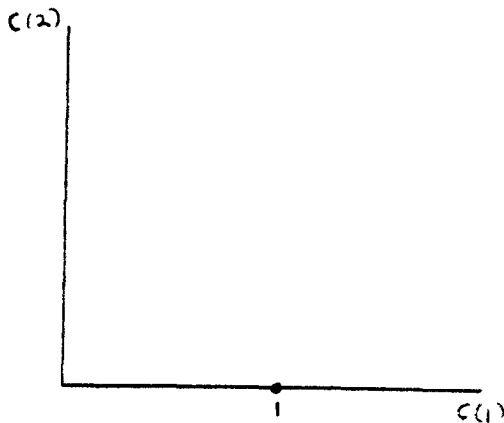

Figure 6

Even where such contingent securities don't literally exist,
it is possible effectively to "trade" them and to compute implicit
prices for them where the number of ordinary securities equals the
number of states of nature. For example, consider our 2-state example
where securities X and Z exist. Security X derives its value from the
value that consumers attach to the consumption stream the security
delivers. Let $p(i)$ be the amount of current output an individual would
sacrifice to obtain one more unit of consumption in state i. Then it
must be so that

$$S_X = X(1)p(1) + X(2)p(2)$$

(7)

$$S_Z = Z(1)p(1) + Z(2)p(2),$$

i.e., the price of each real security must reflect the value of the consumption streams that the security represents a claim on. The above equations can be solved for $p(1)$, $p(2)$, the implicit prices of the contingent securities, so long as $X(1)Z(2) - Z(1)X(2) \neq 0$, i.e., so long as the returns on securities X and Z are not linearly dependent.

More generally suppose there are n states of the world. Suppose that markets for the n contingent securities do not actually exist and that claims promising one dollar if state θ occurs never are traded. Instead, as in the real world, there are n' different companies each selling claims entitling the owner to share in the company's profits. It is easy to show that so long as there are more independent companies than states of nature, it is as if there existed markets in n contingent securities, since by buying and selling actual securities in the proper fashion, the individual can obtain any desired pattern of returns across states of nature.

We suppose that the ith firm's returns across states of nature are given by $X_i(\theta)$, $\theta = 1, \ldots, n$. Select n such firms each of whose patterns of returns across states of nature are not linearly dependent on the returns of the remaining (n-1) firms. That is, for each $i = 1, \ldots, n$ the vector $(X_i(1), X_i(2), \ldots, X_i(n))$ cannot be written as a linear combination of the (n-1) vectors $(X_j(1), X_j(2), \ldots, X_j(n))$ for $j \neq i$. Let the market values of our n firms be $V_1, V_2, \ldots, V_n$. If there were n contingent securities, each promising to pay one dollar in state θ and having price $p(\theta)$, the values of the n firms would have to obey

$$V_1 = \sum_{\theta=1}^{n} X_1(\theta)p(\theta)$$

$$V_2 = \sum_{\theta=1}^{n} X_2(\theta)p(\theta)$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$V_n = \sum_{\theta=1}^{n} X_n(\theta)p(\theta),$$

or in compact notation

(10) $V = Xp$

$$\text{where } V = \begin{bmatrix} V_1 \\ \cdot \\ \cdot \\ \cdot \\ V_n \end{bmatrix}, \ X = \begin{bmatrix} X_1(1)X_1(2)\ldots X_1(n) \\ \cdot \\ \cdot \\ \cdot \\ X_n(1)X_n(2)\ldots X_n(n) \end{bmatrix}, \ p = \begin{bmatrix} p(1) \\ \cdot \\ \cdot \\ \cdot \\ p(n) \end{bmatrix}.$$

Since X is of full rank, (10) can be used to solve for p, giving

(11) $p = X^{-1}V.$

Equation (11) tells us how to unscramble the implicit prices of the n
implicit contingent securities from the market values of n firms and
the patterns of their returns across states of nature.

Liquidity Preference as Behavior Towards Risk[2]

Tobin's explanation of the demand for money as emerging partly as a result of wealthholders' desire to diversify their holdings can be viewed as an application of the theory just described. Suppose that there are two states of the world and that there are two assets: a risky asset that pays off $X(\theta)$ in state $\theta = 1, 2$ for each unit of current output's worth of the asset; and a riskless asset called "money" that pays off one unit of current output, regardless of state, for each unit of current output invested in it. From our preceding discussion, we know that the household will hold at least a little of the risky asset provided that the expected rate of return is positive, i.e., provided that holding the risky security amounts to undertaking a favorable bet. By investing one sure unit of output ("money") in the risky asset, the investor obtains an expected return of

$$\pi(1)X(1) + \pi(2)X(2)$$

which must exceed unity if the security is to offer the individual a favorable bet. Notice that for money to be held, it must be so that either $X(1) < 1$ or $X(2) < 1$, or else the risky asset would dominate money. The expected rate of return on the risky asset, denoted by r, is given by
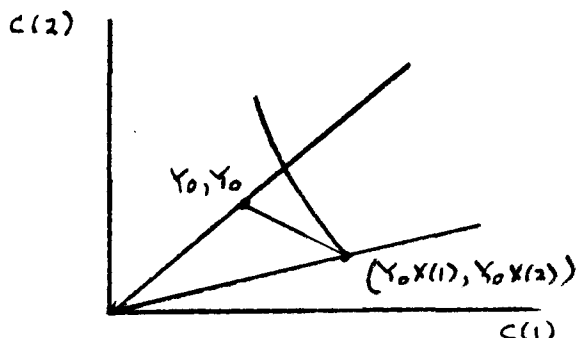
$$r = \pi(1)X(1) + \pi(2)X(2) - 1.$$

It is easy to show that if we start from a position in which $r = 0$, an increase in r, i.e., an increase in either $X(1)$ or $X(2)$, will cause the investor to increase his holdings of the risky asset and decrease his holdings of money. (This is an example of Arrow's proposition that at least a small part of a favorable bet will be undertaken. We leave it

to the reader to work out the details.) Clearly, by risk aversion, if
$r = 0$, the investor will hold his entire portfolio in terms of money.
Notice that we have established that at low enough interest rates, the
investor's holdings of money will vary inversely with the interest rate
on risky assets.

At higher interest rates, an increase in r (i.e., in X(1) or
X(2)) may or may not cause holdings of money to contract. As usual,
there are two effects: a substitution effect inducing a movement along
an indifference curve, an effect which leads to lower money holdings;
and a wealth or income effect, which may or may not offset the substitu-
tion effect, depending on the shape of the investor's indifference
curves.

Notice that it is possible that the investor will want to hold
no money (though if $r > 0$, he will always want to hold some of the risky
asset). This will occur if the rate of return on the risky asset is so
high that the situation is as depicted in Figure 7, where the budget line
is flatter than the indifference curve even where the investor's entire portfolio is in the risky asset.

(Y_0, Y_0) = returns vector if whole portfolio held in money.

$(Y_0 X(1), Y_0 X(2))$ = returns vector if whole portfolio held in risky asset.

Figure 7

The theory described above has often been embodied in a somewhat different form, set forth by Tobin. As above the individual is assumed to maximize expected utility

$$v = \sum_{\theta=1}^{n} \pi(\theta)U(C(\theta)).$$

If we know $C(\theta)$ and $\pi(\theta)$ for each θ, it is straightforward to deduce a probability distribution $g(C)$ which gives the probability that consumption will obtain the value C. In particular,

$$g(\tilde{C}) = \left[\sum_{\theta \in T} \pi(\theta), \quad T = \left\{ \theta \,|\, C(\theta) = \tilde{C} \right\} \right].$$

In the finite-state case currently under discussion, $g(\tilde{C})$ will obtain a nonzero, positive value at only a finite number of $n' \leq n$ values of consumption C. Denote these values of consumption as $C_1, C_2, \ldots, C_{n'}$. Then expected utility v can be written as

$$(8) \qquad v = \sum_{i=1}^{n'} U(C_i)g(C_i),$$

where $\sum_{i=1}^{n'} g(C_i) = 1.$

In a setup with a continuum of states of the world and where consequently C is allowed to take on any real value, the probability associated with consumption occurring in a neighborhood of width ε around C is given by $f(C;B)$ where $f(C;B)$ is the distribution function associated with C and B is a list of parameters determining that distribution. In this case expected utility v is

$$(9) \qquad v = v(B) = \int_{-\infty}^{\infty} U(C)f(C;B)dC.$$

Here expected utility is a function only of the parameters B determining

the distribution of consumption. If there is only one parameter in B,

as would be true if C were distributed according to the Poisson distribution,

then expected utility would depend only on the value of that one parameter.

If there are p parameters in B, then expected utility depends on all p

of them.

The theory has been developed for distributions $f(C;B)$ which

can be characterized by two parameters--one measuring mean or central

tendency, the other measuring variance. The normal distribution is an

example of such a distribution, being completely characterized by the

mean and variance of the distribution. Members of the class of stable

distributions of Paul Levy are also characterized by two parameters.

Following Tobin, suppose B consists of the mean μ_C and standard

deviation σ_C of consumption, so that

$$f(C;B) = f(C;\mu_C, \sigma_C).$$

It greatly facilitates the analysis also to assume that $f(C;B)$ is
a "stable" distribution. A variate Z with density $f(Z;B_2)$ is said
to be stable if when another variate y with the same form of density
$f(y;B_y)$, perhaps with $B_Z \neq B_y$, is added to Z, the result is to produce
a variate $X = Z + y$ obeying the same probability law $f(X;B_X)$. Assuming
the distribution $f(C;B)$ is stable is natural because stable distributions
are the only distributions that serve as the limiting distribution in
central limit theorems. The normal distribution is the best known
of stable distributions. The central limit property of stable distributions
is useful here because the random variable C is often thought of as
representing a linear combination of a large number of independently
distributed returns on various investments, implying that it will
approximately follow a stable distribution.

Assuming that $f(C;B)$ has two parameters, mean μ_C and standard
deviation σ_C, and that it is also a stable distribution amounts to
assuming that it is a normal distribution.[3] That is because the normal
distribution is the only (symmetric) stable distribution for which the
standard deviation exists. Then expected utility v is

$$v(\mu_C, \sigma_C) = \int_{-\infty}^{\infty} U(C) f(C; \mu_C, \sigma_C) dC.$$

Defining the standarized variable Z as

$$Z = \frac{C - \mu_C}{\sigma_C},$$

we have that $C = \mu_C + \sigma_C Z$. Then

$$v = v(\mu_C, \sigma_C) = \int_{-\infty}^{\infty} U(\mu_C + \sigma_C Z) f(Z; 0, 1) dZ,$$

where $f(Z; 0, 1)$ is the standard, unit variance normal distribution. Since expected utility $v(\mu_C, \sigma_C)$ depends only on the two parameters μ_C, σ_C we can define indifference curves in the μ_C, σ_C plane, i.e., combinations of μ_C and σ_C that yield constant levels of expected utility. Along such curves, we have

$$dv = 0 = d\mu_C \int_{-\infty}^{\infty} U'(\mu_C + \sigma_C Z) f(Z; 0, 1) dZ$$

$$+ d\sigma_C \int_{-\infty}^{\infty} Z U'(\mu_C + \sigma_C Z) f(Z; 0, 1) dZ,$$

so that the slope of the indifference curves in μ_C, σ_C plane is

$$(10) \qquad \frac{d\mu_C}{d\sigma_C} = - \frac{\int_{-\infty}^{\infty} Z U'(\mu_C + \sigma_C Z) f(Z; 0, 1) dZ}{\int_{-\infty}^{\infty} U'(\mu_C + \sigma_C Z) f(Z; 0, 1) dZ} .$$

Since $U'' < 0$, while $f(Z; 0, 1)$ is symmetric the numerator on the right is negative so long as $\sigma_C > 0$ (negative Z's being multiplied by larger U's than positive Z's); the denominator is positive since $U' > 0$. Thus the slope (10) is positive. To find the concavity of the indifference curves, we differentiate (10) with respect to σ_C to obtain

$$\frac{d^2\mu_C}{d\sigma_C^2} = - \frac{\int Z^2 U''(\mu_C + \sigma_C Z) f(Z;0,1) dZ}{\int U'(\mu_C + \sigma_C Z) f(Z;0,1) dZ}$$

$$- \frac{\frac{d\mu_C}{d\sigma_C} \int Z U''(\mu_C + \sigma_C Z) f(Z;0,1) dZ}{\int U'(\mu_C + \sigma_C Z) f(Z;0,1) dZ}$$

$$- \frac{d\mu_C}{d\sigma_C} \int Z U''(\mu_C + \sigma_C Z) f(Z;0,1) dZ$$

$$- \left(\frac{d\mu_C}{d\sigma_C}\right)^2 \int U''(\mu_C + \sigma_C Z) f(Z;0,1) dZ.$$

In each case, the limits of integration are $-\infty$, ∞. Since $Z^2 \geq 0$ and $U'' < 0$, the numerator of the first term is negative, making that term positive, since it is preceded by a negative sign. The numerator of the second term is also negative because in the large, U'' must be decreasing in absolute value as C increases in order for $U' > 0$ while $U'' < 0$ for all $C \epsilon [0,\infty)$. This makes the second term positive. Likewise the third and fourth terms are also positive, taking into account the signs preceding them. Thus we have that

$$\frac{d^2\mu_C}{d\sigma_C^2} > 0,$$

which shows that each indifference curve has a slope that increases as we move upward along a curve. An example of a map of such curves is depicted in Figure 8.
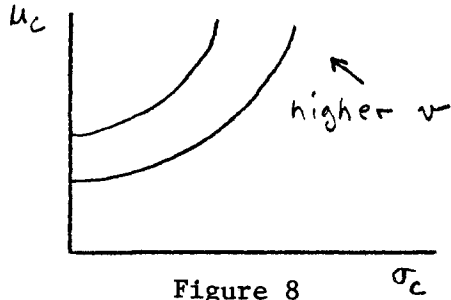
μ_c

higher v

Figure 8 σ_c

It is convenient to use equation (10) to compute the slope of the indifference curves at zero standard deviation. We have

$$\frac{d\mu_c}{d\sigma_c}\Bigg)_{\sigma_c = 0} = -\frac{\int Z U'(\mu_c) f(Z;0,1) dZ}{\int U'(\mu_c) f(Z;0,1) dZ} = \frac{\int Z f(Z;0,1) dZ}{\int f(Z;0,1) dZ} = \frac{-E(Z)}{1} = 0$$

The numerator equals zero, since the normal distribution is symmetric about $Z = 0$, and since $U'(\mu_c)$ is independent of Z. Thus the indifference curves have zero slope for $\sigma_c = 0$. This property of the indifference curves will be seen to reflect that an individual will always take at least a small part of a favorable risk.

To take a specific example, suppose

$$U(C) = -e^{-\lambda C} \qquad \lambda > 0.$$

Notice that

$$U'(C) = \lambda e^{-\lambda C} > 0 \qquad \text{for} \qquad C \epsilon (-\infty, \infty)$$

$$U''(C) = -\lambda^2 e^{-\lambda C} < 0 \qquad \text{for} \qquad C \epsilon (-\infty, \infty).$$

The density function for the normal distribution is

$$f(C; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(C-\mu)^2}{2\sigma^2}}.$$

Consequently expected utility is given by

$$E(U(C)) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} -e^{-\lambda C} e^{-\frac{(C-\mu)^2}{2\sigma^2}} dC$$

$$(11) \qquad = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} -e^{-(\lambda C + \frac{(C-\mu)^2}{2\sigma^2})} dC.$$

Notice that

$$\lambda C + \frac{(C-\mu)^2}{2\sigma^2} = \frac{2\lambda C\sigma^2 + C^2 - 2\mu C + \mu^2}{2\sigma^2}$$

$$= \frac{[C-(\mu-\lambda\sigma^2)]^2 + 2\lambda\mu\sigma^2 - \lambda^2\sigma^4}{2\sigma^2}$$

$$= \frac{[C-(\mu-\lambda\sigma^2)]^2}{2\sigma^2} + \lambda(\mu-(1/2)\lambda\sigma^2).$$

Substituting the above expression into (11) gives

$$E[U(c)] = -e^{-\lambda(\mu-(1/2)\lambda\sigma^2)} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(C-(\mu-\lambda\sigma^2))^2}{2\sigma^2}} \, dC.$$

But we know that

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(C-\mu')^2}{2\sigma^2}} \, dC = 1$$

for any μ' and $\sigma > 0$. So we have

$$E[U(C)] = -e^{-\lambda(\mu-\frac{\lambda\sigma^2}{2})}$$

Along curves of constant expected utility we require

$$\mu - (1/2)\lambda\sigma^2 = \text{constant}.$$

So the mean, standard deviation indifference curves satisfy

$$d_\mu - \lambda_\sigma d_\sigma = 0$$

so that their slope is

$$\frac{d\mu}{d\sigma} = \lambda\sigma > 0 \quad \text{for } \sigma > 0.$$

and their concavity is

$$\frac{d^2\mu}{d\sigma^2} = \lambda > 0.$$

This concludes our anlysis of our specific example for U(C).

Having characterized the individual's preferences, we now describe his opportunities. Suppose that there is available to the individual a safe asset that has the property that if he puts his entire portfolio into this asset he will obtain a consumption stream characterized by mean C_0 and standard deviation zero. Suppose there is also an alternative asset (or maybe a portfolio of other assets) such that if the individual uses his entire portfolio to purchase this asset he obtains consumption goods in an amount $C_0 X$ where X is a normally distributed random variable with mean μ_X and variance σ_X^2. If the individual invests a proportion $(1-\lambda)$ in the risky asset and λ in the safe asset, $0 \leq \lambda \leq 1$ he receives consumption

$$C = \lambda C_0 + (1-\lambda)C_0 X.$$

Then the mean of his consumption would be

$$(12) \qquad \mu_C = \lambda C_0 + (1-\lambda)C_0\mu_X.$$

Notice that

$$C - \mu_c = (1-\lambda)C_0(X-\mu_X),$$

so that

$$(13) \qquad \sigma_c = (1-\lambda)C_0\sigma_X.$$

Solving (13) for λ, substituting into (12) and rearranging we obtain

$$(14) \qquad \mu_c = C_0 + (\frac{\mu_X - 1}{\sigma_X})\sigma_c, \qquad 0 \leq \sigma_c \leq C_0\sigma_X$$

which gives the locus of combinations of μ_c, σ_c attainable by varying

λ. The mean μ_c rises linearly with the standard deviation σ_c, the slope

being $(\mu_X-1)/\sigma_X$. Of course, the expected rate of return (μ_X-1) must

exceed zero for the opportunity locus to have a positive slope. This is

the condition that the risk be favorable. Such an opportunity locus is
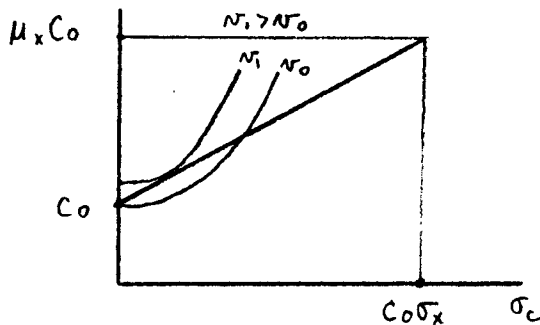
depicted in Figure 9.

Figure 9

As we have seen, the slope of the indifference curves at $\sigma_c = 0$ is zero. That means that if $\mu_X-1 > 0$, the individual will always take at least a small part of the risk, since then the opportunity locus through $(0,C_0)$ has a positive slope, permitting the individual to move to a higher

level of expected utility by taking some risk. It follows that beginning

from a situation where $\mu_X-1 = 0$, an increase in the rate of return on

the risky asset, μ_X-1, will lead to a decrease in the amount held in the

safe asset and an increase in holdings of the risky asset. Thus, for a
low enough rate of return on the risky asset, an increase in that rate
does cause a decrease in the investor's demand for the safe asset. For
higher values of the rate of return on the risky asset, however, an
increase in that rate will not necessarily lead to a decrease in holdings
of the safe asset, there being offsetting substitution and wealth effects.
We leave it to the reader to study these offsetting effects in the
context of the present graphical formulation of the theory. Needless to
say, all of these features of the analysis have their counterparts in
the state-preference version of the theory which we summarized above.

There are several unsatisfactory aspects of the theory that we
have just sketched. For the formulation cast in terms of the mean and
standard deviation of consumption, we have to assume that consumption is
normally distributed, which requires that we act as if consumption can
be an unbounded negative number. It is difficult to imagine negative
consumption. If to circumvent this difficulty we restrict consumption
to be nonnegative, we must pay for this by adopting a probability function
for consumption that lacks the statistical property of stability,
and so greatly weakens the appeal of the theory. But as we have seen
above, the essence of the theory can be cast in terms of the state-
preference analysis where the assumption that consumption has a normal
distribution need play no role.

As a theory of the demand for money, the theory is certainly
of limited applicability. For one thing, the occurrence of unforeseen
price level changes makes money a risky asset in terms of goods, so that
the "money" in the model above does not really correspond with the asset

called money in the real world. For another thing, money is dominated by assets like Treasury bills and savings deposits that are as risk-free as money but offer positive nominal yields. At best, the above theory is one about the demand for such assets, not money. To explain the demand for money it seems essential to take into account the presence of transactions costs.

The Modigliani-Miller Theorem

Throughout these pages, we have assumed that firms have no bonds outstanding, that they retain no earnings, and so they finance all of their investment by issuing equities. It is an implication of the "Modigliani-Miller theorem" that our assumptions about these matters are not restrictive.[4] In particular, Modigliani and Miller's analysis implies that in the absence of a corporate income tax, the firm's cost of capital is independent of whether the firm raises the funds by retaining earnings, issuing bonds, or issuing equities. Moreover, Modigliani and Miller's theorem was proved in the context of a model that explicitly recognized the existence of uncertainty. These notes sketch the reasoning of Modigliani and Miller by using the state-preference presentation of Stiglitz.[5]

We collapse the entire future into a single point in the future. We assume that there is a finite number n of possible future states of the world, each state representing an entire constellation of possible outcomes of all sorts of events in the future. We let $\theta = 1$, 2, ..., n be an index over the possible states. In state $\theta = 1$, for example, it rains two inches in Eugene, Oregon, Ali defeats Foreman in the ring, Nixon wins a third term, and so on. States $\theta = 2$, ..., n correspond to different outcomes of this set of events. An individual's happiness, indexed by U, in the event that state θ prevails depends on the usual way on the amounts of n goods that he consumes:

$$U = U(q_1(\theta), \ldots, q_m(\theta)) \quad \partial U/\partial q_i(\theta) > 0$$

$$U(\ ) \text{ concave}$$

where $q_i(\theta)$ is the amount of the ith good consumed by the individual in state θ, $i = 1, \ldots, m$. We have assumed that the form of the utility function $U(\ )$ is independent of the state θ.

The consumer's notions about the likelihood of various states of the world occurring are supposed to be summarized by a set of subjective probabilities $\pi(1), \pi(2), \ldots, \pi(n)$ that obey

$$\sum_{\theta=1}^{n} \pi(\theta) = 1,$$

where $\pi(\theta)$ is the probability that the consumer assigns to state θ occurring. Individuals are assumed to maximize expected utility v:

$$(1) \qquad v = \sum_{\theta=1}^{n} \pi(\theta)U(q_1(\theta),\ldots,q_m(\theta)).$$

The consumer is assumed to come into a certain endowment q_i^0 of claims to goods $i = 1, \ldots, m$, should state θ occur, $\theta = 1, \ldots, n$. It is assumed that there exist competitive futures markets in which individuals trade claims to the ith good in state θ prior to the occurrence of the state. The individual faces a price $p_i(\theta)$ at which he can buy or sell whatever claims he wishes on the ith good contingent on state θ occurring. The value of the consumer's endowment is

$$\sum_{\theta=1}^{n} \sum_{i=1}^{m} p_i(\theta)q_i^0(\theta).$$

The consumer maximizes expected utility v subject to

$$(2) \qquad \sum_{\theta=1}^{n} \sum_{i=1}^{m} p_i(\theta)q_i^0(\theta) = \sum_{\theta=1}^{n} \sum_{i=1}^{m} p_i(\theta)q_i(\theta),$$

which states that the market value of his endowment equals the market

value of the bundle of contingent commodities that he purchases. Where

λ is an undetermined Lagrange multiplier, the consumer's problem can be

formulated as maximizing

$$J = \sum_{\theta=1}^{n} \left\{ \pi(\theta) U(q_1(\theta),\ldots,q_m(\theta)) + \lambda \left(\sum_{i=1}^{m} p_i(\theta)(q_i(\theta)-q_i^0(\theta)) \right) \right\}.$$

The first order conditions are

(3) $\pi(\theta) \dfrac{\partial U}{\partial q_i(\theta)} + \lambda p_i(\theta) = 0 \qquad i = 1, \ldots, m$

    $\dfrac{\partial J}{\partial \lambda} = 0 \qquad\qquad\qquad\qquad \theta = 1, \ldots, n.$

Dividing (3) for θ and i by (3) for $\tilde{\theta}$ and j, we have

(4) $\dfrac{\dfrac{\partial U}{\partial q_i(\theta)}}{\dfrac{\partial U}{\partial q_j(\tilde{\theta})}} = \dfrac{\pi(\tilde{\theta})}{\pi(\theta)} \dfrac{p_i(\theta)}{p_i(\tilde{\theta})} ,$

which is the analogue of the familiar static marginal equality for the

household. From (4) and the budget constraint (2), demand curves for the

nm contingent commodities can be derived. By aggregating these demand

curves over the set of all consumers, market demand schedules can be

obtained, which together with total market endowments permits computing

a general equilibrium in which the prices $p_i(\theta)$, $i = 1, \ldots, m$, $\theta = 1, \ldots, n$

are determined.

Arrow[6] has shown that consumers are just as well off where

these nm markets in m commodities contingent on state $\theta(=1,\ldots,n)$

occurring are replaced by n markets in "contingent securities," with one

security for each state. Each security promises to pay one dollar

should state θ occur. Following the occurrence of a state, consumers

then trade the m goods as described by the standard static model.

It is straightforward to add production to the model sketched above. We consider such a competitive model in which there exists a complete set of n markets for the n contingent securities, each promising to pay one dollar if state θ occurs in the future. The model is assumed to possess a general equilibrium in which the equilibrium present price of a claim to one dollar in state θ is $p(\theta)$. Notice that the price of a sure dollar next period is $\sum_{\theta=1}^{n} p(\theta)$, which can be interpreted as the reciprocal of one plus the risk-free rate of interest. The assumption that there exist perfect markets in the contingent securities for all n states of the world means that it is possible to insure against any risk. Individuals need bear no risks if that is their preference.

We will assume no taxes are present. Now consider a firm whose prospective returns, net of labor and materials costs, but gross of capital costs, are $X(\theta)$ dollars in state θ. Suppose that the firm issues an amount of B dollars worth of bonds. The firm now promises to pay $(r+1)B$ dollars to its bond holders next period, provided that it does not go bankrupt, i.e., provided that $X(\theta) \geq (r+1)B$. If the firm does go bankrupt, i.e., if $X(\theta) \leq (r+1)B$, then the bond holders receive only $X(\theta)$. Thus the realized rate of return on bonds $r(\theta)$ depends on the state of the world:

$$r(\theta) + 1 = \begin{cases} r + 1 & \text{if } X(\theta) \geq (r+1)B \\ \dfrac{X(\theta)}{B} & \text{if } X(\theta) < (r+1)B. \end{cases}$$

Only if $X(\theta) > (r+1)B$ for all θ, is $r(\theta)$ equal to the promised coupon rate r for all θ.

The value of the firm's bonds must equal the sum of the values of the contingent securities that the bond implicitly consists of. For each state in which the firm doesn't go bankrupt, the bonds will in total pay off (r+1)B. The present value of those returns is

$$(r+1)B \sum_{\theta \varepsilon S} p(\theta) \quad \text{where } S = \left\{ \theta \middle| X(\theta) \geq (r+1)B \right\}$$

For states θ in $S' = \left\{ \theta \middle| X(\theta) < (r+1)B \right\}$, in which the firm goes bankrupt, the bonds pay off $X(\theta)$. So the present value of payments in those states is

$$B \sum_{\theta \varepsilon S'} \frac{X(\theta)}{B} p(\theta).$$

The total present value of the firm's bonds B must thus satisfy

$$B = (r+1)B \sum_{\theta \varepsilon S} p(\theta) + B \sum_{\theta \varepsilon S'} \frac{X(\theta)}{B} p(\theta).$$

Dividing by B and solving for (r+1), we obtain

$$(5) \qquad (r+1) = \frac{1 - \sum_{\theta \varepsilon S'} \frac{X(\theta)}{B} p(\theta)}{\sum_{\theta \varepsilon S} p(\theta)},$$

which tells us that the rate of return a firm's bonds must bear depends on the firm's probability of defaulting, and so on the number of bonds it has issued. Notice that if there is zero probability of the firm's going bankrupt, S' being empty, r equals the risk-free rate of interest.

The firm's equities bear a payout stream across state of nature given by

$$X(\theta)-(r+1)B \quad \text{if} \quad X(\theta) \geq (r+1)B$$
$$0 \quad \text{if} \quad X(\theta) < (r+1)B.$$

As with bonds, the value of the firm's equities must equal the sum of the values of the contingent securities that the equities implicitly represent. So we have that the present value of equities E is

(6) $\qquad E = \sum_{\theta \epsilon S} (X(\theta)-(r+1)B)p(\theta).$

Substituting for (r+1) from (5) in the above expression gives

$$E = \sum_{\theta \epsilon S} p(\theta)X(\theta) - B\left(\frac{1 - \sum_{\theta \epsilon S'} \frac{X(\theta)}{B} p(\theta)}{\sum_{\theta \epsilon S} p(\theta)}\right) \sum_{\theta \epsilon S} p(\theta)$$

$$= \sum_{\theta \epsilon S} p(\theta)X(\theta) - B + \sum_{\theta \epsilon S'} X(\theta)p(\theta)$$

(7) $\qquad E = \sum_{\theta \epsilon S \cup S'} p(\theta)X(\theta) - B$

or

(8) $\qquad E + B = \sum_{\theta \epsilon S \cup S'} p(\theta)X(\theta).$

Equation (8) states that the total value of the firm's debt plus equity equals the present value of the firm's return across state of nature, evaluated at the price of claims to one dollar contingent on the associated states of nature. The total value E + B is therefore independent of the ratio of debt to equity.

Now assume that the firm is contemplating a project that costs C sure dollars today, and that will cause the firm's returns to change by dX(θ) in state θ. The value of stockholders' equity if the project

isn't undertaken is given by (7). If the project is undertaken, the value of the original stockholders' equity will be

$$E' = \sum_\theta p(\theta)X(\theta) - B + \sum_{\theta \epsilon S} p(\theta)dX(\theta) - C.$$

The value of the original stockholders' equity is increased by undertaking the project so long as

$$\sum_\theta p(\theta)dX(\theta) - C > 0;$$

the project ought to be undertaken by the firm so long as the above inequality is met because it will increase the value of the equity of initial stockholders. This is true regardless of whether the project is financed by issuing bonds or more equities. In particular, notice that the rate of interest r on the firm's bonds, which depends on the volume of bonds that the firm has outstanding, is not pertinent in helping the firm determine whether or not to undertake the project.

To be more specific, suppose that the project is financed by issuing bonds. If the project isn't undertaken the value of the firm is

$$E^0 + B^0 = \sum_\theta p(\theta)X^0(\theta).$$

If the project is undertaken the value of the firm will be

$$E' + B' = \sum_\theta p(\theta)(X^0(\theta) + dX(\theta)).$$

Therefore

$$E' + B' = E^0 + B^0 + \sum_\theta p(\theta)dX(\theta)$$

or

$$E' - E^0 = \sum_\theta p(\theta)dX(\theta) - (B' - B^0).$$

If the firm finances by issuing bonds, $B' - B^0 = C$, the cost of the project. Thus the value of equities, depending on whether the project is undertaken, values according to

$$E' - E^0 = \sum_\theta p(\theta)dX(\theta) - C,$$

which should exceed zero in order that the project be undertaken.

If the project is financed by issuing equities, $B' = B^0$, implying that

$$E' = E^0 + \sum_\theta p(\theta)dX(\theta).$$

To finance the project, the firm must issue C dollars worth of new equities. The new value of the equities of the original stockholders will be $E' - C$, which is obtained by subtracting C from both sides of the above equation:

$$E' - C = E^0 + \sum_\theta p(\theta)dX(\theta) - C.$$

It follows that $E' - C$ will exceed E^0 if $\sum_\theta p(\theta)dX(\theta) - C$ exceeds zero, as before.

Effects of a Corporate Income Tax

We now suppose that the firm's profits net of interest payments to bond holders are taxed at a corporate profits tax rate t_K. The returns to stockholders then equal $(1-t_K)(X(\theta)-(r+1)B)$ for states in S, i.e., states satifying $X(\theta) \geq (r+1)B$, and zero for states in which bankruptcy occurs. The interest rate r on the firm's bonds continues to obey (5). The value of the firm's equities is now given by

$$E = \sum_{\theta} (1-t_K)(X(\theta))-(r+1)B)p(\theta).$$

Substituting for r from (5) in the above equation and rearranging gives

$$(9) \qquad E = (1-t_K) \sum_{\theta=1}^{n} X(\theta)p(\theta) - B + t_K B.$$

For $t_K > 0$, the value of the firm, $E + B$, varies directly with the stock of bonds outstanding. Equation (9) thus predicts that it is in stockholders' interest to have the firm levered an indefinitely large amount. The presence of the corporate income tax implies that there is an optimal debt-equity ratio for the firm (one indefinitely large) and thus causes the Modigliani-Miller theorem to fail to hold.

We should note that matters become much more complex when individual income taxes with different rates for interest income and capital gains are included in the analysis.

Footnotes

[1] See Kenneth Arrow, _Essays in the Theory of Risk Bearing_, p. 100.

[2] James Tobin, "Liquidity Preference as Behavior Towards Risk," _Review of Economic Studies_, 1958.

[3] The density function for the normal distribution is

$$f(c; \mu_c, \sigma_c) = \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{(C-\mu_c)^2}{2\sigma_c^2}} \quad .$$

With this distribution, for expected utility to be defined, the utility function U(C) must satisfy

$$|U(C)| \leq Ae^{BC^2}, \quad A > 0, \quad B > 0$$

for some A and B. See J. S. Chipman, "The Ordering of Portfolios in Terms of Mean and Variance," _Review of Economic Studies_, April, 1973.

[4] F. Modigliani and M. H. Miller, "The Cost of Capital, Corporation Finance, and the Theory of Investment," _American Economic Review_, June 1958.

[5] J. Stiglitz, _A Re-Examination of the Modigliani-Miller Theorem_, _American Economic Review_, December 1969.

[6] K. Arrow, "The Role of Securities in the Optimal Allocation of Risk Bearing," _Review Economic Studies_, 1964.

Notes on Sticky Wages

These notes describe a simple model that explains why money wages are "sticky," i.e., why wages don't adjust rapidly enough to assure that labor markets "clear" at every moment, so that layoffs never occur and the supply of labor always equals the demand.* The assumption that money wages are sticky in this sense is a key one in most macroeconomic models that purport to explain fluctuations in the unemployment rate. Here the sticky character of money wages is attributed to different attitudes of firms and workers toward risk bearing.

We consider a competitive firm that will be able to sell all that it wants of a perishable output in period t at the price $p(\theta)$. The price $p(\theta)$ depends on the state of the world θ that prevails at date t. For each date t we assume that there are two states of the world, indexed by θ = 1, 2, and that $p(1) > p(2)$. We assume that the same two prices $p(1)$ and $p(2)$, contingent on states θ = 1, and θ = 2, respectively, hold for all t, a kind of stationarity assumption. The firm and its workers share a common view of the probabilities of states 1 and 2 emerging, denoted by $\pi(1)$ and $\pi(2)$, respectively. We assume that $\pi(1)$ and $\pi(2)$ are the same for each t, which together with our other assumptions imposes stationarity on the system. We assume that $0 < \pi(1) < 1$; of course $\pi(1) + \pi(2) = 1$.

The firm's output in state θ in period t is given by $f(n(\theta))$ where $n(\theta) \geq 0$ is the firm's employment, measured in number of men, in

---

*The model described in these notes is a much simplified version of the one developed by Costas Azariadis in "On the Incidence of Unemployment," (unpublished, 1973) although the setup here is not exactly identical with his.

state θ. The production function f satisfies $f' > 0$, $f'' < 0$; the

marginal product of labor is positive but diminishing. We further

assume that $\lim_{n \to 0} f'(n) = \infty$ and that $\lim_{n \to \infty} f'(n) = 0$.

The firm pays workers a money wage $w(\theta)$ in state θ in period

t. The wage may be dependent on state θ, but is independent of time t.

The latter specification is really no restriction, since our stationarity

assumptions are sufficient to imply it as a consequence of optimal firm

behavior. The firm's profits at time t in state θ are then given by

$$p(\theta)f(n(\theta)) - w(\theta)n(\theta).$$

The firm's objective is to maximize the expected discounted value of its

stream of profits over the time interval t=1, ..., T:

$$V = \sum_{t=1}^{T} \sum_{\theta=1}^{2} \delta^t \pi(\theta)(p(\theta)f(n(\theta)) - w(\theta)n(\theta))$$

$$= D \sum_{\theta=1}^{2} \pi(\theta)(p(\theta)f(n(\theta)) - w(\theta)n(\theta))$$

where δ is the discount factor and $D = \sum_{t=1}^{T} \delta^t$. Since D is fixed, we may

just as well assume that the firm attempts to maximize V/D, subject to

the constraints imposed by the labor market. Positing that the firm

maximizes expected profits means that the firm has a neutral attitude

toward risk and is willing to accept fair bets in unlimited amounts.*

---

*That is, it is as if the firm were trying to maximize its
expected utility of profits, but that its utility functionis linear
in profits.

The representative workers possess a utility function that gives his happiness in period t as a function of the wages that he receives and his leisure. It is assumed that a worker either works full time in period t, in which case his level of leisure is $L=L_o$, or else he is unemployed, in which case he has leisure $L_1 > L_o$. There is no part-time or over-time work. The worker's happiness is given by the utility function

$$U = g(w(\theta), L) ; \qquad\qquad g_1 > 0, \ g_{11} < 0$$
$$g_2 \geq 0, \ g_{22} \leq 0$$

which is assumed to be concave and to possess continuous second partial derivatives. If the worker is employed, his utility can be written solely as a function of his wage.

$$U = U(w(\theta)) = g(w(\theta), L_o) ,$$

since L_o is a paremeter. Our assumptions on g imply that $U' > 0$, $U'' < 0$, so that the worker is assumed to be risk averse. We let r denote the pecuniary value the worker attaches to having leisure L_1 rather than L_o, i.e., it is the amount he would have to be paid to make him indifferent between working and not working. So r is defined by the equality

$$g(0, L_1) = g(r, L_o) .$$
$$= U(r) .$$

The worker's pecuniary income is thus $w(\theta)$ in state θ if he is employed in state θ, and r if he is unemployed. The worker maximizes expected utility, which given our definition of r, can be written as a function only of the distribution of his pecuniary income across states.

The firm employs $n(1)$ workers in state 1 and $n(2)$ workers in state 2 at each date. Its procedure is to offer jobs to a number max $(n(1), n(2))$ workers, and then to lay off in a random fashion workers in the state for which $n(\theta) < \max (n(1), n(2))$. In our case, since $p(1) > p(2)$, there is a presumption that the firm will set $n(1) > n(2)$, so that max $(n(1), n(2)) = n(1)$. Workers are aware of the firm's policy on this matter, and consequently realize that if state 2 occurs in a given period, only $n(2)/n(1)$ of the workers having jobs with the firm will actually be employed in that period. Since the firm lays off workers randomly, if state 2 occurs the worker believes that his chances of working are $n(2)/n(1)$ while his chances of being laid off are $\frac{n(1)-n(2)}{n(1)}$.

Laid off workers receive no wages from the firm.

The worker seeks to maximize his expected discounted utility, which is

$$u = \sum_{t=1}^{T} \delta^t (\pi(1)U(w(1)) + \pi(2)\frac{n(2)}{n(1)}U(w(2)) + \pi(2)\frac{n(1)-n(2)}{n(1)} U(r))$$

or

$$u = D[\pi(1)U(w(1)) + \pi(2)\frac{n(2)}{n(1)} U(w(2)) + \pi(2)(1-\frac{n(2)}{n(1)})U(r)]$$

where as before $D = \sum_{t=1}^{T} \delta^t$. Since D is a constant, we might as well

assume that workers maximize

$$v \equiv \frac{u}{D} = \pi(1)U(w(1)) + \pi(2)\frac{n(2)}{n(1)} U(w(2)) + \pi(2)(1-\frac{n(2)}{n(1)})U(r)$$

The firm is assumed to be able to hire as large a labor force as it wants subject to the restraint that its jobs offer workers a level of expected utility v at least as great as $\bar{v}$, where $\bar{v}$ is a market-determined level of expected utility that workers can obtain by accepting jobs with other firms. The firm's problem is thus to maximize expected profits

(1) $\pi(1)(p(1)f(n(1)) - w(1)n(1)) + \pi(2)(p(1)f(n(2)) - w(2)n(2))$

subject to the constraint

(2) $\bar{v} = \pi(1)U(w(1)) + \pi(2)\frac{n(2)}{n(1)} U(w(2)) + \pi(2)(1- \frac{n(2)}{n(1)})U(r)$.

The firm chooses w(1), w(2), n(1), n(2) so as to maximize (1) subject to (2). The firm's problem is thus to choose a wage and layoff policy to maximize its profits subject to the constraints imposed by the labor market. Notice that according to (2), workers are willing to sacrifice

some wages in state 2 for more security of employment, i.e., a higher $n(2)/n(1)$. Maximization of (1) subject to (2) is carried out through unconstrained maximization of

(3) $\quad J(n(1), n(2), w(1), w(2), \lambda) = \pi(1)(p(1)f(n(1))-w(1)n(1))$

$$+\pi(2)(p(2)f(n(2))-w(2)n(2))+\lambda[\bar{v}-\pi(1)U(w(1))-\pi(2)\frac{n(2)}{n(1)}$$

$$U(w(2))-\pi(2)(1-\frac{n(2)}{n(1)})U(r)]$$

where λ is an undetermined Lagrange multiplier. The first-order conditions for a maximum of (3) are:

(4) $\quad \dfrac{\partial J}{\partial n(1)} = \pi(1)p(1)f'(n(1))-\pi(1)w(1)+\lambda\pi(2)\dfrac{n(2)}{n(1)^2}(U(w(2))-U(r)) = 0$

(5) $\quad \dfrac{\partial J}{\partial n(2)} = \pi(2)p(2)f'(n(2))-\pi(2)w(2)-\lambda\pi(2)\dfrac{1}{n(1)}(U(w(2))-U(r)) = 0$

(6) $\quad \dfrac{\partial J}{\partial w(1)} = -\pi(1)n(1)-\lambda\pi(1)U'(w(1)) = 0$

(7) $\quad \dfrac{\partial J}{\partial w(2)} = -\pi(2)n(2)-\lambda\pi(2)\dfrac{n(2)}{n(1)}U'(w(2)) = 0$

(8) $\quad \dfrac{\partial J}{\partial \lambda} = \bar{v}-\pi(1)U(w(1))-\pi(2)\dfrac{n(2)}{n(1)}U(w(2))-\pi(2)(1-\dfrac{n(2)}{n(1)})U(r) = 0$

Equation (6) can be written

$$n(1) = -\lambda U'(w(1)) ,$$

while equation (7) can be written

$$n(1) = -\lambda U'(w(2)) .$$

Both of these equations can be satisfied only if

(9) $\quad U'(w(2)) = U'(w(1)) .$

But since U'(w) is a monotone function of w--recall that we have assumed that U'' < 0--equation (9) can be satisfied only if

(10) $w(1) = w(2) = w$

According to (10), the wage rate should be independent of the state that occurs. The firm should offer a fixed wage w and not adjust it, say, downward in state 2 simply because $p(1) > p(2)$.

To indicate the forces that lead to the wage being constant across states of nature, we study the worker's and the firm's indifference curves with respect to $w(1)$ and $w(2)$. Holding expected utility v and $n(2)/n(1)$ both constant, $w(1)$ and $w(2)$ can vary so long as they satisfy

$$0 = dv = \pi(1)U'(w(1))dw(1) + \pi(2)\frac{n(2)}{n(1)}U'(w(2))dw(2).$$

Hence, in the $w(1)$, $w(2)$ plane, the slope of the worker's indifference curve is

$$\frac{dw(2)}{dw(1)} = -\frac{\pi(1)}{\pi(2)}\frac{n(1)}{n(2)}\frac{U'(w(1))}{U'(w(2))} < 0.$$

Differentiating the above equation with respect to $w(1)$ gives

$$\frac{d^2w(2)}{dw(1)^2} = -\frac{\pi(1)}{\pi(2)}\frac{n(1)}{n(2)}\frac{U''(w(1))}{U'(w(2))} - \left(\frac{\pi(1)}{\pi(2)}\frac{n(1)}{n(2)}\right)^2 \frac{U'(w(1))^2 U''(w(2))}{U'(w(2))^3} > 0.$$

which shows that the indifference curves are convex.

Holding both expected profits V/D and $n(1)$ and $n(2)$ constant, the firm is content with movements in $w(1)$ and $w(2)$ that satisfy

$$-\pi(1)n(1)dw(1) - \pi(2)n(2)dw(2) = d\left(\frac{V}{D}\right) = 0$$

or

$$\frac{dw(2)}{dw(1)} = -\frac{\pi(1)}{\pi(2)}\frac{n(1)}{n(2)} < 0,$$

which gives the slope of the firm's indifference curves in the w(1),

w(2) plane, along which expected profits are constant. Differentiating

the above equation with respect to w(1) gives

$$\frac{d^2w(2)}{dw(1)^2} = 0 \; ,$$

which shows that the firm's indifference curves are straight lines, a

consequence of the firm's neutral attitude toward risk. Higher expected

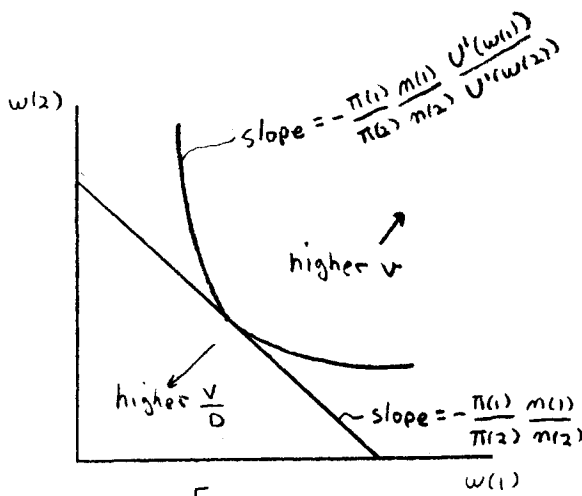profits correspond to firm indifference curves closer to the origin,

Figure I

while higher expected utility
for workers correspond to indiffer-
ence curves further from the origin.
(See figure (1)). The firm maxi-
mizes its expected profits subject
to a fixed level of v by equating
the slope of its iso-expected
profit line to the slope of the
worker's indifference curve:

$$\frac{-\pi(1)}{\pi(2)} \frac{n(1)}{n(2)} = - \frac{\pi(1)}{\pi(2)} \frac{n(1)}{n(2)} \frac{U'(w(1))}{U'(w(2))} \; ,$$

or

$$U'(w(1)) = U'(w(2)) \; .$$

The above condition can only be satisfied where w(1) = w(2). Thus,

given n(1) and n(2), the firm maximizes its expected profits subject to

a constant level of v by eliminating any risk that the wage depend

on the state of demand for the firm's product (i.e., on p(θ)).

We now turn to analyze the determinants of $n(1)$ and $n(2)$, and so the probability of employment, $n(2)/n(1)$. To start with an interesting extreme case, we begin by assuming that $r = 0$, so that workers derive no additional utility from their additional leisure when unemployed at time t.   Under this assumption, it happens that $n(2) = n(1)$, so that employment is constant across states and the probability of employment is unity. To show this in a convenient way, note that if the firm employs $n(\theta)^o$ men in state θ, $n(2)^o < n(1)^o$, and pays a wage of w^o that is fixed across states of nature, its expected profits are

$$(\frac{V}{D})^o = \pi(1)p(1)f(n(1)^o) + \pi(2)p(2)f(n(2)^o) - n(1)^o w^o$$
$$(\pi(1) + \pi(2)\frac{n^o(2)}{n^o(1)}) \ .$$

Since workers are risk-averse, they prefer a certain wage of

$$\tilde{w} = w^o(\pi(1) + \pi(2)\frac{n^o(2)}{n^o(1)}) < w^o \ ,$$

to the uncertain wage package of w^o with probability $\pi(1) + \pi(2)\frac{n^o(2)}{n^o(1)}$ zero with probability $\pi(2)(1 - n^o(2)/n^o(1))$. So if the firm employs $n^o(1)$ workers in each state, paying them the reduced wage $\tilde{w}$, workers are better off. The firm's expected wage bill is unchanged since the new expected wage bill is

$$n(1)^o\tilde{w} = n(1)^o w^o(\pi(1) + \pi(2)\frac{n^o(2)}{n^o(1)}) \ .$$

The firm's expected profits with $n(1) = n^o(1)$, $n(2) = n^o(1)$ and the certain money wage $\tilde{w}$ are given by

$$(\frac{V}{D})' = \pi(1)p(1)f(n^o(1)) + \pi(2)p(2)f(n^o(1)) - n^o(1)\tilde{w} \ .$$

Subtracting $(V/D)^o$ from $(V/D)'$ gives

$$(\frac{V}{D})' - (\frac{V}{D})^o = \pi(2)p(2)(f(n^o(1)) - f(n^o(2))) \ ,$$

which is positive so long as $n^o(1) > n^o(2)$. Thus, the firm's expected

profits are greater and workers are happier with $n(2) = n^o(1)$, so that

in a sense additional workers in state 2 are free to the firm so long as

$n(2) < n(1)$. Since workers have a positive marginal product, it pays

the firm to set $n(1) = n(2)$.

So with $r = 0$, employment, output, and the wage rate are

each constant across states, this in spite of the fact that the price of

the firm's output varies across states. Evidently, then, the firm's

supply curve is vertical, while its demand for workers is independent of

the real wage in terms of the firm's own good. Notice that the equilibrium

in which $w(1) = w(2)$ and $n(1) = n(2)$ is one in which the worker bears no

risk since he receives a certain wage of $w(1)$. The firm bears all the

risk. This is a consequence of the firm's risk neutrality and the

worker's risk-aversion. The firm is willing to accept whatever fair

bets are offered to it, while the worker attempts to avoid any fair

bets. There is thus incentive for the firm and individuals to trade

risks so that the firm accepts whatever risks must be borne. The character

of our results stems directly from the asymetrical attitudes toward risk

that we have attributed to the firm, on the one hand, and to the worker,

on the other hand.

We now examine the case in which $r > 0$. We begin by taking

the total differential of the firm's expected profits,

$$d(\frac{V}{D}) = -(\pi(1)n(1)+\pi(2)n(2))dw+\pi(1)[p(1)f'(n(1))-w]dn(1)$$

$$+\pi(2)[p(2)f'(n(2))-w]dn(2) .$$

Assume that $n(1) = n(2)$ initially and hold $n(1)$ fixed (i.e., set $dn(1) = 0$). Then we have

$$d(\frac{V}{D}) = -n(1)dw+\pi(2)[p(2)f'(n(2))-w]dn(2) .$$

The firm is willing to bear variations in w and $n(2)$ so long as expected profits remain unchanged, $(d(\frac{V}{D}) = 0)$, i.e., so long as the variations of w and $n(2)$ satisfy

(11) $$dw = \pi(2)(p(2)f'(n(2))-w)\frac{dn(2)}{n(1)} .$$

Since $p(2)f'(n(2))-w < 0$,* the firm is willing to increase w if it can decrease $n(2)$. Notice that

(12) $$d(\frac{n(2)}{n(1)}) = \frac{dn(2)}{n(1)} - \frac{n(2)}{n(1)} \frac{dn(1)}{n(1)} ,$$

so that if $dn(1) = 0$, as we are assuming, then $d(n(2)/n(1)) = dn(2)/n(1)$.

Where $r > 0$, is it still true that contracts will be written so that $n(2) = n(1)$? The worker's expected pecuniary income is

$$E(\tilde{w}) = \pi(1)w+\pi(2)\frac{n(2)}{n(1)} w+\pi(2)(1- \frac{n(2)}{n(1)})r ,$$

where now $r > 0$. Taking the total differential of the above equation and setting $n(2)/n(1) = 1$, we have

---

*The marginal conditions (6) and (7) imply $\lambda = -n(1)/U'(w)$. Substituting this into marginal condition (5) gives

$$p(2)f'(n(2))-w = -(\frac{U(w)-U(r)}{U'(w)}) < 0 ,$$

since $U(w)-U(r) > 0$ and $U'(w) > 0$.

(13) $dE(\tilde{w}) = dw + \pi(2)[w-r]d(\frac{n(2)}{n(1)})$.

Now if $n(2) = n(1)$ initially, firms are just willing to raise wages and decrease $n(2)$ so long as dw and $dn(2)$ obey (11), which with use of (12) can be written as

(14) $dw = \pi(2)(p(2)f'(n(2))-w)d(\frac{n(2)}{n(1)})$.

The effects on the expected income of the worker of such a variation in an initial w and $n(2)$, starting from a position where $n(1) = n(2)$, are found by substituting (14) into (13):

(15) $dE(\tilde{w}) = \pi(2)[p(2)f'(n(2))-r]d(\frac{n(2)}{n(1)})$.

Now if $p(2)f'(n(2))-r < 0$, the worker's expected pecuniary income will increase as a result of a _lowering_ of $n(2)/n(1)$, since then sign $dE(\tilde{w})=$ $-$ sign $d(\frac{n(2)}{n(1)})$. This means that if $p(2)f'(n(2))-r < 0$ and $\frac{n(2)}{n(1)} = 1$, the firm, in effect, is in a position to offer the worker a favorable bet. By bearing a little uncertainty i.e., accepting a decrease in $n(2)/n(1)$ below unity, the worker can increase his expected pecuniary income. As we have seen above, a risk-averse individual who behaves as our worker does will always take at least a small part of a favorable bet (this is Arrow's proposition, which we have encountered in several guises already). This means that the worker will be anxious to get, and the firm willing to offer, a contract in which $n(2)/n(1) < 1$. Thus, if $p(2)f'(n(2))-r < 0$, $n(2)/n(1)$ cannot equal unity; some workers will be unemployed in state 2.

The following considerations provide a heuristic way of understanding what is going on here. Suppose we begin from a situation where $n(2)/n(1) = 1$. Since $p(2)f'(n(2))-w < 0$, given w, the firm would have

higher expected profits if n(2) were smaller. By lowering n(2) by dn(2), the firm's expected profits would increase by (p(2)f'(n(2))-w)dn(2). So the firm would actually be willing to pay unemployment compensation in an amount up to $-[p(2)f'(n(2))-w]$ per man in order to have dn(2) fewer people working in state 2; i.e., the firm would be willing to pay this much in order to have some people not work. In order not to work in state 2, a worker would want the firm to pay him at least w-r, the excess of his pecuniary income when he is working over that when he isn't. Then the firm is willing to pay the worker not to work more than he requires to be induced not to work if

$$(w-p(2)f'(n(2)))-(w-r) > 0$$

$$(16) \qquad r-p(2)f'(n(2)) > 0 \; ,$$

which is our condition for n(2)/n(1) to be less than unity. In our setup, the firm is constrained from actually paying unemployment compensation, but part of the "surplus" indicated if condition (16) is met is distributed to workers in the form of the firm's offering workers what amounts to a favorable bet for them. In our setup, workers can share in the "surplus" only by bearing some risk. That is, we have ruled out the possibility that the firm directly offers to insure workers against unemployment, offering to pay them some amount when they are unemployed. Since workers are risk averse there seems to be an incentive for such an institution to emerge.

To show that this indeed can be the case, suppose that firms now consider paying workers an amount per worker of w(3) for not working in state 2. The firm's expected profits are then

$$\frac{V}{D} = \pi(1)(p(1)f(n(1))-w(1)n(1))+\pi(2)(p(2)f(n(2))-w(2)n(2))$$

$$-\pi(2)w(3)(n(1)-n(2)) \ ,$$

where $n(1)-n(2)$ is the number of men in state 2 receiving unemployment compensation from the firm. The worker's expected utility is

$$v = \pi(1)U(w(1))+\pi(2)\frac{n(2)}{n(1)} \ U(w(2))+\pi(2)(1- \frac{n(2)}{n(1)})g(w(3)), \ L_1)$$

The firm's problem can then be formulated as the unconstrained maximization of

$$J = \pi(1)(p(1)f(n(1))-w(1)n(1))+\pi(2)(p(2)f(n(2))-w(2)n(2))$$

$$-\pi(2)w(3)(n(1)-n(2))$$

$$+\lambda[\bar{v}-\pi(1)U(w(1))-\pi(2)\frac{n(2)}{n(1)} \ U(w(2))-\pi(2)(1- \frac{n(2)}{n(1)})g(w(3)), \ L_1)]$$

where λ is again a Lagrange multiplier. The marginal conditions for $w(1)$ and $w(2)$ are identical with (6) and (7), which can be written as

(6') $n(1) = -\lambda U'(w(1))$

(7') $n(1) = -\lambda U'(w(2))$.

Equating to zero the partial derivative of J with respect $w(3)$ gives

$$-\pi(2)(n(1)-n(2))-\lambda\pi(2)(\frac{n(1)-n(2)}{n(1)})\frac{\partial g(w(3), \ L_1)}{\partial w(3)} \ = \ 0$$

or

(17) $n(1) = -\lambda\dfrac{\partial g(w(3), \ L_1)}{\partial w(3)}$

Together, equations (6'), (7') and (17) imply

$$(18) \qquad U'(w(1)) = U'(w(2)) = \frac{\partial g(w(3), L_1)}{\partial w(3)} \qquad .$$

Given the monotone nature of $U'(\ )$, the above equality implies that $w(1) = w(2)$.

Consider now the particular utility function

$$(19) \qquad U = g(w, L) = h(w+BL) \ ; \qquad B > 0; \ h' > 0, \ h'' < 0.$$

This utility function is characterized by straight-line indifference curves between wages and leisure, so that wages and leisure are perfect substitutes. Given this utility function, $U(w)$ is given by

$$U(w) = h(w+BL_0) \ ;$$

r is defined by

$$h(r+BL_0) = h(BL_1)$$

or

$$r+BL_0 = BL_1$$

$$r = B(L_1-L_0) \quad .$$

Notice, that

$$g(w(3), L_1) = h(w(3)+BL_1)$$

$$= h(w(3)+r+BL_0)$$

$$= g(w(3)+r, L_0)$$

$$= U(w(3)+r) \ .$$

This implies that

$$\frac{\partial g(w(3), L_1)}{\partial w(3)} = U'(w(3)+r) \quad .$$

Given the above equality, equation (18) becomes

$$U'(w(1)) = U'(w(2)) = U'(w(3)+r)$$

which together with the monotone nature of $U'(\ )$ implies

$$w(1) = w(2) = w(3)+r = w.$$

Then if at $n(2) = n(1)$, $r > p(2)f'(n(2))$, condition (16), the firm will set $n(2) < n(1)$ but pay unemployment compensation at the rate $w(3) = w-r$. In this fashion, for the particular utility function (19), the labor contract is fashioned so that workers bear no risks, trading them all to the firm.

For utility functions not of the form (19), it will not in general be true that $w(3)+r = w(1) = w(2)$. Still, for many utility functions it will be true that (19) can be satisfied with $w(3) > 0$, so that the firm will opt to set $n(1) > n(2)$ and pay unemployment compensation if condition (16) is met. Evidently, this institutional arrangement is Pareto superior to the one posited at the beginning of these notes, which had the effect of preventing the firm from offering workers unemployment insurance, thereby ruling out a certain "market."

Notes on Difference Equations and Lag Operators

The backward shift or lag operator is defined by

$$L\, X_t = X_{t-1}$$

(1) $\qquad L^n\, X_t = X_{t-n}$ $\qquad\qquad\qquad$ for $n = \ldots\, -2,-1,0,1,2,\ldots$

Multiplying a variable X_t by L^n thus gives the value of X shifted back
n periods. Notice that if $n < 0$ in (1), the effect of multiplying X_t by
L^n is to shift X <u>forward</u> in time by $(-n)$ periods.

We shall consider polynomials in the lag operator

$$A(L) = a_0 + a_1\, L + a_2\, L^2 + \ldots$$

$$= \sum_{j=0}^{\infty} a_j\, L^j,$$

where the a_j's are constants and $L^0 \equiv 1$. Multiplying X_t by $A(L)$ yields
a moving sum of X's:

$$A(L)\, X_t = (a_0 + a_1 L + a_2 L^2 + \ldots)\, X_t$$

$$= a_0\, X_t + a_1\, X_{t-1} + a_2\, X_{t-2} + \ldots$$

$$= \sum_{j=0}^{\infty} a_j\, X_{t-j}.$$

It is generally convenient to work with polynomials $A(L)$ that
are "rational," meaning that they can be expressed as the ratio of two
(finite order) polynomials in L:

$$A(L) = \frac{B(L)}{C(L)}$$

where

$$B(L) = \sum_{j=0}^{m} b_j L^j, \quad C(L) = \sum_{j=0}^{n} C_j L^J \quad \text{where}$$

the b_j's and C_j's are constant. Assuming that A(L) is rational amounts to imposing a more economical and restrictive parameterization on the a_j's.

To take the simplest example of a rational polynomial in L, consider

$$(2) \qquad A(L) = \frac{1}{1-\lambda L} \; .$$

For the scalar $|C| < 1$, we know that

$$(3) \qquad \frac{1}{1-C} = 1 + C + C^2 + \dots .$$

This suggests treating λL of (2) exactly like the C of (3) to get

$$(4) \qquad \frac{1}{1-\lambda L} = 1 + \lambda L + \lambda^2 L^2 + \dots,$$

an expansion which is sometimes only "useful" so long as $|\lambda| < 1$. To prove that the equality (4) is true, multiply both sides of (4) by $(1-\lambda L)$ to obtain

$$\frac{1-\lambda L}{1-\lambda L} = 1 = (1+\lambda L+\lambda^2 L^2+\dots) - \lambda L(1+\lambda L+\lambda L^2+\dots) = 1,$$

which holds for any value of λ, not just values of λ obeying $|\lambda| < 1$. The reason that sometimes we say that (4) is "useful" only if $|\lambda| < 1$ derives from the following argument. We intend often to multiply $1/(1-\lambda L)$ by X_t to obtain the infinite moving sum

$$(5) \qquad \frac{1}{1-\lambda L} X_t = (1+\lambda L+\lambda^2 L^2+\dots) X_t$$

$$= \sum_{i=0}^{\infty} \lambda^i X_{t-i}.$$

Consider this sum for a path of X which is constant over time, so that $X_{t-i} = \bar{X}$ for all i and all t. Then the sum of (5) becomes

$$\frac{1}{1-\lambda L} X_t = \bar{X} \sum_{i=0}^{\infty} \lambda^i$$

The sum $\sum_{i=0}^{\infty} \lambda^i$ equals $(1/(1-\lambda))$ if $|\lambda| < 1$.

But if $|\lambda| \geq 1$ that sum is unbounded, being $+\infty$ if $\lambda \geq 1$. We will sometimes (though not always) be applying the polynomial in the lag operator (4) in situations in which it is appropriate to go infinitely far back in time; and we sometimes find it necessary to insist that in such cases the infinite sum in (5) exist where X has been constant through time. This is what leads to the requirement sometimes imposed that $|\lambda| < 1$ in (4). As we shall see, however, in standard analyses of difference equations, which take the starting point of all processes as some point only finitely far back into the past, the requirement that $|\lambda| < 1$ need not be imposed in (4).

It is useful to note that there is an alternative expansion for the "geometric" polynomial $1/(1-\lambda L)$. For notice that

$$(6) \qquad \frac{1}{1-\lambda L} = \frac{-\frac{1}{\lambda L}}{1 - \frac{1}{\lambda} L^{-1}}$$

$$= \frac{-1}{\lambda L} (1 + \frac{1}{\lambda} L^{-1} + (\frac{1}{\lambda})^2 L^{-2} + \ldots)$$

$$= \frac{-1}{\lambda} L^{-1} - (\frac{1}{\lambda})^2 L^{-2} - (\frac{1}{\lambda})^3 L^{-3} - \ldots,$$

an expansion which is especially "useful" where $|\lambda| > 1$, i.e., where $|1/\lambda| < 1$. So (6) implies that

$$\frac{1}{1-\lambda L} X_t = -\frac{1}{\lambda} X_{t+1} - \frac{1}{\lambda 2} X_{t+2} - \ldots$$

$$= -\sum_{i=1}^{\infty} (\frac{1}{\lambda})^i \; X_{t+i},$$

which shows $(1/(1-\lambda L)) \; X_t$ to be a geometrically declining weighted sum of _future_ values of X. Notice that for this infinite sum to be finite for a constant time path $X_{t+i} = \bar{X}$ for all i and t, the series

$$- \sum_{i=1}^{\infty} (\frac{1}{\lambda})^i$$

must be convergent, which requires that $|\frac{1}{\lambda}| < 1$.

To illustrate how polynomials in the lag operator can be manipulated, consider the difference equation

$$(7) \qquad Y_t = \lambda Y_{t-1} + b \; X_t + a \qquad\qquad t = -\infty, \ldots, 0, 1, 2, \ldots$$

where X_t is an exogenous variable and Y_t is an endogenous variable. Write the above equation as

$$(1-\lambda L) \; Y_t = a + b \; X_t.$$

Dividing both sides of the equation by $(1-\lambda L)$ gives

$$Y_t = \frac{a}{1-\lambda L} + \frac{b}{1-\lambda L} X_t.$$

$$(8) \qquad Y_t = \frac{a}{1-\lambda} + b \sum_{i=0}^{\infty} \lambda^i \; X_{t-i},$$

since $a/(1-\lambda L) = \sum_{i=0}^{\infty} \lambda^i \; a = a \sum_{i=0}^{\infty} \lambda^i = a/(1-\lambda)$ provided $|\lambda| < 1$. So the first-order difference equation (7) and the geometric distributed lag equation (8) are equivalent. Equation (8) can be regarded as the "solution" to (7), since it describes the entire path of Y associated with a given time path for X. Notice that for the Y_t defined by (8) to be

finite, $\lambda^i X_{t-i}$ must be "small" for large i. More precisely, we require

(9) $\qquad \lim\limits_{n\to\infty} \sum\limits_{i=n}^{\infty} \lambda^i X_{t-i} = 0$, for all t.

For the case of X constant for all time, $X_{t-i} = \bar{X}$ all i and t, this condition becomes

$$\lim\limits_{n\to 0} \bar{X} \frac{\lambda^n}{1-\lambda} = 0,$$

which requires $|\lambda| < 1$. Notice also that the infinite sum $a \sum\limits_{i=0}^{\infty} \lambda^i$ in

(8) is finite only if $|\lambda| < 1$, in which case it equals $a/(1-\lambda)$, or if $a = 0$, in which case it equals zero regardless of the value of λ. We tentatively assume that $|\lambda| < 1$.

For analyzing difference equations with arbitrary initial conditions given, it is convenient to rewrite (8) for $t > 0$ as

$$Y_t = a \sum\limits_{i=0}^{t-1} \lambda^i + a \sum\limits_{i=t}^{\infty} \lambda^i + b \sum\limits_{i=0}^{t-1} \lambda^i X_{t-i}$$

$$+ b \sum\limits_{i=t}^{\infty} \lambda^i X_{t-i}.$$

$$= \frac{a(1-\lambda^t)}{1-\lambda} + \frac{a\lambda^t}{1-\lambda} + b \sum\limits_{i=0}^{t-1} \lambda^i X_{t-i}$$

$$+ b \lambda^t \sum\limits_{i=0}^{\infty} \lambda^i X_{0-i}$$

(10) $\qquad Y_t = \frac{a(1-\lambda^t)}{1-\lambda} + b \sum\limits_{i=0}^{t-1} \lambda^i X_{t-i}$

$$+ \lambda^t \{\frac{a}{1-\lambda} + b \sum\limits_{i=0}^{\infty} \lambda^i X_{0-i}\} . \qquad\qquad t \geq 1$$

The term in braces equals Y_0, as reference to expression (8) will confirm. So (10) becomes

$$Y_t = \frac{a(1-\lambda^t)}{1-\lambda} + b \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \lambda^t Y_0$$

or

$$(11) \qquad Y_t = \frac{a}{1-\lambda} + \lambda^t (Y_0 - \frac{a}{1-\lambda}) + b \sum_{i=0}^{t-1} \lambda^i X_{t-i}. \qquad t \geq 1$$

Now textbooks on difference equations often analyze the special

case in which $X_t = 0$ for all $t > 0$. Under this special circumstance

(11) becomes

$$(12) \qquad Y_t = \frac{a}{1-\lambda} + \lambda^t (Y_0 - (\frac{a}{1-\lambda})),$$

which is the solution of the first-order difference equation $Y_t = a + \lambda Y_{t-1}$

subject to the initial condition that Y equals the arbitrarily given

value Y_0 at time. Notice that if $Y_0 = a/(1-\lambda)$, then (12) implies

$Y_t = Y_0$ for all $t \geq 0$, which shows $a/(1-\lambda)$ to be a "stationary point"

or long-run equilibrium value of Y. Notice also that if, as we are

assuming, $|\lambda| < 1$, then (12) implies that

$$\lim_{t \to \infty} Y_t = \frac{a}{1-\lambda} ,$$

which shows that the system is "stable," tending to approach the

stationary point as time passes.

Now consider the first-order system (7) under the assumption

that $a = 0$, so that $a \sum_{i=0}^{\infty} \lambda^i$ equals zero regardless of the value of λ.

Then the appropriate counterpart to (10) is

$$Y_t = b \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \lambda^t b \sum_{i=0}^{\infty} \lambda^i X_{0-i}.$$

Assuming that condition (9) is met even where $|\lambda| > 1$ (so that the second term in the equation is finite), the above equation becomes

$$Y_t = b \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \lambda^t Y_0 \qquad t \geq 1$$

As before we analyze the special case where $X_t = 0$ for all $t > 0$. Then the above equation becomes

$$Y_t = \lambda^t Y_0. \qquad t \geq 1$$

The stationary point of this solution is zero, since if $Y_0 = 0$, Y will remain equal to zero forever, regardless of the value of λ. However, if $|\lambda| > 1$, the system will diverge farther and farther from this stationary point if either $Y_0 > 0$, or $Y_0 < 0$. If $\lambda > 1$, Y_t will tend toward $+\infty$ as $t \to \infty$ provided $Y_0 > 0$; Y_t will tend toward $-\infty$ as $t \to \infty$ if $Y_0 < 0$. If $\lambda < -1$, Y_t will display explosive oscillations of periodicity two time periods.

Where an arbitrary initial condition finitely far back in time is not supplied, so that the process is thought of as starting up infinitely far back in time, equation (8) is the solution to (7) provided that $|\lambda| < 1$ and that condition (9) is met. (We require $|\lambda| < 1$ so that a $\sum_{i=0}^{\infty} \lambda^i$ be finite.) It may seem that to the right side of (8) we could add a term $\alpha\lambda^t$, where α is arbitrary, to get

$$(8') \qquad Y_t = \frac{a}{1-\lambda} + b \sum_{i=0}^{\infty} \lambda^i X_{t-i} + \lambda^t \alpha,$$

which seems to be a solution of (7). To see this, notice that (8') implies

$$\lambda Y_{t-1} = \frac{a\lambda}{1-\lambda} + b \sum_{i=1}^{\infty} \lambda^i X_{t-i} + \lambda^t \alpha.$$

Subtracting the above equation from (8') yields equation (7), so that

(8') is indeed a solution of (7). However, notice that we are requiring

(8') to be a solution for all t. But if $\alpha > 0$, for example, then

$$\lim_{t \to -\infty} \lambda^t \alpha = \lim_{i \to \infty} \lambda^{-i} \alpha = \infty$$

if $0 < \lambda < 1$. (If $1 < \lambda < 0$, the λ^{-i} will display unbounded,

undamped oscillations as $t \to -\infty$). Thus, if we require that the limit of

$|Y_t|$ must be finite as $t \to -\infty$, we must in general have that $\alpha = 0$ in (8'),

so that (8') collapses to (8). The condition that $\lim_{t \to -\infty} |Y_t| < \infty$ is,

in effect, an initial condition that we are imposing on the solution.

If $|\lambda| > 1$, then (8) is not the appropriate solution for

(7). A solution can be obtained by solving in the forward direction,

using equation (6). The solution to (7) is then

$$(8'') \qquad Y_t = - \sum_{i=1}^{\infty} (\frac{1}{\lambda})^i X_{t+i} - \frac{\frac{a}{\lambda}}{1 - \frac{1}{\lambda}}$$

where we require that

$$(9') \qquad \lim_{n \to \infty} \sum_{i=n}^{\infty} (\frac{1}{\lambda})^i X_{t+i} = 0,$$

so that the above infinite sum is finite.

As before, (8'') remains a solution to (7) if the term

$\alpha \lambda^t$, α arbitrary, is added to the right side of (8''):

$$Y_t = - \sum_{i+1}^{\infty} (\frac{1}{\lambda})^i X_{t+i} - \frac{\frac{a}{\lambda}}{1 - \frac{1}{\lambda}} + \alpha \lambda^t.$$

To see this, subtract λY_{t-1} from both sides of the above equation.

But since $|\lambda| > 1$, if $\alpha \neq 0$, then the above equation implies that

for many X paths satisfying condition (9') (e.g., a path for which X is constant for all times),

$$\lim_{t \to \infty} |Y_t| = \infty.$$

This occurs because for $\alpha \neq 0$,

$$\lim_{t \to \infty} |\alpha \lambda^t| = \infty,$$

since $|\lambda| > 1$. Since we want Y to be finite for all t, we will impose the requirement

$$\lim_{t \to \infty} |Y_t| < \infty$$

which implies that $\alpha = 0$. So (8'') is the solution to (7) for $|\lambda| > 1$ that satisfies the "terminal condition" summarized by the above inequality.

Second-Order Difference Equations

Consider the second-order difference equation

(13) $Y_t = t_1 Y_{t-1} + t_2 Y_{t-2} + a + b X_t.$

Using lag operators, (13) can be written as

$$(1 - t_1 L - t_2 L^2) Y_t = a + b X_t$$

or

(14) $\qquad Y_t = \dfrac{a}{1-t_1L-t_2L^2} + \dfrac{b}{1-t_1L-t_2L^2} X_t.$

by long division it is easy to verify that

(15) $\qquad \dfrac{b}{1-t_1L-t_2L^2} = \sum\limits_{i=0}^{\infty} w_i L^i$

where $w_0 = b_0$

$\qquad w_1 = b_0 t_1$

$\qquad w_j = t_1 w_{j-1} + t_2 w_{j-2} \quad$ for $j \geq 2.$

That is,

$$
1-t_1L-t_2L^2 \,\Big|\!\overline{\begin{array}{l} \dfrac{1+t_1L+(t_2+t_1^2)L^2 + (t_1(t_2+t_1^2)+t_1t_2)L^3}{} +... \\[2pt] 1 \\[4pt] \dfrac{1-t_1L-t_2L^2}{} \\[4pt] t_1L + t_2L^2 \\[4pt] \dfrac{t_1L - t_1^2 L^2 - t_1t_2L^3}{} \\[4pt] (t_2+t_1^2)L + t_1t_2L^3 \\[4pt] \dfrac{(t_2+t_1^2)L - t_1(t_2+t_1^2)L^3 - t_2(t_2+t_1^2)L^4}{} \quad ... \end{array}}
$$

Notice that the weights in (15) follow a geometric pattern if $t_2 = 0$, as we would expect, since then (13) collapses to a first-order equation.

It is convenient to write the polynomial $(1-t_1 L-t_2L^2)$ in an alternative way, given by the "factorization"

(16) $\qquad 1-t_1 L-t_2 L^2 = (1-\lambda_1 L)(1-\lambda_2 L)$

$\qquad\qquad\qquad\quad = (1-(\lambda_1+\lambda_2)L + \lambda_1\lambda_2 L^2),$

so that $\lambda_1 + \lambda_2 = t_1$ and $-\lambda_1 \lambda_2 = t_2$. To see how λ_1 and λ_2 are related to the "roots" or "zeroes" of $(1-t_1 L - t_2 L^2)$, notice that

$$(1-\lambda_1 L)(1-\lambda_2 L) = \lambda_1 \lambda_2 (\frac{1}{\lambda_1} -L)(\frac{1}{\lambda_2} -L).$$

Therefore the equation

$$0 = (1-\lambda_1 L)(1-\lambda_2 L) = \lambda_1 \lambda_2 (\frac{1}{\lambda_1} -L)(\frac{1}{\lambda_2} -L)$$

is satisfied at the two "roots" $L = \frac{1}{\lambda_1}$ and $L = \frac{1}{\lambda_2}$. Given the polynomial

$1-t_1 L - t_2 L^2$, the roots $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$ are found from solving the

"characteristic equation"

$$1-t_1 L - t_2 L^2 = 0 \text{ or } t_2 L^2 + t_1 L - 1 = 0$$

for two values of L. The roots are given by the quadratic formula

$$(17) \qquad L = \frac{-t_1 \pm \sqrt{t_1^2 + 4t_2}}{2t_2}.$$

Formula (17) enables us to obtain the reciprocals of λ_1 and λ_2 for given values of t_1 and t_2.

So without loss of generality, we can write the second-order difference equation as

$$(1-\lambda_1 L)(1-\lambda_2 L)Y_t = a + b X_t.$$

$$(18) \qquad Y_t = \frac{a}{(1-\lambda_1 L)(1-\lambda_2 L)} + \frac{b}{(1-\lambda_1 L)(1-\lambda_2 L)} X_t.$$

Notice that if $\lambda_1 \neq \lambda_2$

$$\frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)} = \frac{1}{\lambda_1-\lambda_2} (\frac{\lambda_1}{1-\lambda_1 L} - \frac{\lambda_2}{1-\lambda_2 L}),$$

which can be verified directly. Thus (18) can be written

$$Y_t = \frac{a}{(1-\lambda_1)(1-\lambda_2)} + \frac{\lambda_1 b}{\lambda_1 - \lambda_2} \cdot \frac{1}{1-\lambda_1 L} X_t - \frac{\lambda_2 b}{\lambda_1 - \lambda_2} \cdot \frac{1}{1-\lambda_2 L} X_t$$

$$(19) \qquad Y_t = a \sum_{i=0}^{\infty} \lambda_1^i \sum_{j=0}^{\infty} \lambda_2^j + \frac{\lambda_1 b}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_1^i X_{t-i}$$

$$- \frac{\lambda_2 b}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_2^i X_{t-i},$$

where we are making use of the fact that for a constant a

$$H(L)a = \sum_{i=0}^{\infty} h_i L^i a$$

$$= a \sum_{i=0}^{\infty} h_i = a H(1).$$

Notice that

$$\frac{1}{1-\lambda_1 L} \cdot \frac{1}{1-\lambda_2 L} = \sum_{i=0}^{\infty} \lambda_1^i L^i \sum_{j=0}^{\infty} \lambda_2^j L^j$$

so that the sum of the distributed lag weights

$$\sum_{i=0}^{\infty} \lambda_1^i \sum_{i=0}^{\infty} \lambda_2^i \text{ is finite and equals } \frac{1}{(1-\lambda_1)(1-\lambda_2)}$$

provided that both $|\lambda_1| < 1$, $|\lambda_2| < 1$. So in writing (19), we require either that both $|\lambda_1|$ and $|\lambda_2|$ be less than unity or that

$a = 0$, so that $a \sum_{i=0}^{\infty} \lambda_1^i \sum_{j=0}^{\infty} \lambda_2^j$ is defined. Furthermore, we

require that

(9') $\lim\limits_{n \to \infty} \sum\limits_{i=n}^{\infty} \lambda_j^i X_{t-i} = 0$, all t,

hold for j = 1, 2, so that the geometric sums in (19) are both

finite.

Suppose that a = 0. On this assumption write (19) as

(20) $y_t = \dfrac{\lambda_1 b}{\lambda_1 - \lambda_2} \sum\limits_{i=0}^{t-1} \lambda_1^i X_{t-i} - \dfrac{\lambda_2 b}{\lambda_1 - \lambda_2} \sum\limits_{i=0}^{t-1} \lambda_2^i X_{t-i}$, $t \geq 1$

$$+ \frac{\lambda_1^{t+1}}{\lambda_1 - \lambda_2} \theta_0 + \frac{\lambda_2^{t+1}}{\lambda_1 - \lambda_2} n_0$$

where $\theta_0 = b \sum\limits_{i=0}^{\infty} \lambda^i X_{0-i}$

$n_0 = -b \sum\limits_{i=0}^{\infty} \lambda^i X_{0-i}$.

The case in which $X_t = 0$ for $t \geq 1$ is often analyzed, as for the

first-order case. On this assumption, (20) becomes

(21) $Y_t = \dfrac{\lambda_1^{t+1}}{\lambda_1 - \lambda_2} \theta_0 + \dfrac{\lambda_2^{t+1}}{\lambda_1 - \lambda_2} n_0$, $t \geq 1$.

If $\theta_0 = n_0 = 0$, $Y_t = 0$ for all $t \geq 1$, regardless of the values of

λ_1 and λ_2. So Y = 0 is the stationary point or long-run equilibrium

value of (21).

If λ_1 and λ_2 are real, then $\lim\limits_{t \to \infty} Y_t$ will equal zero if

and only if both $|\lambda_1| < 1$ and $|\lambda_2| < 1$, regardless of the values of the

parameters θ_0 and n_0, so long as they are finite. If, however,

$|\lambda_1| > 1$, $|\lambda_2| < |\lambda_1|$, and $\theta_0 > 0$, then $\lim\limits_{t \to \infty} Y_t = +\infty$. If $|\lambda_2| < |\lambda_1| > 1$,

and $\theta_0 < 0$, then $\lim\limits_{t \to \infty} Y_t = -\infty$. Thus, Y will tend toward the stationary

point zero as time passes provided that both $|\lambda_1| < 1$ and $|\lambda_2| < 1$. If

one or both of the λ's exceed one in absolute value, the behavior of Y

will eventually be "dominated" by the term in (21) associated with

the λ that is larger in absolute value; that is, eventually Y will

grow approximately as λ_m^t, where λ_m is the λ_j with the larger absolute

value.

Now suppose that the roots are complex. If the roots are

complex, they will occur as a complex conjugate pair, as the quadratic

formula (17) verifies. So assume that the roots are complex, and write them

as

$$\lambda_1 = re^{iw} = r \cos w + i \sin w$$

$$\lambda_2 = re^{-iw} = r \cos w - i \sin w$$

where the real part is $r \cos w$ and the imaginary part is $\pm r \sin w$.

Notice that

$$(22) \qquad \lambda_1 - \lambda_2 = r(e^{iw} - e^{-iw}) = 2r \, i \sin w.$$

Furthermore, notice that equation (21) can be written

$$(23) \qquad Y_t = \frac{b \, \lambda_1^{t+1}}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_1^i \, X_{0-i} - \frac{b \, \lambda_2^{t+1}}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_2^i \, X_{0-i}$$

$$= \frac{b}{\lambda_1 - \lambda_2} \sum_{j=0}^{\infty} (\lambda_1^{t+j+1} - \lambda_2^{t+j+1}) \, X_{0-j}.$$

Notice that

$$\lambda_1^{t+j+1} - \lambda_2^{t+j+1} = (re^{iw})^{t+j+1} - (re^{-iw})^{t+j+1}$$

$$= r^{t+j+1} (e^{iw(t+j+1)} - e^{-iw(t+j+1)})$$

$$= r^{t+j+1} (2i \sin (w(t+j+1))).$$

But from trigonemtric formulas $\sin(wt+w(j+1)) = \sin wt \cos w(j+1)$ $+\cos wt \sin w(j+1)$.[*] Substituting this into the above formula gives

$$\lambda_1^{t+j+1} - \lambda_2^{t+j+1} = r^{t+j+1} (2i[\sin wt \cos w(j+1) + \cos wt \sin w(j+1)]).$$

Substituting the above equation and (22) into (23) gives

---

[*] Notice that

$$e^{iw_1} = \cos w_1 + i \sin w_1$$

$$e^{iw_1} e^{iw_2} = (\cos w_1 + i \sin w_1)(\cos w_2 + i \sin w_2)$$

$$= (\cos w_1 \cos w_2 - \sin w_1 \sin w_2)$$

$$+ i(\sin w_1 \cos w_3 + \sin w_2 \cos w_1).$$

Also notice that

$$e^{i(w_1+w_2)} = \cos (w_1+w_2) + i \sin (w_1+w_2).$$

Therefore

$$\sin (w_1+w_2) = \sin w_1 \cos w_2 + \sin w_2 \cos w_1$$

and

$$\cos (w_1+w_2) = \cos w_1 \cos w_2 - \sin w_1 \sin w_2.$$

$$Y_t = \frac{br^{t+1}}{2ri\ \sin w} \sum_{j=0}^{\infty} r^j \cdot (2i\,[\sin(wt)\ \cos w(j+1) + \cos wt\ \sin w(j+1))]\ X_{0-j}$$

$$Y_t = \frac{br^t}{\sin w} [\sin wt \sum_{j=0}^{\infty} r^j \cos w(j+1) + \cos wt \sum_{j=0}^{\infty} r^j \sin w(j+1)]\ X_{0-j}$$

$$t \geq 1$$

or

$$(24) \qquad Y_t = \frac{br^t}{\sin w} \cdot \sin wt \cdot Z_0 + \frac{br^t}{\sin w}\ \cos wt \cdot Z_1,$$

where $Z_0 = \sum_{j=0}^{\infty} r^j \cos w(j+1)\ X_{0-j}$

$$Z_1 = \sum_{j=0}^{\infty} r^j \sin w(j+1)\ X_{0-j}.$$

As before $Y = 0$ is the stationary point of the difference equation. For arbitrary initial conditions, i.e., for arbitrary values of the parameters Z_0 and Z_1, Y_t will approach zero as time passes provided that $r < 1$; for (24) describes the evolution of Y over time as the sum of "damped" sin and cosine functions, the damping factor being r^t. (Notice that for it to be possible to divide by sin w in (24), it is necessary that $\sin w \neq 0$, which means that w cannot equal zero, π, 2π,.... This will be satisfied so long as the roots are complex (remember $\lambda_1 = r \cos w + i \sin w$)). If $r > 1$, the oscillations are explosive, while if $r < 1$, the oscillations are damped, and Y_t approaches its stationary value of zero in an oscillatory fashion as time passes.

Notice that if λ_1 and λ_2 are complex, the distributed lag weights of (19) oscillate. Rewrite (19) as

(19') $\qquad Y_t = a \sum\limits_{i=0}^{\infty} \lambda_1^i \sum\limits_{j=0}^{\infty} \lambda_2^j + \dfrac{b}{\lambda_1 - \lambda_2} \sum\limits_{j=0}^{\infty} (\lambda_1^{j+1} - \lambda_2^{j+1}) \ X_{t-j},$

which using calculations similar to those above can be rewritten as

$$Y_t = a \sum_{i=0}^{\infty} \lambda_1^i \sum_{i=0}^{\infty} \lambda_2^i + \frac{b}{\sin w} \sum_{j=0}^{\infty} r^j \sin w \ (j+1) \ X_{t-j}$$

Notice that the damping factor multiplying the sin curve is r^j, so that the range of the weights decreases as the lag j increases, provided that $r < 1$.

As noted above, the roots λ_1 and λ_2 are the reciprocals of the roots of the polynomial

(61) $\qquad 1 - t_1 \ L - t_2 \ L^2 = 0.$

For we know that $1 - t_1 \ L - t_2 \ L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L)$, with roots $1/\lambda_1$ and $1/\lambda_2$. Alternatively, multiply the above equation by L^{-2} to obtain

$$L^{-2} - L^{-1} \ t_1 - t_2 = 0 = (L^{-1} - \lambda_1)(L^{-1} - \lambda_2)$$

or

(62) $\qquad x^2 - t_1 \ X - t_2 = 0$

where $X = L^{-1}$. Notice that the roots of (62) are the reciprocals of the roots of (61). Thus, λ_1 and λ_2 are the roots of (62).

It is interesting to know what values of t_1 and t_2 yield complex roots. Using the quadratic formula we have that the roots of (62) are

$$\lambda_i = X = \frac{t_1 \pm \sqrt{t_1^2 + 4t_2}}{2} \; .$$

For the roots to be complex, the term whose square root is taken must be negative, i.e.,

$$(63) \qquad t_1^2 + 4t_2 < 0,$$

which implies that $t_2 < 0$. In case (63) is satisfied, the roots are

$$\lambda_1 = \frac{t_1}{2} + \frac{i\sqrt{-(t_1^2 + 4t_2)}}{2} = a + bi$$

$$\lambda_2 = \frac{t_1}{2} - \frac{i\sqrt{-(t_1^2 + 4t_2)}}{2} = a - bi$$

To write $a + bi$ in polar form we recall that

$$a + bi = r \cos w + ri \sin w = re^{iw}$$

where $r = a^2 + b^2$ and where $\cos w = a/r$. Thus we have that

$$r = \sqrt{\left(\frac{t_1}{2}\right)^2 - \frac{(t_1^2 + 4t_2)}{4}}$$

$$= \sqrt{-t_2} \; .$$

We also have that

$$\cos w = \frac{t_1}{\sqrt{2} \; \sqrt{-t_2}} \qquad \text{or} \quad w = \cos^{-1}\left(\frac{t_1}{2\sqrt{-t_2}}\right)$$

For the oscillations to be damped we require that $r = \sqrt{-t_2} < 1$, which requires that $-t_2 < 1$.

The periodicity of the oscillations is $2\pi/\cos^{-1}(t_1/\sqrt{-t_2})$; i.e., this is the number of periods from peak to peak in the oscillations.

If the roots are real, movements will be damped if both roots are less than one in absolute value. That requires

$$-1 < \frac{t_1 + \sqrt{t_1^2 + 4t_2}}{2} < 1$$

and

$$-1 < \frac{t_1 - \sqrt{t_1^2 + 4t_2}}{2} < 1.$$

The condition

$$\frac{t_1 + \sqrt{t_1^2 + 4t_2}}{2} < 1$$

implies

$$\sqrt{t_1^2 + 4t_2} < 2 - t_1$$

$$t_1^2 + 4t_2 < 4 + t_1^2 - 4t_1$$

(64) $t_1 + t_2 < 1$

The condition

$$\frac{t_1 - \sqrt{t_1^2 + 4t_2}}{2} > -1$$

implies that

$$-\sqrt{t_1^2 + 4t_2} > -2 - t_1$$

$$\sqrt{t_1^2 + 4t_2} < 2 + t_1$$

$$t_1^2 + 4t_2 < t_1^2 + 4 + 4t_1$$

(65) $t_2 < 1 + t_1.$

Conditions (64) and (65) must be satisfied for the roots, if real, to be less than unity in absolute value.

Notice that both roots are negative and real if $t_1^2 + 4t_2 > 0$ and

$$\frac{t_1 + \sqrt{t_1^2 + 4t_2}}{2} < 0 \text{ which implies}$$

$$t_1 < -\sqrt{t_1^2 + 4t_2}$$

$$t_1^2 > t_1^2 + 4t_2$$

$$0 > t_2.$$

Figure 1 depicts regions of the t_1, t_2 plane for which conditions (63), (64), or (65) are or are not satisfied. The graph shows combinations of t_1 and t_2 that give rise to damped oscillations, explosive oscillations, etc.

An Example

Maybe the most famous second-order difference equation in economics is the one associated with Samuelson's multiplier accelerator model. Samuelson posited the model

$$C_t = c\, Y_{t-1} + \alpha \qquad\qquad 1 > c > 0 \text{ (consumption function)}$$

FIGURE 1

$t_2 = 1 + t_1$

$t_2 < 1 + t_1$

$t_1 + t_2 = 1$

$t_1 + t_2 < 1$

EXPLOSIVE
OSCILLATIONS

EXPLOSIVE
GROWTH

-2.0   -1.0   1.0   2.0 t_1

DAMPED OSCILLATIONS

$t_1^2 + 4t_2 = 0$

$t_1^2 + 4t_2 < 0$

-1.0

EXPLOSIVE OSCILLATIONS
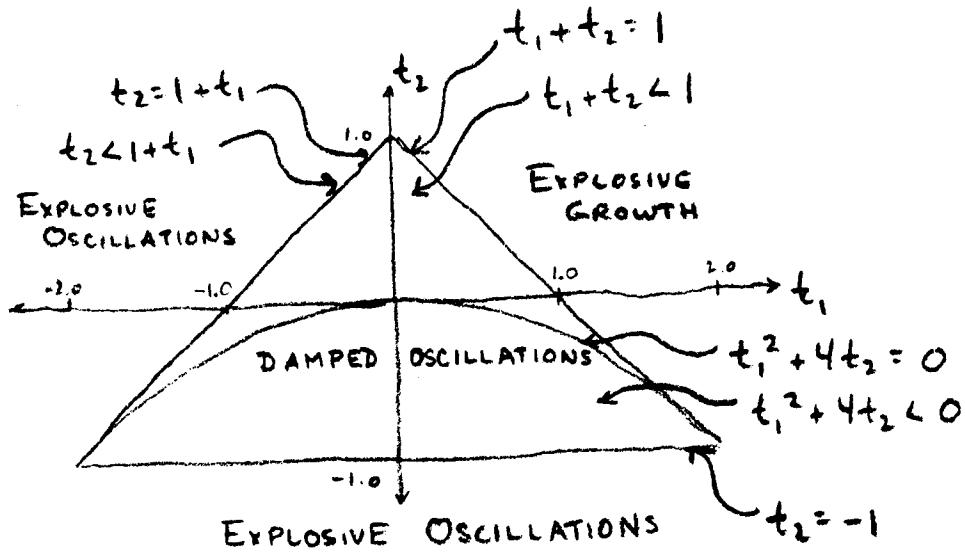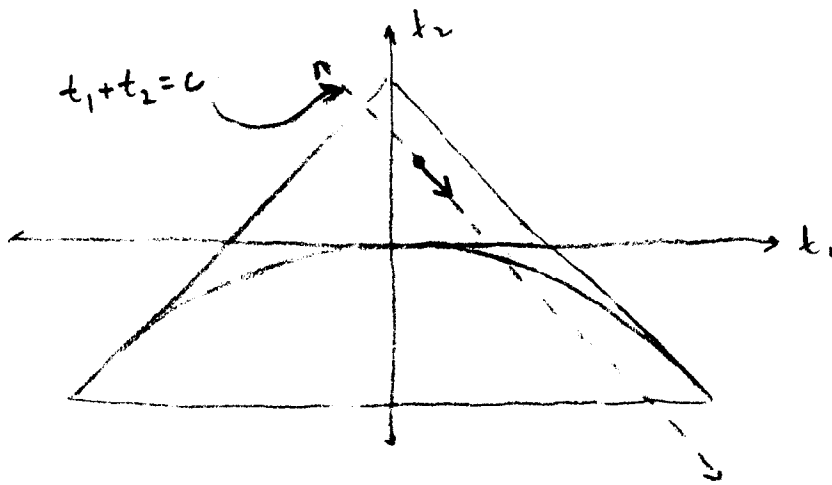
$t_2 = -1$

FIGURE 2

$t_1 + t_2 = 0$

t_2

t_1

$$I_t = \gamma(Y_{t-1} - Y_{t-2}) \qquad\qquad \gamma > 0 \text{ (accelerator)}$$

$$C_t + I_t = Y_t$$

where C_t is consumption and I_t is investment. Substituting the first two equations into the third gives

$$Y_t = (c+\gamma)\, Y_{t-1} + \gamma Y_{t-2} + \alpha$$

or

$$Y_t = t_1\, Y_{t-1} + t_2\, Y_{t-2},$$

where $t_1 = c + \gamma$, $t_2 = -\gamma$. Notice that $t_1 + t_2 = c$. So variations in the parameter γ move the parameters t_1 and t_2 downward and to the right along the line $t_1 + t_2 = c$ in figure 2. Using figure 2, the values of c and γ compatible with damped oscillations, explosive oscillations, and so on, can easily be determined.

Figure 3 shows the path of Y over time for various values of c and γ, and for the initial conditions $Y_0 = Y_1 = 10$.

Second-Order Difference Equations (Equal Roots)

The preceding treatment assumed that $\lambda_1 \neq \lambda_2$. (Notice that we divided by $\lambda_1 - \lambda_2$ to obtain (19).) If $\lambda_1 = \lambda_2$, then the polynomial we must study is

$$\frac{1}{(1-\lambda L)(1-\lambda L)} = \frac{1}{1-\lambda L}\,(1+\lambda L+\lambda^2 L^2 +\ldots)$$

$$= (1+\lambda L+\lambda^2 L^2 +\ldots) + \lambda L(1+\lambda L+\lambda^2 L^2 +\ldots)$$

$$+ \lambda^2 L^2(1+\lambda L+\lambda^2 L^2 +\ldots) +\ldots$$

$$= 1+2\lambda L+3\lambda^2 L^2+...$$

(25)
$$\frac{1}{(1-\lambda L)^2} = \sum_{i=0}^{\infty} (i+1)\lambda^i L^i.$$

The polynomial in (25) is called a second-order Pascal lag distribution.
It is the product of two geometric lag distributions with the same decay
parameter λ.

With the aid of (25) we can study the solution to difference equations of the form

$$(1- \lambda L)^2 Y_t = a + b X_t.$$

The solution is

$$(26) \qquad Y_t = a \sum_{i=0}^{\infty} (i+1)\lambda^i + b \sum_{i=0}^{\infty} (i+1)\lambda^i X_{t-i}.$$

For a $\sum_{i=0}^{\infty} (i+1)\lambda^i$ to be finite, either $|\lambda| < 1$ or $a = 0$ must be satisfied.

To aid in studying difference equations with arbitrary initial conditions, we assume that $a = 0$ and rewrite (26) as

$$(27) \qquad Y_t = b \sum_{i=0}^{t-1} (i+1)\lambda^i X_{t-i} + b \sum_{i=t}^{\infty} (i+1)\lambda^i X_{t-i}.$$

The second sum can be written as

$$b \sum_{j=0}^{\infty} (j+1+t)\lambda^{t+j} X_{0-j}$$

$$= b \sum_{j=0}^{\infty} (j+1)\lambda^{t+j} X_{0-j} + b \sum_{j=0}^{\infty} t\lambda^{t+j} X_{0-j}$$

$$= b \lambda^t \sum_{j=0}^{\infty} (j+1)\lambda^j X_{0-j} + b t\lambda^t \sum_{j=0}^{\infty} \lambda^j X_{0-j}.$$

So for the special case $X_t = 0$ for $t > 0$, (27) becomes

$$(28) \qquad Y_t = \lambda^t \theta_0 + t\lambda^t n_0$$

$$\text{where } \theta_0 = b \sum_{j=0}^{\infty} (j+1)\lambda^j X_{0-j}$$

and

$$n_0 = b \sum_{j=0}^{\infty} \lambda^j x_{0-j}.$$

The stationary point of the equation is zero. For arbitrary initial conditions θ_0 and n_0, Y_t will approach the stationary point as time passes if $|\lambda| < 1$. If $|\lambda| > 1$, the value of Y_t will diverge from the stationary point zero as time passes, unless $\theta_0 = n_0$. Thus, if $|\lambda| > 1$, the stationary point is a "razor's edge" equilibrium.

N^{th}-Order Difference Equations (distinct roots)

Consider a rational polynomial with n^{th} order denominator:

$$A(L) = \frac{F(L)}{G(L)} = \frac{F(L)}{(1-\lambda_1 L)(L-\lambda_2 L)\dots(L-\lambda_n L)}$$

We assume that the degree of the numerator polynomial is less than the degree of the denominator polynomial. The zeros of $G(L)$ are $L_1 = 1/\lambda_1$, $L_2 = 1/\lambda_2, \dots, L_n, \dots, L_n = 1/\lambda_n$, since each of these values for L satisfies the equation

$$G(L) = (1-\lambda_1 L)(1-\lambda_2 L)\dots(1-\lambda_n L) = 0.$$

Suppose the n roots are distinct. Now the method of partial fractions enables us to express $A(L)$ as

(29)
$$\frac{F(L)}{(1-\lambda_1 L)\dots(1-\lambda_n L)} = \frac{A_1}{1-\lambda_1 L} + \frac{A_2}{1-\lambda_2 L} + \dots + \frac{A_n}{1-\lambda_n L}$$

where A_1, A_2, ..., A_n are constants to be determined. To determine them, multiply both sides of the above equation by $(1-\lambda_1 L)\dots(1-\lambda_n L)$ to get

$$F(L) = A_1(1-\lambda_2 L)\dots(1-\lambda_n L) + A_2(1-\lambda_1 L)(1-\lambda_3 L)\dots(1-\lambda_n L) +$$

$$\dots + A_n(1-\lambda_1 L)\dots(1-\lambda_{n-1} L).$$

Evaluating the above equation at $L = \frac{1}{\lambda_1}$, the first root of $G(L)$, gives

$$A_1 = \frac{F(\frac{1}{\lambda_1})}{(1-\frac{\lambda_2}{\lambda_1})\ldots(1-\frac{\lambda_n}{\lambda_1})} \; .$$

For the general term A_i we obtain

$$(29') \qquad A_i = \frac{F(\frac{1}{\lambda_i})}{(1-\frac{\lambda_1}{\lambda_i})\ldots(1-\frac{\lambda_{i-1}}{\lambda_i})(1-\frac{\lambda_{i+1}}{\lambda_i})\ldots(1-\frac{\lambda_n}{\lambda_i})} .$$

As an example, consider applying (29) to the second-order denominator polynomial

$$A(L) = \frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)} = \frac{A_1}{(1-\lambda_1 L)} + \frac{A_2}{1-\lambda_2 L} \; .$$

We have

$$A_1 = \frac{1}{1-\frac{\lambda_2}{\lambda_1}} = \frac{\lambda_1}{\lambda_1-\lambda_2}$$

and

$$A_2 = \frac{1}{1-\frac{\lambda_1}{\lambda_2}} = \frac{\lambda_2}{\lambda_2-\lambda_1} \; .$$

Thus we obtain

$$\frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)} = \frac{1}{\lambda_1-\lambda_2} [\frac{\lambda_1}{1-\lambda_1 L} - \frac{\lambda_2}{1-\lambda_2 L}],$$

Which we earlier used on the way to deriving (19).

Suppose we have an nth order difference equation

(30) $\qquad (1-\lambda_1 L)(1-\lambda_2 L)\ldots(1-\lambda_n L)Y_t = b\ X_t.$

The solution to (30) is obtained by dividing by $(1-\lambda_1 L)\ldots(1-\lambda_n L)$ to obtain

$$Y_t = \frac{b}{(1-\lambda_1 L)\ldots(1-\lambda_n L)}\ X_t.$$

We suppose that the λ_j's are all distinct. Then application of (29) to the above equation gives

$$Y_t = b\ \sum_{r=1}^{n} \left(\frac{A_r}{1-\lambda_r L}\right) X_t$$

(31) $\qquad Y_t = b\ \displaystyle\sum_{r=1}^{n} A_r \sum_{i=0}^{\infty} \lambda_r^i\ X_{t-i}$

which shows that Y_t can be expressed as the weighted sum of n geometric distributed lags with decay coefficients $\lambda_1,\ \lambda_2,\ldots,\ \lambda_n$.

Given n initial values of Y, and assuming $X_t = 0$ always, it is possible to start up difference equation (30) finitely far back in the past, and to obtain a solution of the form

$$Y_t = \lambda_1^t\ n_1 + \lambda_2^t\ n_2 + \ldots + \lambda_n^t\ n_n$$

where $n_1,\ldots,\ n_n$ are constants chosen to satisfy the n initial values. The above equation can be derived from (31) by applying calculations analogous to those applied above in the first and second order cases.

N^{th}-Order Difference Equations (N equal roots)

Consider the nth order difference equation

(32) $(1-\lambda L)^n Y_t = b X_t$

which has the solution

$$Y_t = \frac{b}{(1-\lambda L)^n} X_t.$$

The polynomial $\dfrac{1}{(1-\lambda L)^n}$ is the one associated with an nth order Pascal

lag distribution, which is formed by multiplying (convolving) n geometric

lag distributions with the same decay parameter λ. We have already

studied the second-order Pascal distribution. The binomial expansion

with negative exponent is

$$(1+x)^{-n} = 1 + nx + \frac{n(n+1)}{2!} x^2 + \frac{n(n+1)(n+2)}{3!} x^3 + \dots +$$

$$\frac{n(n+1)\dots(n+i-1)}{i!} x^i + \dots \qquad |x| < 1.$$

Notice that $n(n+1)\dots(n+i-1) = (n+i-1)!/(n-1)!$ The coefficient on x in

the above expansion is thus

$$\binom{n+i-1}{i} \equiv \frac{(n+i-1)!}{i!(n-1)!} \quad .$$

Therefore, the expansion can be written

$$(1+x)^{-n} = \sum_{i=0}^{\infty} \binom{n+i-1}{i} x^i.$$

Consequently, for a lag-generating function with n equal roots we have

(33) $$\frac{1}{(1-\lambda L)^n} = \sum_{i=0}^{\infty} \binom{n+i-1}{i} \lambda^i L^i,$$

which agrees with our earlier formulas for the special cases n=1 and n=2.

To arrive at (33) in an alternative way, write

$$f(\lambda L) = \frac{1}{(1-\lambda L)} = \sum_{i=0}^{\infty} (\lambda L)^i.$$

Differentiating with respect to (λL) gives

$$f'(\lambda L) = \frac{1}{(1-\lambda L)^2} = \sum_{i=0}^{\infty} i(\lambda L)^{i-1} = \sum_{i=0}^{\infty} (i+1)(\lambda L)^i,$$

which is (33) with n=2. Differentiating again with respect to (λL) gives

$$f''(\lambda L) = \frac{2}{(1-\lambda L)^2} = \sum_{i=0}^{\infty} i(i-1)(\lambda L)^{i-2}$$

or

$$\frac{1}{(1-\lambda L)^3} = \sum_{i=0}^{\infty} \frac{(i+1)(i+2)}{2} (\lambda L)^i = \sum_{i=0}^{\infty} \binom{i+2}{i} (\lambda L)^i,$$

which is (33) with n=3.

With the aid of (33), the solution to (32) can be written

$$Y_t = b \sum_{i=0}^{\infty} \binom{n+i-1}{i} \lambda^i X_{t-i}$$

Using calculations like those for the first and second order cases for the special case in which $X_t = 0$ for all t, and in which n arbitrary initial values are supplied to start up the process, it is straight-forward to show that the solution obeys

$$Y_t = \lambda^t n_1 + t \lambda^t n_2 + \ldots + t^{n-1} \lambda^t n_n$$

where $n_1, \ldots, n_n$ are constants chosen to satisfy the n initial values.

The General Case: n Distinct Roots, m Identical Roots

Consider the polynomial

$$\frac{F(L)}{G(L)} = \frac{F(L)}{(1-\lambda_1 L)(1-\lambda_2 L)\ldots(1-\lambda_n L)(1-\lambda_{n+1} L)^m}$$

where the order of the numerator polynomial is less than that of the denominator; $\lambda_1, \ldots, \lambda_n$ are distinct roots while λ_{n+1} is a root that is repeated m times. Using partial fractions we write

$$(75) \qquad \frac{F(L)}{(1-\lambda_1 L)\ldots(1-\lambda_n L)(1-\lambda_{n+1} L)^m} = \frac{A_1}{1-\lambda_1 L} + \ldots + \frac{A_n}{1-\lambda_n L} + \frac{A_{n+1}}{(1-\lambda_{n+1} L)^m}.$$

Multiplying both sides by the denominator of the left side gives

$$F(L) = A_1(1-\lambda_2 L)\ldots(1-\lambda_n L)(1-\lambda_{n+1} L)^m + A_2(1-\lambda_1 L)(1-\lambda_3 1)\ldots(1-\lambda_{n+1} L)^m +$$

$$\ldots + A_n(1-\lambda_1 L)\ldots(1-\lambda_{n-1} L)(1-\lambda_{n+1} L)^m + A_{n+1}(1-\lambda_1 L)\ldots(1-\lambda_n L).$$

Evaluating the above expression successively at $L = \frac{1}{\lambda_1}$, $L = \frac{1}{\lambda_2}$, $\ldots$, $L = \frac{1}{\lambda_{n+1}}$ gives

$$A_1 = \frac{F(\frac{1}{\lambda_1})}{(1-\frac{\lambda_2}{\lambda_1})\ldots(1-\frac{\lambda_{n+1}}{\lambda_1})^m}$$

$$A_2 = \frac{F(\frac{1}{\lambda_2})}{(1-\frac{\lambda_1}{\lambda_2})(1-\frac{\lambda_3}{\lambda_2})\ldots(1-\frac{\lambda_{n+1}}{\lambda_2})^m}$$

$$\vdots$$

$$A_n = \frac{F(\frac{1}{\lambda_n})}{(1-\frac{\lambda_1}{\lambda_n})\ldots(1-\frac{\lambda_{n-1}}{\lambda_n})(1-\frac{\lambda_{n+1}}{\lambda_n})^m}$$

$$A_{n+1} = \frac{F(\frac{1}{\lambda_{n+1}})}{(1-\frac{\lambda_1}{\lambda_{n+1}})\ldots(1-\frac{\lambda_n}{\lambda_{n+1}})} \; .$$

Applying (33) to the last term of (75) then gives

$$\frac{F(L)}{G(L)} = \frac{A_1}{1-\lambda_1 L} + \ldots + \frac{A_n}{1-\lambda_n L} + \frac{A_{n+1}}{(1-\lambda_{n+1}L)^m}$$

(76) $\qquad \dfrac{F(L)}{G(L)} = A_1 \displaystyle\sum_{i=0}^{\infty} \lambda_1^i L^i + \ldots + A_n \displaystyle\sum_{i=0}^{\infty} \lambda_n^i L^i + A_{n+1} \displaystyle\sum_{i=0}^{\infty} \binom{m+i-1}{i} \lambda_{n+1}^i L^i .$

Consider the difference equation

$$(1-\lambda_1 L)\ldots(1-\lambda_n L)(1-\lambda_{n+1}L)^m Y_t = bX_t \; .$$

With the aid of (76) its solution can be written

$$Y_t = A_1 \sum_{i=0}^{\infty} \lambda_1^i X_{t-1} + \ldots + A_n \sum_{i=0}^{\infty} \lambda_n^i X_{t-i} +$$

$$A_m \sum_{i=0}^{\infty} \binom{m+i-1}{i} \lambda_{n+1}^i X_{t-i} \; .$$

For the special case in which $X_t = 0$ for all t, and in which arbitrary initial values are supplied to initiate the process, it is straightforward to show that the solution obeys

$$Y_t = \lambda_1^t \eta_1 + \lambda_2^t \eta_2 + \ldots + \lambda_n^t \eta_n + \lambda_{n+1}^t (\xi_1 + t\xi_2 + t^2 \xi_3 +$$

$$\ldots + t^{m-1} \xi_m)$$

where η_1, ..., η_n, ξ_1, ..., ξ_m are constants chosen to satisfy m+n initial conditions.

An Example of a First Order System

Consider the following model studied by Cagan.[*] Let m_t be the log of the money supply, p_t the log of the price level and p^e_{t+1} the log of the price expected to prevail at time t+1 given information available at time t. The model is

(34) $m_t - p_t = \alpha(p^e_{t+1} - p_t)$ $\alpha < 0$,

which is a portfolio equilibrium condition. The demand for real balances varies inversely with expected inflation $p^e_{t+1} - p_t$. The variable m_t is exogenous.

Suppose first that

(35) $p^e_{t+1} - p_t = \gamma(p_t - p_{t-1})$

so that the public expects inflation next period to be the current rate of inflation, $p_t - p_{t-1}$ multiplied by the constant γ. Then (34) becomes

$$m_t - p_t = \alpha\gamma\, p_t - \alpha\gamma\, p_{t-1}.$$

Using lag operators, this can be written as

$$[(\alpha\gamma+1) - \alpha\gamma L]\, p_t = m_t$$

or

$$[1 - \frac{\alpha\gamma}{1+\alpha\gamma} L]\, p_t = \frac{1}{1+\alpha\gamma}\, m_t.$$ The solution can be written

_____

[*]"The Monetary Dynamics of Hyperinflation"

$$P_t = \frac{1}{1+\alpha\gamma} \sum_{i=0}^{\infty} (\frac{\alpha\gamma}{1+\alpha\gamma})^i m_{t-i},$$

which will be finite for the time path $m_t = \bar{m}$ for all t, provided that

$$\left|\frac{\alpha\gamma}{1+\alpha\gamma}\right| < 1.$$

The above inequality is in the spirit of the "stability condition" developed by Cagan in his paper. It is a condition that delivers a finite p_t for all t for a certain time path of m. Notice that

$$\frac{1}{1+\alpha\gamma} \sum_{i=0}^{\infty} (\frac{\alpha\gamma}{1+\alpha\gamma})^i = \frac{\frac{1}{1+\alpha\gamma}}{1 - \frac{\alpha\gamma}{1+\alpha\gamma}} = 1.$$

Thus, the long-run effect of a once-and-for-all jump in m is to drive p up by an equal amount (provided the above "stability condition" is met).

Returning to (34), let us abandon (35) and now assume perfect foresight:

(36) $\qquad p_{t+1}^e = p_{t+1}.$

Substituting (36) into (34) gives

$$m_t - p_t = \alpha p_{t+1} - \alpha p_t$$

or $\qquad \alpha p_{t+1} + (1-\alpha) p_t = m_t.$

Write this as

$$(L^{-1} + \frac{1-\alpha}{\alpha}) p_t = \frac{1}{\alpha} m_t$$

or

(37) $\qquad (1 - \frac{\alpha-1}{\alpha} L) p_t = \frac{1}{\alpha} m_{t-1}$

Notice that since $\alpha < 0$, it follows that $\frac{\alpha-1}{\alpha} > 1$. This fact is an invitation to solve (37) in the "forward" direction, that is, to use (6). Dividing both sides of (37) by $(1 - \frac{\alpha-1}{\alpha} L)$ gives

$$p_t = (\frac{\frac{1}{\alpha}}{1 - \frac{\alpha-1}{\alpha}L}) \, m_{t-1},$$

which using (6) becomes

$$p_t = (\frac{\frac{1}{\alpha} \cdot -(\frac{\alpha}{\alpha-1})L^{-1}}{1 - \frac{\alpha}{\alpha-1} L^{-1}}) \, m_{t-1} = -\frac{1}{\alpha-1} (\sum_{i=0}^{\infty} (\frac{\alpha}{\alpha-1})^i L^i) \, m_t$$

(38) $$p_t = \frac{1}{1-\alpha} \cdot \sum_{i=0}^{\infty} (\frac{\alpha}{\alpha-1})^i \, m_{t+i}$$

Notice that since $\alpha < 0$, $0 < \frac{\alpha}{\alpha-1} < 1$, so that the sum of the lag weights is finite. Equation (38) expresses the log of the current price as a moving sum of current and __future__ values of the log of the money supply. Notice that

$$\frac{1}{1-\alpha} \sum_{i=0}^{\infty} (\frac{\alpha}{\alpha-1})^i = \frac{\frac{1}{1-\alpha}}{1 - \frac{\alpha}{\alpha-1}} = 1,$$

so that p is a weighted __average__ of current and future values of m.

An Example of a Second Order System

Consider the following model studied by Muth.[*] Let p_t be the price of a commodity at t, C_t the demand for current consumption, I_t the stock of inventories of the commodity, Y_t the output of the commodity, and p_t^e the price previously expected to prevail at time t;

---

[*]"Rational Expectations and the Theory of Price Movements."

X_t represents the effects of the weather on supply. The model is

$$C_t = -\beta \, p_t, \qquad \beta > 0 \qquad \text{demand curve}$$

$$Y_t = \gamma p_t^e + X_t, \qquad \gamma > 0 \qquad \text{supply curve}$$

$$I_t = \alpha(p_{t+1}^e - p_t) \qquad \alpha > 0 \qquad \text{inventory demand}$$

$$Y_t = C_t + (I_t - I_{t-1}) \qquad \text{market clearing}$$

Let us suppose that there is perfect foresight so that $p_t^e = p_t$ for all t. Making this assumption and substituting the first three equations into the fourth gives

$$\gamma \, P_t + X_t = \alpha(p_{t+1} - p_t) - \alpha(p_t - p_{t-1}) - \beta \, p_t$$

or

$$\gamma \, P_{t+1} - (2\alpha + \beta + \gamma) \, p_t + \alpha p_{t-1} = X_t.$$

Dividing by α gives

$$\left(p_{t+1} - \frac{(2\alpha + \beta + \gamma)}{\alpha} \, p_t - \alpha p_{t-1} \right) = \frac{1}{\alpha} X_t$$

or

$$(L^{-1} - \phi + L) \, p_t = \frac{1}{\alpha} X_t$$

where $\phi = \dfrac{\beta + \gamma}{\alpha} + 2 > 0$. Multiplying by L gives

$$(39) \qquad (1 - \phi L + L^2) \, p_t = \frac{1}{\alpha} X_{t-1}.$$

We need to factor the polynomial $(1 - \phi L + L^2)$ as

$$(1-\phi L + L^2) = (1-\lambda_1 L)(1-\lambda_2 L)$$

$$= (1-(\lambda_1+\lambda_2)L + \lambda_1\lambda_2 L^2)$$

so that we require that

$$\lambda_1 + \lambda_2 = \phi$$

$$\lambda_1 \lambda_2 = 1.$$

The second equality establishes that $\lambda_1 = 1/\lambda_2$,

so that the two roots appear as a reciprocal pair. So we can write

$$(1-\phi L + L^2) = (1-\lambda L)(1- \frac{1}{\lambda}L)$$

where λ is chosen to satisfy $\lambda + \frac{1}{\lambda} = \phi$.

So (39) can be written

$$(39') \qquad (1-\lambda L)(1- \frac{1}{\lambda}L) \; p_t = \frac{1}{\alpha} X_{t-1}.$$

Since $\frac{B+\gamma}{\alpha} > 0$, it follows that $\lambda = \frac{B+\gamma}{\alpha} + 2 > 2$. That implies that

λ does not equal 1, since $\lambda + 1/\lambda = \phi$. Notice that if $\lambda > 1$, $1/\lambda < 1$.

So one of our roots necessarily exceeds 1, the other necessarily is

less than 1.

We divide both sides of (39') by $(1-\lambda L)(1-(1/\lambda)L)$ to obtain

$$p_t = \frac{1}{\alpha} \frac{1}{(1-\lambda L)(1- \frac{1}{\lambda}L)} X_{t-1}.$$

Without loss of generality, suppose $\lambda < 1$ and let $\lambda_2 = 1/\lambda$. Use

(6) and (19) to write the solution as

$$
p_t = \frac{\frac{1}{\alpha}\lambda}{\lambda-\lambda_2} \left(\frac{1}{1-\lambda L}\right) X_{t-1} - \frac{\frac{1}{\alpha}\lambda_2}{(\lambda-\lambda_2)} \left(\frac{-\frac{1}{\lambda_2 L}}{1-\frac{1}{\lambda_2}L^{-1}}\right) X_{t-1}
$$

$$
= \frac{\frac{1}{\alpha}\lambda}{\lambda-\frac{1}{\lambda}} \left(\frac{1}{1-\lambda L}\right) X_{t-1} + \frac{\frac{1}{\alpha}\cdot\frac{1}{\lambda}}{\lambda-\frac{1}{\lambda}} \left(\frac{\lambda L^{-1}}{1-\lambda L^{-1}}\right) X_{t-1}
$$

$$
= \frac{\frac{1}{\alpha}\lambda}{\lambda-\frac{1}{\lambda}} \left(\frac{1}{1-\lambda L}\right) X_{t-1} + \frac{\frac{1}{\alpha}}{\lambda-\frac{1}{\lambda}} \left(\frac{1}{1-\lambda L^{-1}}\right) X_t
$$

$$
p_t = \frac{\frac{1}{\alpha}}{\lambda-\frac{1}{\lambda}} \sum_{i=1}^{\infty} \lambda^i X_{t-i} + \frac{\frac{1}{\alpha}}{\lambda-\frac{1}{\lambda}} \sum_{i=0}^{\infty} \lambda^i X_{t+i}
$$

$$
(40) \qquad p_t = \frac{\frac{1}{\alpha}}{\lambda-\frac{1}{\lambda}} \sum_{i=-\infty}^{\infty} \lambda^{|i|} X_{t-i}
$$

The solution (40) expresses p_t as a "two-sided" distributed lag of X, that is, as a weighted sum of past, present, and future values of X. In this model, the current price depends on the entire path of the exogenous shock X over the entire past _and_ the entire future.

Naive Business Cycle Theory

Deterministic (nonrandom) difference operations of low order
can generate "cycles," but not of the kind ordinarily thought to charac-
terize economic variables. For example, we have seen that second order
difference equations can generate cycles of constant periodicity that
are damped, explosive, or, in the very special case where the amplitude
$r = 1$, of constant-amplitude. But the "cycles" in economic variables
seem neither damped nor explosive, and they don't have a constant period
from one cycle to the other; e.g., some recessions last one year, some
last for one and a half years. The "business cycle" is the tendency of
certain economic variables to possess persistent cycles of approximately
constant amplitude and somewhat irregular periodicity from one "cycle"
to the other. The National Bureau of Economic Research has inspected
masses of data indicating the presence of a business cycle of average
length of about three years from peak to peak in many important economic
aggregates for the U.S.

Figure 1 graphs the 91 day Treasury Bill rate and the unemploy-
ment rate over the postwar period for quarterly data. The "business
cycle" shows up in both series, interest rates tending to be high and
unemployment low in "booms," and interest rates tending to be low and
unemployment high in recessions. Clearly the "cycles" are irregular in
length and don't "look like" those generated by our low order difference
equations.

While low order deterministic difference equations don't
provide an adequate model for explaining the cycles in economic data,
low order stochastic or random difference equations do. If the initial

condition of a deterministic difference equation is subjected to repeated

random shocks of a certain kind, there emerges the possibility of

persistent cycles of the kind seemingly infesting economic data. This

is an important idea in macroeconomics, and owes its origin to Slutsky

and Frisch. These notes sketch the elements of that idea.

A basic building block is the serially uncorrelated random

process ε_t, which satisfies

la $E(\varepsilon_t) = 0$ all t

lb $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ all t

lc $\text{cov} (\varepsilon_t, \varepsilon_{t-s}) = E(\varepsilon_t \varepsilon_{t-s}) = 0$ all t and all s $\neq$ 0

where E is the mathematical expectations operator. According to (1),

the mean of ε_t, which is zero for all t, and the variance of ε_t, which

is σ_ε^2 for all t, both are independent of time. According to (1c), ε_t

is uncorrelated (i.e., has zero covariance) with itself lagged s $= \pm 1$,

$\pm 2, \ldots$ times and is said to be "serially uncorrelated." The variate ε_t

is also said to be a "white noise." The schedule of covariances

$E(\varepsilon_t \varepsilon_{t-s})$ for s $= 0, \pm 1, \pm 2, \ldots$ is a function only of s, and not of t,

a characteristic called "covariance stationarity." The schedule of

covariances $E(\varepsilon_t \varepsilon_{t-s})$ is called the "covariogram" of the ε process.

Notice that the covariogram, viewed as a function of s, is symmetric

about zero since $E(\varepsilon_t \varepsilon_{t-s}) = E(\varepsilon_t \varepsilon_{t+s})$, an implication of $E(\varepsilon_t \varepsilon_{t-s})$

depending only on s and not on t.

Now consider the random process y_t defined by

(a) $y_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}$
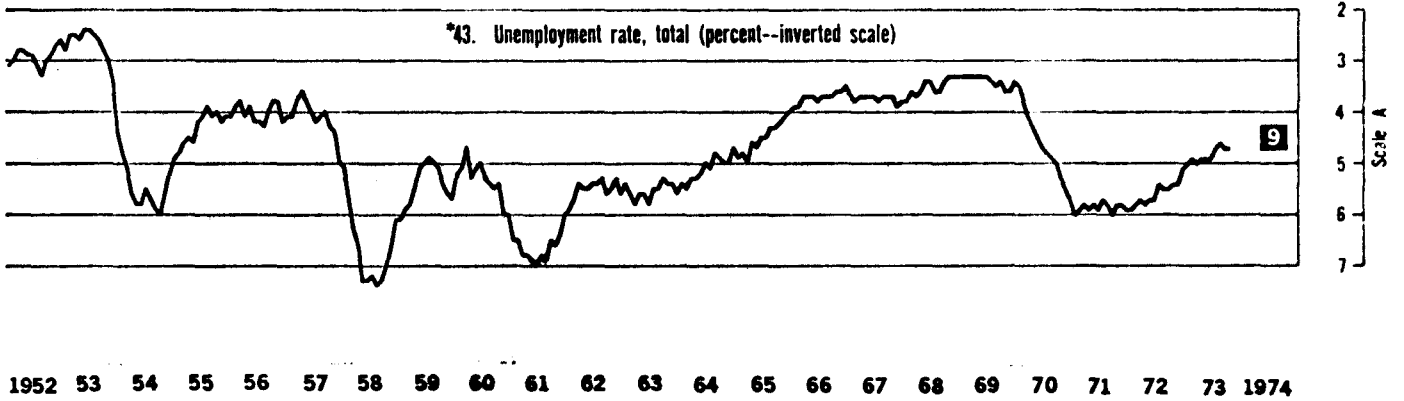
           $= B(L)\varepsilon_t$

FIGURE 1.

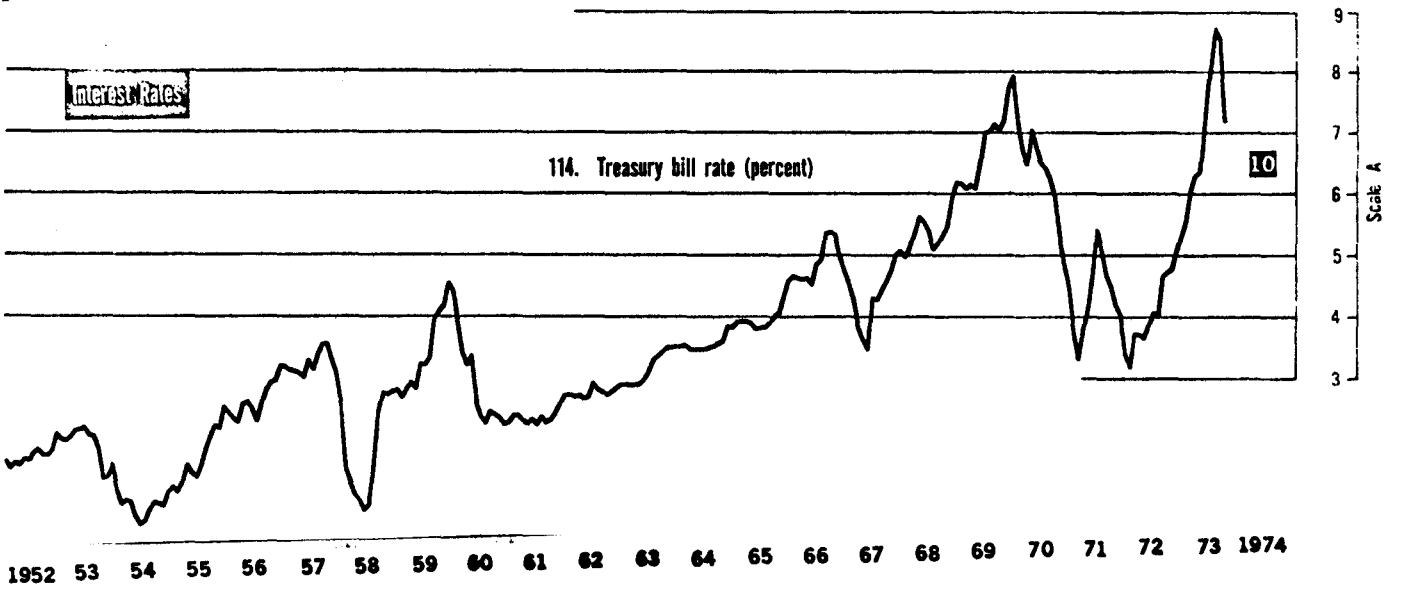| (July) | (Aug.) | | (July) | (Apr.) | | (May) | (Feb.) | | (Nov.) | (Nov.) |
| P | T | | P | T | | P | T | | P | T |

Comprehensive Unemployment

*43. Unemployment rate, total (percent--inverted scale)

9

Scale A

1952 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 1974

Interest Rates

114. Treasury bill rate (percent)

10

Scale A

1952 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 1974

where $B(L) = \sum\limits_{j=0}^{\infty} b_j L^j$, and where we assume $\sum\limits_{j=0}^{\infty} b_j^2 < \infty$,

a requirement needed to assure that the variance of y is finite. We

assume that the ε process is "white" and thus satisfies properties (1).

Equation (2) says that the y process is a one-sided moving sum of a

white noise process, ε.

We seek the covariogram of the y process, i.e., we seek the

values of $c_y(k) = E(y_t y_{t-k})$ for all k. It will be convenient to obtain

the "covariance generating function" $g_y(z)$ which is defined by

$$(3) \qquad g_y(z) = \sum_{k=-\infty}^{\infty} c_y(k) z^k.$$

The coefficient on z^k in (3) is the kth lagged covariance, $c_y(k)$.

First notice that taking mathematical expectations on both

sides of (2) gives

$$E(y_t) = \sum_{j=0}^{\infty} b_j E(\varepsilon_{t-j})$$

$$= 0 \qquad \text{for all t.}$$

It therefore follows that

$$c_y(k) = E\{(y_t - Ey_t)(y_{t-k} - Ey_{t-k})\}$$

$$= Ey_t \cdot y_{t-k} \qquad \text{for all k.}$$

Notice $y_t \cdot y_{t-k}$ is

$$y_t y_{t-k} = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j} \sum_{b=0}^{\infty} b_h \varepsilon_{t-k-h}$$

$$= (b_0 \varepsilon_t + b_1 \varepsilon_{t-1} + b_2 \varepsilon_{t-2} + \ldots)(b_0 \varepsilon_{t-k} + b_1 \varepsilon_{t-k-1} + b_2 \varepsilon_{t-k-2} + \ldots)$$

$$y_r y_{t-k} = \{b_0 b_k \varepsilon_{t-k}^2 + b_1 b_{k+1} \varepsilon_{t-k-1}^2 + b_2 b_{k+2} \varepsilon_{t-k-2}^2 + \dots \} +$$

crossproduct terms whose expectations are zero.

Thus

$$(4) \qquad c_y(k) = E y_t y_{t-k} = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j b_{j+k}.$$

The covariance generating function is then

$$g_y(z) = \sum_{k=-\infty}^{\infty} z^k c_y(k)$$

$$= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} z^k \sum_{j=0}^{\infty} b_j b_{j+k}$$

$$= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} b_j b_{j+k} z^k$$

$$g_y(z) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} b_j b_{j+k} z^k.$$

Let $h = j + k$, so that $k = h - j$. Writing the above line in terms of the index h then gives

$$g_y(z) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} b_j b_h z^{h-j}$$

$$= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j z^{-j} \sum_{h=0}^{\infty} b_h z^h.$$

The last equation gives the convenient expression

$$(5) \qquad g_t(z) = \sigma_\varepsilon^2 B(z^{-1}) B(z)$$

where $B(z^{-1}) = \sum_{j=0}^{\infty} b_j z^{-j}$, $B(z) = \sum_{j=0}^{\infty} b_j z^j$.

Equation (5) gives the covariance generating function $g_y(z)$ in terms of the b_j's of (2) and the variance σ_ε^2 of the white noise ε.

To take an example that illustrates the usefulness of (5), consider the first order process

$$
(6) \qquad y_t = \left(\frac{1}{1 - \lambda L}\right)\varepsilon_t = \sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i}, \qquad |\lambda| < 1
$$

where, as always, ε is a white noise process with variance σ_ε^2. We have

$$
B(L) = \frac{1}{1 - \lambda L} ,
$$

$$
B(z) = \frac{1}{1 - \lambda z} = 1 + \lambda z + \lambda^2 z^2 + \ldots
$$

$$
B(z^{-1}) = \frac{1}{1 - \lambda z^{-1}} = 1 + \lambda z^{-1} + \lambda^2 z^{-2} + \ldots.
$$

(Thus, $B(z)$ is found by replacing L in $B(L)$ by z.) So applying (5), we have

$$
(7) \qquad g_y(z) = \sigma_\varepsilon^2 \left(\frac{1}{1 - \lambda z^{-1}}\right)\left(\frac{1}{1 - \lambda z}\right) .
$$

From our experience with difference equations we know that the expression (7) can be written as a sum

$$
(8) \qquad g_y(z) = \frac{k_1 \sigma_\varepsilon^2}{1 - \lambda z} + \frac{k_2 \sigma_\varepsilon^2 z^{-1}}{1 - \lambda z^{-1}}
$$

where k_1 and k_2 are certain constants. To find out what the constants must be, notice that (8) implies

$$
g_y(z) = \sigma_\varepsilon^2 k_1 (1 + \lambda z + \lambda^2 z^2 + \ldots)
$$

$$
+ \sigma_\varepsilon^2 k_2 (z^{-1} + \lambda z^{-2} + \lambda^2 z^{-3} + \ldots),
$$

so that $c_y(0) = k_1 \sigma_\varepsilon^2$ and $c_y(1) = \sigma_\varepsilon^2 \lambda k_1 = \sigma_\varepsilon^2 k_2$.

By direct computation using (6) we note that

$$Ey_t^2 = \sum_{i=0}^{\infty} \lambda^{i2} E\varepsilon_t^2 = \frac{\sigma_\varepsilon^2}{1 - \lambda^2}$$

$$Ey_t y_{t-1} = E\sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i} \sum_{i=1}^{\infty} \lambda^{i-1} \varepsilon_{t-i} = E\sum_{i=1}^{\infty} \lambda^i \lambda^{i-1} \varepsilon_{t-i}^2$$

$$= \sigma_\varepsilon^2 \lambda \sum_{i=1}^{\infty} \lambda^{(i-1),2} = \frac{\sigma_\varepsilon^2}{1 - \lambda^2} \; .$$

So for (8) to be correct, we require that

$$k_1 = \frac{1}{1 - \lambda^2}$$

$$k_2 = \frac{\lambda}{1 - \lambda^2} \; .$$

With these values of k_1 and k_2, we can verify directly that

$$\sigma_\varepsilon^2 \left[\frac{\frac{1}{1-\lambda^2}}{1-\lambda z} + \frac{z^{-1}\left(\frac{\lambda}{1-\lambda^2}\right)}{1-\lambda z^{-1}} \right]$$

$$= \sigma_\varepsilon^2 \cdot \frac{1}{1-\lambda^2} \; \left[\frac{(1-\lambda z^{-1}) + \lambda z^{-1} - \lambda^2}{(1-\lambda z)(1-\lambda z^{-1})} \right]$$

$$= \sigma_\varepsilon^2 \frac{1}{(1-\lambda z)(1-\lambda z^{-1})} \; ,$$

so that (8) and (7) are equivalent.

Expression (8) is the more convenient of the two expressions since it yields quite directly,

$$g_y(z) = \sigma_\varepsilon^2 \frac{1}{1-\lambda^2} \left[\frac{1}{1-\lambda z} - \frac{\lambda z^{-1}}{1-\lambda z^{-1}} \right]$$

(9) $\qquad = \sigma_\epsilon^2 \dfrac{1}{1-\lambda^2} \;[\{1+\lambda z+\lambda^2 z^2+\ldots\} + \{\lambda z^{-1}+\lambda^2 z^{-2}+\lambda^3 z^{-3}+\ldots\}\;].$

Thus, we have that for the "geometric" process (6),

$$c_y(k) = \frac{\sigma_\epsilon^2}{1-\lambda^2} \cdot \lambda^{|k|} \qquad\qquad k=0,\pm1,\pm2,\ldots$$

The covariance declines geometrically with increases in $|k|$. We require $|\lambda|<1$ in order that the y process have a finite variance.

To get this result more directly multiply y_t by y_{t-k}, $k > 0$, to obtain

$$y_t y_{t-k} = \lambda y_{t-1} y_{t-k} + \epsilon_t y_{t-k}.$$

Taking expected values on both sides and noting that $E\epsilon_t y_{t-k} = 0$ gives

$$E(y_t y_{t-k}) = \lambda E(y_{t-1} y_{t-k})$$

or

$$c_y(k) = \lambda c_y(k-1) \qquad\qquad k > 0$$

which implies

$$c_y(k) = \lambda^k c_y(0) \qquad\qquad k > 0$$

As a second example, consider the second-order process

$$(10) \qquad y_t = \left(\frac{1}{1-\lambda_1 L}\right)\left(\frac{1}{1-\lambda_2 L}\right)\epsilon_t, \qquad |\lambda_1+\lambda_2| < 1$$

where ϵ_t is white noise with variance σ_ϵ^2.

For (10) we have

$$B(L) = \left(\frac{1}{1-\lambda_1 L}\right)\left(\frac{1}{1-\lambda_2 L}\right)$$

$$B(z) = \left(\frac{1}{1-\lambda_1 z} \right) \left(\frac{1}{1-\lambda_2 z} \right)$$

$$B(z^{-1}) = \left(\frac{1}{1-\lambda_1 z^{-1}} \right) \left(\frac{1}{1-\lambda_2 z^{-1}} \right).$$

Applying formula (5), we have that the covariance generating function is

$$(11) \qquad g_y(z) = \sigma_\epsilon^2 \frac{1}{(1-\lambda_1 z)} \frac{1}{(1-\lambda_2 z)} \frac{1}{(1-\lambda_1 z^{-1})} \frac{1}{(1-\lambda_2 z^{-1})} \;.$$

Notice that (10) can be written

$$y_t = \frac{\lambda_1}{\lambda_1 - \lambda_2} \left(\frac{1}{1-\lambda_1 L} \right) \epsilon_t - \left(\frac{\lambda_2}{\lambda_1 - \lambda_2} \right) \left(\frac{1}{1-\lambda_2 L} \right) \epsilon_t$$

$$(12) \qquad y_t = \frac{\lambda_1}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_1^i \epsilon_{t-i} - \frac{\lambda_2}{\lambda_1 - \lambda_2} \sum_{i=0}^{\infty} \lambda_2^i \epsilon_{t-i} \;.$$

For y_{t-k}, $k \geq 0$, we have

$$(13) \qquad y_{t-k} = \frac{\lambda_1}{\lambda_1 - \lambda_2} \sum_{i=k}^{\infty} \lambda_1^{i-k} \epsilon_{t-i} - \frac{\lambda_2}{\lambda_1 - \lambda_2} \sum_{i=k}^{\infty} \lambda_2^{i-k} \epsilon_{t-i} \;.$$

Multiplying (12) and (13) together and taking expectations gives

$$E(y_t y_{t-k}) = \sigma_\epsilon^2 \left\{ \frac{\lambda_1^2}{(\lambda_1 - \lambda_2)^2} \sum_{i=0}^{\infty} \lambda_1^{k+i} \lambda_1^i + \frac{\lambda_2^2}{(\lambda_1 - \lambda_2)^2} \sum_{i=0}^{\infty} \lambda_2^{k+i} \lambda_2^i \right.$$

$$\left. - \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2} \sum_{i=0}^{\infty} \lambda_1^{k+i} \lambda_2^i - \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2} \sum_{i=0}^{\infty} \lambda_2^{k+i} \lambda_1^i \right\}$$

$$(14) \qquad E y_t y_{t-k} = \left(\frac{1}{\lambda_1 - \lambda_2} \right)^2 \sigma_\epsilon^2 \left[\frac{\lambda_1^{2+k}}{(1-\lambda_1^2)} + \frac{\lambda_2^{2+k}}{(1-\lambda_2^2)} - \frac{\lambda_1 \lambda_2}{1-\lambda_1 \lambda_2} (\lambda_1^k + \lambda_2^k) \right]$$

$$k \geq 0.$$

So (14) and the symmetry of $g_y(z)$ suggests that the appropriate factorization of (11) is

(15) $\quad g_y(z) = \left(\frac{1}{\lambda_1 - \lambda_2}\right)^2 \sigma_\varepsilon^2 \left\{ \left(\frac{\lambda_1^2}{(1-\lambda_1^2)} - \frac{\lambda_1 \lambda_2}{1-\lambda_1 \lambda_2}\right) \left(\frac{1}{1-\lambda_1 z} + \frac{\lambda_1 z^{-1}}{1-\lambda_1 z^{-1}}\right) \right.$

$$+ \left(\frac{\lambda_2^2}{1-\lambda_2^2} - \frac{\lambda_1 \lambda_2}{1-\lambda_1 \lambda_2}\right) \left(\frac{1}{1-\lambda_2 z} + \frac{\lambda_2 z^{-1}}{1-\lambda_2 z^{-1}}\right) \Big\} \ .$$

According to (14) and (15) the covariogram of a y process governed by the second-order process (10) consists of a weighted sum of two geometric decay processes, the decay parameters being λ_1 and λ_2, the inverse roots of the polynomial $(1-\lambda_1 L)(1-\lambda_2 L)$. Expression (14) implies that the covariogram displays damped oscillations if the roots λ_1 and λ_2 are complex conjugates. This can be shown by substituting $\lambda_1 = re^{-iw}$ and $\lambda_2 = re^{iw}$ into (14), and proceeding to analyze (14) as we above analyzed the solution of the deterministic (nonrandom) second order difference equation. An alternative way to reach the same conclusion is as follows. Multiply both sides of (10) by $(1-\lambda_1 L)(1-\lambda_2 L)$ to get

(16) $\quad y_t = t_1 y_{t-1} + t_2 y_{t-2} + \varepsilon_t$

where $t_1 = (\lambda_1 + \lambda_2)$ and $t_2 = -\lambda_1 \lambda_2$. Multiply (16) by y_{t-k} for $k \geq 0$ to get

$$y_t y_{t-k} = t_1 y_{t-1} y_{t-k} + t_2 y_{t-2} y_{t-k} + \varepsilon_t y_{t-k}.$$

Since $E \varepsilon_t y_{t-k} = 0$, we have

$$E(y_t y_{t-k}) = t_1 E(y_{t-1} y_{t-k}) + t_2 (y_{t-2} y_{t-k}) \qquad k \geq 0$$

which shows that $c_y(k)$ obeys the difference equation

(17) $\quad c_y(k) = t_1 c_y(k-1) + t_2 c_y(k-2).$

So the covariogram of a second (n^{th}) order process obeys the solution to

the deterministic second (n^{th}) order difference equation examined above.

In particular, corresponding to (17) we consider the polynomial

(18) $1 - t_1 k - t_2 k^2 = 0,$

which has roots $1/\lambda_1$ and $1/\lambda_2$. (We know that $1-t_1 k-t_2 k$ equals

$(1-\lambda_1 k)(1-\lambda_2 k)$, with roots $1/\lambda_1$ and $1/\lambda_2$.) Alternatively, multiply (18)

by k^{-2} to obtain

$$k^{-2} - t_1 k^{-1} - t_2 = 0$$

(19) $x^2 - t_1 x - t_2 = 0$ where $x = k^{-1}$.

Notice that the roots of (19) are the reciprocals of the roots of (18),

so λ_1 and λ_2 are the roots of (19).

The solution to the deterministic difference equation (17) is,

as we have seen,

(20) $c_y(k) = \lambda_1^k z_0 + \lambda_2^k z_1,$ $k \geq 0$

where z_0 and z_1 are certain constants chosen to make $c_y(0)$ and $c_y(1)$

equal the proper quantities. If the roots λ_1 and λ_2 are complex, we

know from our work with deterministic difference equations that (20)

becomes

(21) $\qquad c_y(k) = z_0 \dfrac{r^k}{\sin w} \sin wk + z_1 \dfrac{r^k}{\sin w} \cos wk$

where $\lambda_1 = re^{iw}$ and $\lambda_2 = re^{-iw}$. Accordingly to (21), the covariogram displays damped (we require $r < 1$) oscillations with angular frequency w. A complete cycle occurs as wk goes from zero ($k=0$) to 2π ($k=2\pi/w$, if that is possible). So the cycles in the covariogram occur with period from peak to peak of $2\pi/w$ periods. The restrictions on t_1 and t_2 needed to deliver complex roots and so an oscillatory covariogram can be read directly from Figure 1 of "Notes on Difference Equations."

Figure 1 displays some realizations of second order processes for various values of t_1 and t_2, values for which the roots are complex. Notice the tendency of these series to cycle, but with a periodicity that is somewhat variable from cycle to cycle.

The foregoing suggests one definition of the business cycle: a series may be said to possess a "cycle" if its covariogram is characterized by oscillations. The typical "length" of the cycle can be measured by the number of periods it takes for the covariogram to experience one full cycle. To be labelled a business cycle the cycle should exceed a year in length. (Cycles of one year in length are termed "seasonals.")

The Spectrum

An alternative to the preceding definition of the business cycle is based on the spectrum of the y process. Recall the covariance generating function of y defined in (3),

$$(3) \qquad g_y(z) = \sum_{k=-\infty}^{\infty} c_y(k) z^k \ .$$

For the process $y_t = B(L)\varepsilon_t$ we have seen that

$$g_y(z) = B(z)B(z^{-1}) \sigma_\varepsilon^2 \ .$$

If we evaluate (3) at the value $z = e^{-iw}$, we have

$$(22) \qquad g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) e^{-iwk} \qquad -\pi < w < \pi \ .$$

Viewed as a function of angular frequency w, $g_y(e^{-iw})$ is called the spectrum of y.

The spectrum $g_y(e^{-iw})$ is itself a covariance generating function, which is hardly surprising. Given an expression for $g_y(e^{iw})$, it is easy to recover the covariances $c_y(k)$. To see this, we multiply (22) by e^{iwh} and integrate with respect to w from $-\pi$ to π:

$$(23) \qquad \int_{-\pi}^{\pi} g_y(e^{-iw}) e^{iwh} dw = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} c_y(k) e^{iw(h-k)} dw$$

$$= \sum_{k=-\infty}^{\infty} c_y(k) \int_{-\pi}^{\pi} e^{iw(h-k)} dw.$$

Now for $h = k$ we have

$$\int_{-\pi}^{\pi} e^{iw(h-k)} dw = \int_{-\pi}^{\pi} 1 dw = 2\pi.$$

For $h \neq k$ we have,

$$\int_{-\pi}^{\pi} e^{iw(h-k)} dw = \int_{-\pi}^{\pi} \cos w \ (h-k) dw + i \int_{-\pi}^{\pi} \sin w \ (h-k) \ dw$$

$$= -\sin w(h-k)]_{-\pi}^{\pi} + i \cos w \ (h-k)]_{-\pi}^{\pi}$$

$$= 0.$$

Therefore (23) becomes

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) e^{iwh} dw = 2\pi c_y(h).$$

Thus multiplying the spectrum by e^{iwh} and integrating from $-\pi$ to π gives the hth lagged covariance times 2π. In particular, notice that for $h = 0$, we have

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) dw = 2\pi c_y(0),$$

so that the area under the spectrum from $-\pi$ to π equals 2π times the variance of y. This fact motivates the interpretation of the spectrum as a device for decomposing the variance of a series by frequency. The portion of the variance of the series occurring between any two frequencies is given by the area under the spectrum between those two frequencies.

Notice that from (22) we have

$$g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) \ e^{-iwk}$$

$$(34) \qquad = c_y(0) + \sum_{k=1}^{\infty} c_y(k)(e^{iwk} + e^{-iwk})$$

$$= c_y(0) + 2 \sum_{k=1}^{\infty} c_y(k) \cos wk.$$

According to (34) the spectrum is real valued at each frequency, and is obtained by multiplying the covariogram of y by a cosine function of the frequency in question. Notice also that since $\cos x = \cos -x$, it follows from (34) that

$$g_y(e^{iw}) = g_y(e^{-iw}),$$

so that the spectrum is symmetric about $w=0$.

Notice also that since $\cos(w+2\pi k) = \cos(w)$, $k=0, \pm 1$, $\pm 2, \ldots$, it follows that the spectrum is a periodic function of w with period 2π. Therefore we can confine our attention to the interval $[-\pi, \pi]$, or even $[0, \pi]$ by virtue of the symmetry of the spectrum about $w=0$.

The fact that the spectrum can be viewed as decomposing the variance of a series by frequency motivates our second definition of the business cycle. A series is said to display a cycle of a given periodicity if its spectrum possesses a peak at that periodicity. A series displays a "business cycle" if that periodicity is of about three years. If a peak occurs in a spectrum at a certain frequency, it indicates that a relatively large amount of the variance of the series occurs (can be explained by cosine functions) at that frequency. The sharper is the peak in the spectrum, the more regular are the cycles occurring in the series.

To motivate further the interpretation of the spectrum as a decomposition of variance by frequency, suppose that we have T observations on y_t, $t=0, 1, \ldots, T-1$. Suppose that T is an even number. We consider computing the following regression of y_t on sine and cosine functions of angular frequency $w_j = \frac{2\pi j}{T}$ where $j=0, 1, \ldots, T/2$:
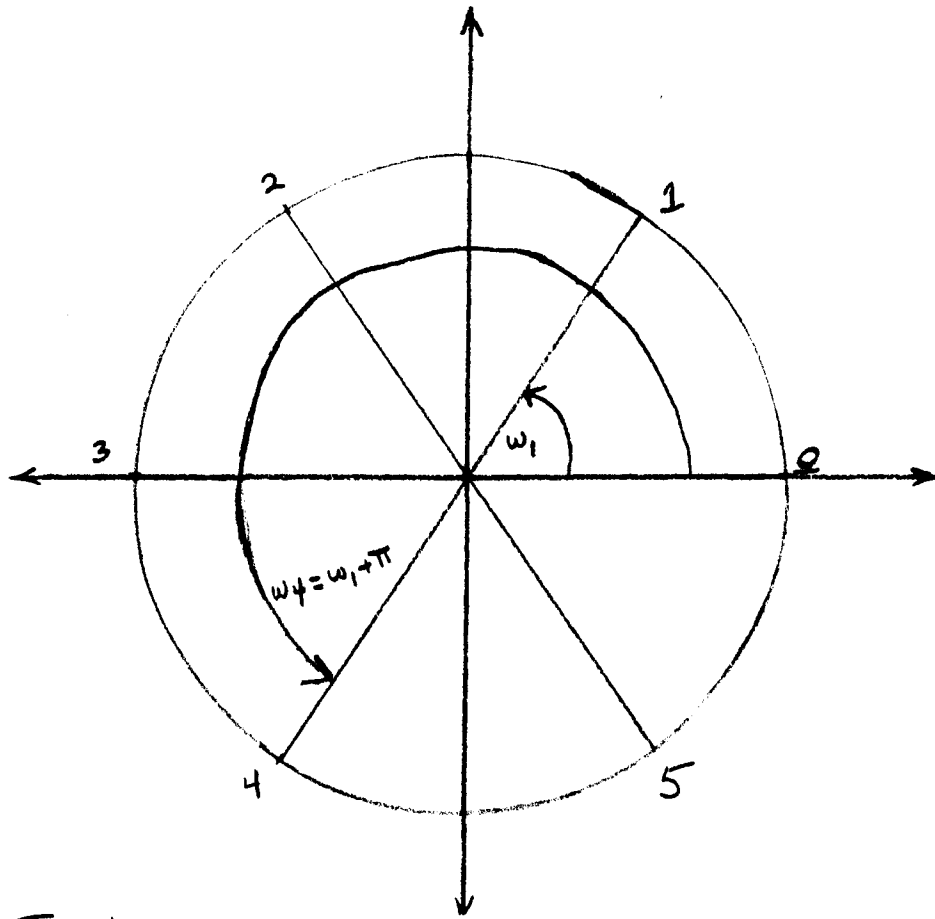
$T = 6$

$$(25) \qquad y_t = \sum_{k=0}^{T/2} \hat{\alpha}(w_k) \cos w_k t + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k) \sin w_k t, \quad w_k = \frac{2\pi k}{T} .$$

There are T observations and T independent variables in (25). The independent variables of (25) are mutually orthogonal. For we know that

$$\cos w_k t \cos w_j t + \sin w_k t \sin w_j t = \cos (w_k - w_j)t$$

$$\cos w_k t \cos w_j t - \sin w_k t \sin w_j t = \cos (w_k + w_j)t .$$

Summing both equations and adding we have for $j \neq k$

$$(26) \qquad 2 \sum_{t=0}^{T-1} \cos w_k t \cos w_j t = \sum_{t=0}^{T-1} \cos (w_k - w_j)t + \sum_{t=0}^{T-1} \cos (w_k + w_j)t$$

$$= 0,$$

since the angles $(w_k + w_j)t = \dfrac{2\pi(j+k)t}{T}$, $t=0$, 1, ..., T-1 are spaced evenly about the circle in the fashion depicted in Figure ___. The angles appear in pairs, w', $w' + \pi$, so that for each cosine in the sum of angle w', there is another offsetting cosine associated with the angle $w' + \pi$. From (26) it follows that

$$\sum_{t=0}^{T-1} \cos w_k t \cos w_j t = 0, \qquad\qquad k \neq j$$

so that $\cos w_k t$ and $\cos w_j t$ are orthogonal. In a similar fashion, it can be shown that

$$\sum_{t=0}^{T-1} \sin w_k t \cos w_j t = \sum_{t=0}^{T-1} \sin w_k t \sin w_j t = 0 \text{ for } j \neq k,$$

so that the independent variables are mutually orthogonal.

Where the independent variables are orthogonal, the (multivariate) least squares estimator of the regression coefficients is identical with the vector of simple least squares estimates. These are given by

$$\hat{\alpha}(w_k) = \frac{\displaystyle\sum_{t=0}^{T-1} y_t \cos w_k t}{\displaystyle\sum_{t=0}^{T-1} \cos^2 w_k t} \qquad k=0,1,\ldots T/2$$

(27)

$$\hat{\beta}(w_k) = \frac{\displaystyle\sum_{t=0}^{T-1} y_t \sin w_k t}{\displaystyle\sum_{t=0}^{T-1} \sin^2 w_k t} \; . \qquad k=1,2,\ldots T/2-1$$

Notice that

$$\sum_{t=0}^{T-1} \cos^2 w_0 t = \sum_{t=0}^{T-1} 1 = T$$

$$\sum_{t=0}^{T-1} \cos^2 w_{T/2} = \sum_{t=0}^{T-1} \cos^2 (\pi t) = T \; .$$

and for $k = 1, 2, \ldots, T/2 - 1$

$$\sum_{t=0}^{T-1} \cos^2 w_k t = \sum_{t=0}^{T-1} (\cos 2 w_k t + \sin^2 w_k t) = \sum_{t=0}^{T-1} \sin^2 w_k t$$

$$= \sum_{t=0}^{T-1} (1-\sin^2 w_k t)$$

which implies that

$$\sum_{t=0}^{T-1} \sin^2 w_k t = \sum_{t=0}^{T-1} \cos^2 w_k t = T/2 \quad \text{for } k = 1, 2, \ldots, T/2 - 1.$$

Thus, (27) becomes

$$\hat{\alpha}(w_0) = \frac{\displaystyle\sum_{t=0}^{T-1} y_t}{T}$$

(28) $\qquad \hat{\alpha}(w_{T/2}) = \frac{1}{T} \sum_{t=0}^{T-1} y_t (-1)^t$

$$\hat{\alpha}(w_k) = \frac{2}{T} \sum_{t=0}^{T-1} y_t \cos w_k t \qquad\qquad k = 1, 2, \ldots, T/2 - 1$$

$$\hat{\beta}(w_k) = \frac{2}{T} \sum_{t=0}^{T-1} y_t \sin w_k t \qquad\qquad k = 1, 2, \ldots, T/2 - 1.$$

Since (25) represents a regression of T observations on y against T orthogonal independent variables (which guarantees that the X'X matrix of the linear statistical model is of full rank), we know that the regression fits the data exactly, i.e., it gives a perfect fit. So what we have achieved is a decomposition of y_t {t=0, ..., T-1} into a weighted sum of sine and cosine terms of angular frequencies $w_k = \frac{2\pi k}{T}$, k = 0, ..., T/2. The least squares regression coefficients $\hat{\alpha}(w_k)$ and $\hat{\beta}(w_k)$ give a measure of how important the various frequencies are in composing the series y_t. To make this more precise, notice that from (25), the sample variance of the y's can be written

$$\frac{1}{T} \sum_{t=0}^{T-1} \left(y_t - \frac{\sum_{t=0}^{T-1} y_t}{T}\right)^2 = \frac{1}{T} \sum_{t=0}^{T-1} \left(y_t - \hat{\alpha}(w_0)\right)^2$$

$$= \frac{1}{T} \left\{ \sum_{k=1}^{T/2-1} \hat{\alpha}(w_k)^2 \sum_{t=0}^{T-1} \cos^2(w_k t) + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k)^2 \sum_{t=0}^{T-1} \sin^2(w_k t)\right.$$

$$\left. + \hat{\alpha}(w_{T/2})^2 \sum_{t=0}^{T-1} \cos^2(w_{T/2} t) \right\},$$

which follows by virtue of the orthogonality of sines and cosines of different frequencies. From our earlier calculations of $\sum_{t=0}^{T-1} \cos^2 w_k t$ and $\sum_{t=0}^{T-1} \sin^2 w_k t$, the above equation becomes

$$\frac{1}{T} \sum_{t=0}^{T-1} (y_t - \frac{\sum y_t}{T})^2 = \frac{1}{T} \{ [T/2 \sum_{k=1}^{T/2-1} \hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)]$$

$$+ T \hat{\alpha}^2(w_{T/2})\} = \frac{1}{2} \{ \sum_{t=1}^{T/2-1} [\hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)] + 2 \hat{\alpha}^2(w_{T/2})\}.$$

Thus, the term $1/2 [\hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)]$ measures the contribution of sine and

cosine terms of frequency w_k to the sample variance of y.

To look at the coefficients $\hat{\alpha}(w_k)$ and $\hat{\beta}(w_k)$ from a slightly

different perspective, consider the quantities

$$(29) \qquad A(w_k) = \frac{1}{T} \sum_{t=0}^{T-1} y_t e^{iw_k t} = \frac{1}{T} \sum_{t=0}^{T-1} y_t \cos w_k t + i \frac{1}{T} \sum_{t=0}^{T-1} y_t \sin w_k t$$

$$\equiv a(w_k) + ib(w_k) \qquad w_k = \frac{2\pi k}{T} , \quad k=0, 1, \ldots, T-1$$

where $a(w_k) = \frac{1}{T} \sum_{t=0}^{T-1} y_t \cos w_k t$, $b(w_k) = \frac{1}{T} \sum_{t=0}^{T-1} y_t \sin w_k t$.

The list of $A(w_k)$'s for $k=0, 1, \ldots, T-1$ is known as the Fourier

transform of the series $y_t \{t=0, \ldots, T-1\}$.

Now consider the quantity

$$(30) \qquad \sum_{k=0}^{T-1} A(w_k) \cdot e^{-iw_k t}$$

$$= \sum_{k=0}^{T-1} (a(w_k) + ib(w_k))(\cos w_k t - i \sin w_k t).$$

Notice that $w_{T-k} = \frac{2\pi(T-k)}{T} = 2\pi - \frac{2\pi k}{T} = 2\pi - w_k$.

Since $\sin(x+2\pi) = \sin(x)$ and $\cos(x+2\pi) = \cos x$, it follows that

$A(w_k) = A(w_k+2\pi)$. Furthermore, since $-\sin x = \sin(-x)$ and $\cos(x)$

$= \cos(-x)$, it follows that

$$A(w_{T-k}) = A(2\pi - w_k) = A(-w_k)$$

$$= a(-w_k) + ib(-w_k)$$

$$= a(w_k) - ib(w_k) \equiv \overline{A(w_k)}$$

where $\overline{A(w_k)}$ is the complex conjugate of $A(w_k)$. Consequently, (30) can be written as

$$\sum_{k=0}^{T-1} A(w_k)e^{-iw_k t} = (a(w_{T/2}) + ib(w_{T/2}))(\cos w_{T/2} t - i \sin w_{T/2} t)$$

$$- (a(w_0) + ib(w_0))(\cos w_0 t - i \sin w_0 t)$$

$$+ \sum_{k=0}^{T/2-1} (a(w_k)+ib(w_k))(\cos w_k t - i \sin w_k t)$$

$$+ \sum_{k=0}^{T/2-1} (a(w_k)-ib(w_k))(\cos w_k t + i \sin w_k t)$$

which, since $\sin w_{T/2} = 0$, equals

$$(31) \qquad - a(w_0) + a(w_{T/2}) \cos (w_{T/2} t)$$

$$+ \sum_{k=0}^{T/2-1} 2 a(w_k) \cos w_k t + \sum_{k=0}^{T/2-1} 2 b(w_k) \sin w_k t.$$

Comparing $a(w_k)$ and $b(w_k)$ with our earlier least squares estimates $\hat{\alpha}(w_k)$ and $\hat{\beta}(w_k)$ we notice that

$$a(w_0) = \frac{1}{T} \sum_{t=0}^{T-1} y_t = \hat{\alpha}(w_0)$$

$$a(w_k) = \frac{1}{2} \hat{\alpha}(w_k) \qquad k = 1, 2, \ldots, T/2 - 1$$

$$(32)$$

$$a(w_{T/2}) = \hat{\alpha}(w_{T/2})$$

$$b(w_k) = \frac{1}{2} \hat{\beta}(w_k) \qquad k = 1, 2, \ldots, T/2 - 1.$$

Consequently (31) equals the least squares regression

$$\sum_{k=0}^{T/2} \hat{\alpha}(w_k) \cos w_k t + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k) \sin w_k t$$

which we know equals y_t by virtue of the perfect fit of (25). We have therefore proved that

$$\sum_{k=0}^{T-1} A(w_k) e^{-iw_k t} = y_t,$$

which is a theorem due to Fourier. The real and imaginary parts of

$A(w_k) = a(w_k) + i \, b(w_k)$ are (apart from a scaler for $k = 1, \ldots, T/2-1$)

the regression coefficients in (25), as is summarized in (32).

A "natural" measure of the importance of the cosine and sine

of the frequency w_k in composing y_t is the squared amplitude of $A(w_k)$,

which is

$$A(w_k)\overline{A(w_k)} = \left| A(w_k) \right|^2$$

$$= (a(w_k)+ib(w_k))(a(w_k)-ib(w_k))$$

$$= a^2(w_k)+b^2(w_k).$$

The higher is this quantity, the larger is the weight put on the sine

and cosine of frequency w_k in (25) in making up y_t. As it turns out,

the quantity $\left| A(w_k) \right|^2$ can be used to estimate the value of the spectrum

of y at frequency w_k.

Consider the quantity

$$T \, A(w_k)\overline{A(w_k)} = \frac{1}{T} \sum_{t=0}^{T-1} y_t e^{iw_k t} \sum_{j=0}^{T-1} y_j e^{-iw_j j}$$

$$= \frac{1}{T} \sum_{t=0}^{T} \sum_{j=0}^{T} y_j y_t e^{iw_k(t-j)}$$

$$= \frac{1}{T} \{ \ y_0(y_0 + y_1 e^{-iw_k 1} + y_2 e^{-iw2} + \ldots + y_{T-1} e^{-iw(T-1)})$$

$$+ \ y_1(y_0 e^{iw1} + y_1 + y_2 e^{-iw1} + \ldots + y_{T-1} e^{-iw(T-2)})$$

$$+ \ y_2(y_0 e^{iw2} + y_2 e^{iw1} + y_2 + \ldots + y_{T-1} e^{-iw(T-3)})$$

$$+$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$+ \ y_{T-1}(y_0 e^{iw(T-1)} + y e^{iw(T-2)} + \ldots + y_{T-1})$$

$$= \frac{1}{T} \sum_{t=0}^{T-1} y_t^2 + \frac{1}{T} \sum_{j=1}^{T-1} \sum_{t=j}^{T-1} y_t y_{t-j} (e^{iw_k j} + e^{-iw_k j})$$

$$TA(w_k) \overline{A(w_k)} = \frac{1}{T} \sum_{t=0}^{T-1} y_t^2 + \frac{1}{T} \sum_{j=1}^{T-1} \sum_{t=j}^{T-1} y_t y_{t-j} \ 2 \cos w_k j \ .$$

We are taking the view that y_t { $t = 0, 1, \ldots, T-1$ } is the realization of a random process, so that it is appropriate to inquire about the expected value of $TA(w_k)\overline{A(w_k)}$, which is a random variable itself, being a function of the random y_t's. Taking expected values on both sides of the above equation gives

$$\mathrm{ETA}(w_k)\overline{A(w_k)} = ET(a^2(w_k) + b^2(w_k))$$

$$= c_y(0) + 2 \sum_{j=1}^{T-1} \frac{1}{T} \sum_{t=j}^{T-1} E(y_t y_{t-j}) \cos w_k j$$

$$(33) \qquad \mathrm{ETA}(w_k)\overline{A(w_k)} = c_y(0) + 2 \sum_{j=1}^{T-1} (1-\frac{j}{T}) c_y(j) \cos w_k j.$$

Recall that the spectrum of y is

$$g_y(e^{-iw_k}) = c_y(0) + 2 \sum_{j=1}^{\infty} c_y(j) \cos w_k j.$$

Now as $T \to \infty$, the term $1 - j/T \to 1$ for fixed j. Thus if $c_y(j)$ approaches zero fast enough as $j \to \infty$, we have that

$$\mathrm{ETA}(w_k)\overline{A(w_k)} \to g_y(e^{-iw_k})$$

for <u>all</u> frequencies w_k (not just those for which $w_k = 2\pi k/T$, integer k). (It is possible to show that for any T,

$$\mathrm{ETA}(w_k)\overline{A(w_k)} = g_y(e^{-iw_k})$$

for $w_k = 2\pi k/T$, k an integer. See, e.g., Melvin Hinich, "Introduction to Fourier Analysis of Data," Center for Naval Analyses, 1969, p. 22.)

We have thus showed that the variable $(a^2(w_k) + b^2(w_k))$ bears an intimate relation to the spectrum. The quantities $(a^2(w_k) + b^2(w_k))$ are called periodogram ordinates, and a graph of them for various w_k against w_k is known as the periodogram. (In fact computing the periodogram is where one often begins on the way to estimating a spectrum.) It is the relation of the quantities $a(w_k)$ and $b(w_k)$ to the regression (25) of y_t on sines and cosines, on the one hand, and to the spectrum of y on the other hand, that motivates the interpretation of the spectrum as a decomposition of the variance of y by frequency.

Let us examine the spectra of some simple processes. First consider the white noise process

$$y_t = \varepsilon_t$$

ε_t white so that $c_y(0) = \sigma_\varepsilon^2$, $c_y(h) = 0$ for $h \neq 0$.

For this process the covariance generating function is simply

$$g_y(z) = \sigma_\varepsilon^2,$$

so that the spectrum is

$$g_y(e^{-iw}) = \sigma_\varepsilon^2, \qquad -\pi \leq w \leq \pi$$

so that the spectrum is flat, as depicted in Figure 3, and equals σ_ε^2 at each frequency. Notice that

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) dw = 2\pi \sigma_\varepsilon^2,$$

as expected. So a white noise has a flat spectrum, indicating that all frequencies between $-\pi$ and π are equally important in accounting for its variance.

Next consider the first order process

$$y_t = B(L) \varepsilon_t = \frac{1}{1-\lambda L} \varepsilon_t. \qquad -1 < \lambda < 1.$$

For this process the covariance generating function is

$$g_y(z) = \left(\frac{1}{1-\lambda z} \right) \left(\frac{1}{1-\lambda z^{-1}} \right) \sigma_\varepsilon^2.$$

Therefore the spectrum is

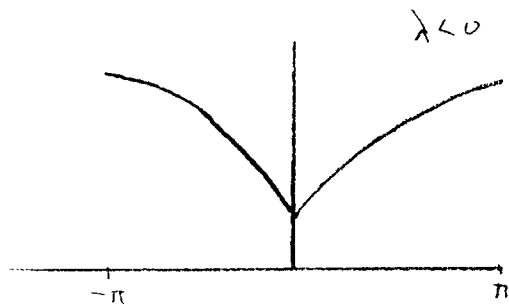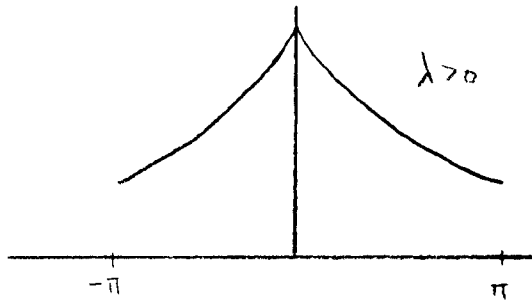$$g_y(e^{-iw}) = (\frac{1}{1-\lambda e^{-iw}})(\frac{1}{1-\lambda e^{iw}}) \; \sigma_\epsilon^2$$

$$= \frac{1}{(1-2\lambda(e^{iw}+e^{-iw}) + \lambda^2)}$$

$$g_y(w) = \frac{1}{1-2\lambda \cos w + \lambda^2} \quad .$$

Notice that

$$\frac{dg_y(w)}{dw} = - (1-2\lambda\cos w + \lambda^2)^{-2}(2\lambda \sin w) \; .$$

The first term in parenthesis is positive. Since $\sin w > 0$ for

$0 < w < \pi$, the second term is negative on $(0,\pi]$ if $\lambda < 0$ and positive on

$(0,\pi]$ if $\lambda > 0$. Therefore if $\lambda > 0$, the spectrum decreases on $(0,\pi]$ as w

increases; if $\lambda < 0$, the spectrum increases on $(0,\pi]$ as w increases.

Thus if $\lambda > 0$, low frequencies (i.e., low values of w) are relatively important in composing the variance of w, while if $\lambda < 0$, high frequencies are the more important. It is easy to verify that the higher in absolute value is λ, the steeper is the spectrum. Notice that the first order process can have a peak in its spectrum only at w=0 or w=$\pm\pi$. A peak at w=π corresponds to a periodicity of

$2\pi/w = 2\pi/\pi = 2$ periods. A peak at w=0, corresponds to a cycle with

"infinite" periodicity, which is unobservable and hence not a cycle at all.

With quarterly data, a business cycle corresponds to a peak in the spectrum at a periodicity of about 12 quarters. A first order process is capable of having a peak only at two quarters or at "infinite" quarters, and so is not capable of rationalizing a business cycle in the sense of a peak in the spectrum at about 12 quarters. As we saw above, a first order process cannot possess a covariogram with a periodicity other than two periods, and so with quarterly data cannot rationalize a business cycle in the sense of an oscillatory covariogram.

Next consider the second order process

$$y_r = \frac{1}{1-t_1 L-t_2 L^2} \varepsilon_t,$$

ε_t white noise. For this process the covariance generating function is

$$g_y(z) = \frac{1}{1-t_1 z-t_2 z^2} \cdot \frac{1}{1-t_1 z^{-1}-t_2 z^{-2}} \sigma_\varepsilon^2.$$

Therefore the spectrum of the process is

$$g_y(e^{-iw}) = \frac{1}{1-t_1 e^{-iw}-t_2 e^{-2iw}} \frac{1}{1-t_1 e^{iw}-t_2 e^{2iw}} \sigma_\varepsilon^2$$

$$= \frac{\sigma_\varepsilon^2}{1+t_1^2+t_2^2+(t_2 t_1-t_1)(e^{iw}+e^{-iw})-t_2(e^{-2iw}+e^{2iw})}$$

$$= \frac{\sigma_\varepsilon^2}{1+t_1^2+t_2^2-2t_1(1-t_2)\cos w-2t_2 \cos 2w} = \frac{\sigma_\varepsilon^2}{h(w)}.$$

Differentiating with respect to w, we have

$$\frac{dg_y(e^{-iw})}{dw} = -\sigma_\varepsilon^2 h(w)^{-2}(2t_1(1-t_2)\sin w + 4t_2 \sin 2w)$$

$$+ -\sigma_\varepsilon^2 h(w)^{-2}(2 \sin w \cdot [t_1(1-t_2)+4t_2 \cos w]).$$

We know that $h(w) > 0$. For the above derivative to be zero at a w belonging to $(0, \pi)$, we must have the term in brackets equal to zero:

$$t_1(1-t_2) + 4t_2 \cos w = 0$$

or

$$(35) \qquad \cos w = \frac{-t_1(1-t_2)}{4t_2}$$

so that

$$(35') \qquad w = \cos^{-1} \left(\frac{-t_1(1-t_2)}{4t_2} \right) .$$

Equation (35) can be satisfied only if

$$(36) \qquad \left| \frac{-t_1(1-t_2)}{4t_2} \right| < 1,$$

since $|\cos x| \leq 1$ for all x. If (36) is met, the spectrum of y does achieve a maximum on $(0,\pi)$. Condition (36) is slightly more restrictive than the condition that the roots of the deterministic difference equation be complex so that the covariogram display oscillations. Let us write (36) as

$$(37) \qquad -1 < \frac{-t_1(1-t_2)}{4t_2} < 1.$$

The boundaries of the region (37) are

$$(38) \qquad -t_1(1-t_2) = 4t_2$$

and

$$(39) \qquad -t_1(1-t_2) = -4t_2 .$$

The points $(t_1, t_2) = (0,0)$ appear on both boundaries, while the point $(t_1, t_2) = (2, -1)$ appears on (38) and $(t_1, t_2) = (-2, -1)$ appears on (39). Differentiating (38) implicitly with respect to t_1 gives

$$\frac{dt_2}{dt_1} = \frac{t_2-1}{4-t_2} \, ,$$

so that along (38)

$$\left.\frac{dt_2}{dt_1}\right)_{t_1=t_2=0} = -\frac{1}{4}$$

and

$$\left.\frac{dt_2}{dt_1}\right)_{\substack{t_1=2 \\ t_2=-1}} = -1.$$
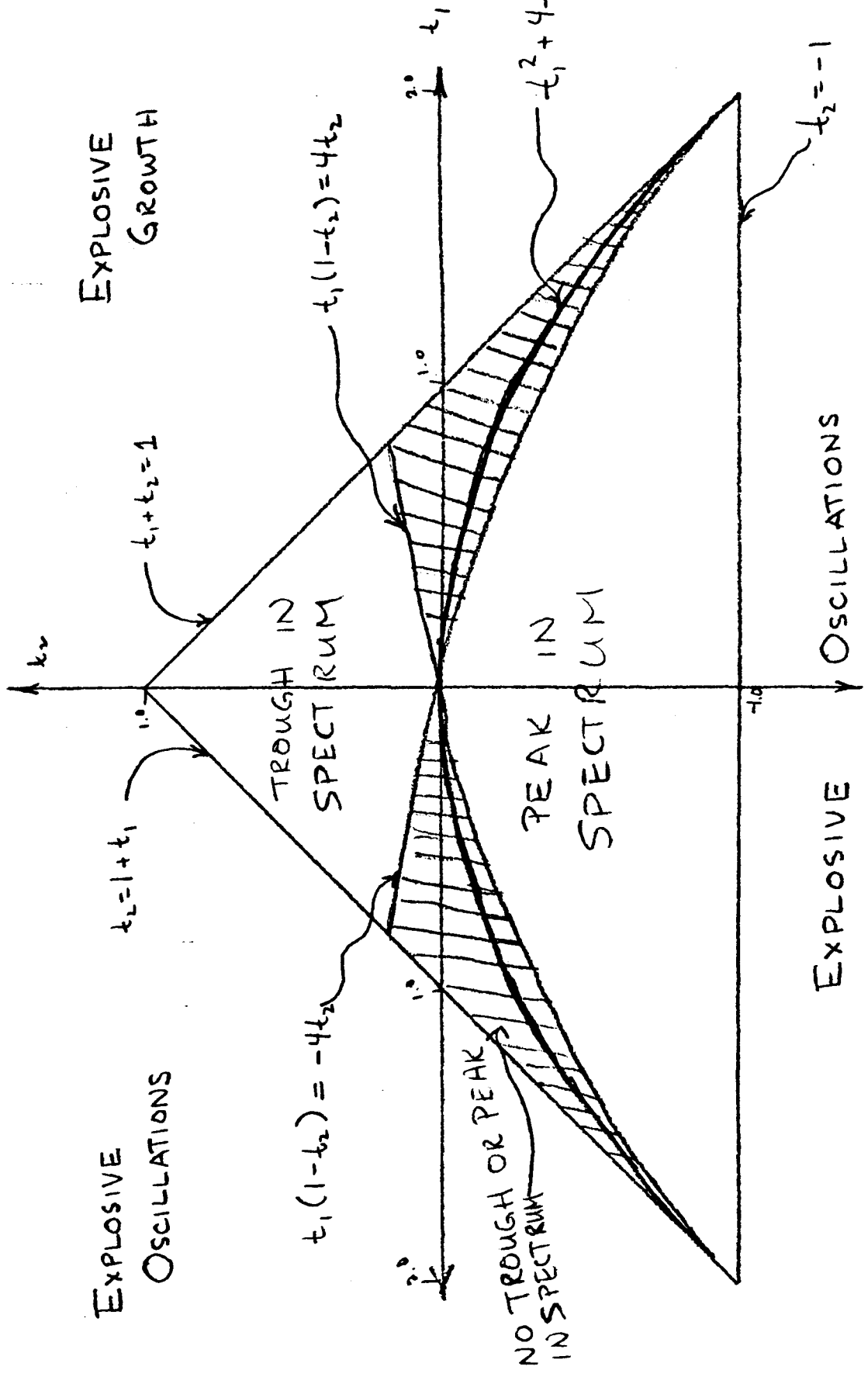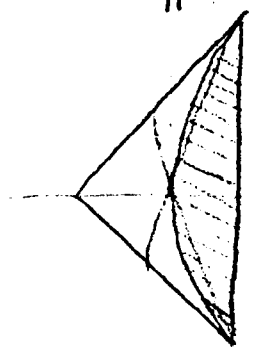
Differentiating (39) with respect to t_1 gives

$$\frac{dt_2}{dt_1} = \frac{1-t_1}{4+t_1}$$

so that along (39)

$$\left.\frac{dt_2}{dt_1}\right)_{t_1=t_2=0} = \frac{1}{4}$$

$$\left.\frac{dt_2}{dt_1}\right)_{\substack{t_1=-2 \\ t_2=-1}} = 1.$$

- 329 -

= OSCILLATORY COVARIOGRAM

EXPLOSIVE GROWTH

$t_1^2 + 4t_2^2 = 0$

$t_2 = -1$

$t_1(1-t_2) = 4t_2$

$t_1 + t_2 = 1$

TROUGH IN SPECTRUM

PEAK IN SPECTRUM

OSCILLATIONS

EXPLOSIVE OSCILLATIONS

$t_2 = 1 + t_1$

$t_1(1-t_2) = -4t_2$

NO TROUGH OR PEAK IN SPECTRUM

EXPLOSIVE

Such calculations show that the boundaries of region (37) are as depicted in Figure 4. To be in region (37) with $t_2 < 1$ (a requirement of covariance stationarity) implies that the roots of the difference equation are complex. However, complex roots don't imply that (37) is satisfied. Consequently, our two definitions of the business cycle aren't quite equivalent.

The "Slutsky Effect" and Kuznets' Transformation

In the above examples, we have seen that if

$$(40) \qquad y_t = B(L)\varepsilon_t,$$

where ε_t is white noise, then the spectrum of y is related to the spectrum of ε_t by

$$g_y(e^{-iw}) = B(e^{-iw})B(e^{iw}) \, \sigma_\varepsilon^2$$

or

$$(41) \qquad g_y(e^{-iw}) = B(e^{-iw})B(e^{iw})g_\varepsilon(e^{-iw})$$

since for the white noise ε, $g_y(e^{-iw}) = \sigma_\varepsilon^2$. It is straightforward to show that for any ε_t, not necessarily a white one, affecting y via (40), the spectrum of y is related to the spectrum of ε by (41). Thus assume that y is related to X by

$$(42) \qquad y_t = \sum_{s=-p}^{q} b_s X_{t-s} \equiv B(L)X_t \qquad\qquad p \geq 0, \ q \geq 0$$

and that the spectrum of X is defined. From (42) we know that

$$y_t y_{t-j} = \sum_{s=-p}^{q} b_s X_{t-s} \sum_{r=-p}^{q} b_r X_{t-j-r}$$

$$= \sum_{s=-p}^{q} \sum_{r=-p}^{q} b_s b_r X_{t-s} X_{t-j-r}.$$

Taking expected values on both sides gives

$$c_y(j) = E(y_t y_{t-j}) = \sum_{s=-p}^{q} \sum_{r=-p}^{q} b_s b_r c_x(j+r-s).$$

The spectrum of y is defined as

$$g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) e^{-iwk}$$

(43)
$$= \sum_{k=-\infty}^{\infty} \sum_{s=-p}^{\infty} \sum_{r=-p}^{\infty} b_r b_s c_x(k+r-s) e^{-iwk}.$$

Define the index $h = k + r - s$, so that $k = h - r + s$. Notice that

(44) $$e^{-iwk} = e^{-iw(h-r+s)} = e^{-iwh} e^{-iws} e^{iwr} .$$

Substituting (44) into (43) gives

$$g_y(e^{-iw}) = \sum_{r=-p}^{q} b_r e^{iwr} \sum_{s=-p}^{q} b_s e^{-iws} \sum_{h=-\infty}^{\infty} c_x(h) e^{-iwh}$$

(45) $$g_y(e^{-iw}) = B(e^{iw}) B(e^{-iw}) g_x(e^{-iw})$$

or

$$g_y(e^{-iw}) = \left| B(e^{-iw}) \right|^2 g_x(e^{-iw}) ,$$

which shows that the spectrum of the "output" y equals the spectrum of the "input" x multiplied by the real number $B(e^{iw}) B(e^{-iw})$.

Relation (45) can be used to show the famous "Slutsky" effect. Slutsky considered the effects of starting with a white noise ε_t,

taking a 2 period moving sum n times, and then taking first differences
m times. That is, Slutsky considered forming the series

$$Z_t = (1+L)(1+L)\ldots(1+L)\varepsilon_t = (1+L)^n\varepsilon_t$$

and

$$y_t = (1-L)(1-L)\ldots(1-L)Z_t = (1-L)^m Z_t$$

(46) $$y_t = (1+L)^n(1-L)^m\varepsilon_t.$$

Applying (45) to (46) we have

$$g_y(e^{-iw}) = (1+e^{iw})^n(1+e^{-iw})^n(1-e^{iw})^m(1-e^{-iw})^m\sigma_\varepsilon^2$$

$$= [(1+e^{iw})(1+e^{-iw})]^n[(1-e^{-iw})(1-e^{iw})]^m\sigma_\varepsilon^2$$

$$= [(2+(e^{iw}+e^{-iw})]^n[(2-(e^{iw}+e^{-iw})]^m\sigma_\varepsilon^2$$

(47) $$g_y(e^{-iw}) = \sigma_\varepsilon^2 2^n[1+\cos w]^n\, 2^m[1-\cos w]^m .$$

Consider first the special case where m = n. Then (47) becomes

$$g_y(e^{-iw}) = \sigma_\varepsilon^2 4^n[1- \cos^2 w]^n$$

(48) $$= \sigma_\varepsilon^2 4^n[\sin^2 w]^n .$$

On $[0,\pi]$, the spectrum of y has a peak, at $w = \pi/2$, since there
$\sin w = 1$. Notice that since $\sin w \leq 1$, (48) implies that as n becomes
large, the peak in the spectrum of y at $\pi/2$ becomes sharp. In the
limit, as $n \to \infty$, the spectrum of y becomes a "spike" at $\pi/2$, which means
that y behaves like a cosine of angular frequency $\pi/2$.

Similar behavior results for fixed m/n as n becomes large where $m \neq n$. Consider (47) and set $dg_y(e^{-iw})/dw$ equal to zero in order to locate the peak in the spectrum:

$$\frac{dg_y}{dw} = \sigma_\epsilon^2 2^{m+n} \{ n[1-\cos w]^m [1+\cos w]^{n-1} (-\sin w)$$

$$+ m(1-\cos w)^{m-1} (\sin w)[1+\cos w]^n \}$$

$$= \sigma_\epsilon^2 2^{m+n} \sin w \{ (1-\cos w)^{m-1} (1+\cos w)^{n-1})$$

$$\cdot [m(1+\cos w)-n(1-\cos w)] \}.$$

This expression can equal zero on $(0,\pi)$ only if the expression in brackets equals zero:

$$m(1+\cos w)-n(1-\cos w) = 0$$

which implies

$$\cos w = \frac{1 - \frac{m}{n}}{1 + \frac{m}{n}} \quad ,$$

or

$$w = \cos^{-1}(\frac{1-m/n}{1+m/n})$$

which tells us the frequency at which the spectrum of y attains a peak. For fixed m/n, the spectrum of y approaches a spike as $n \to \infty$. This means that as $n \to \infty$, y tends to behave more and more like a cosine of angular frequency $\cos^{-1}((1-m/n)/(1+m/n))$.

What Slutsky showed, then, is that by successively summing and then successively differencing a serially uncorrelated or "white noise" process ϵ_t, a series with "cycles" is obtained.

Another use of (45) is in the analysis of transformations that have been applied to data. An example is Howrey's analysis of the transformations used by Kuznets. Data constructed by Kuznets have been inspected to verify the existence of "long swings," long cycles in economic activity of around twenty years. Before analysis, however, Kuznets subjected the data to two transformations. First, he took a five year moving average:

$$Z_t = \frac{1}{5} [L^{-2}+L^{-1}+1+L+L^2]X_t \equiv A(L)X_t .$$

Then he took the centered first difference of the (nonoverlapping) five year moving average:

$$y_t = Z_{t+5} - Z_{t-5} = [L^{-5}-L^5]Z_t = B(L)Z_t .$$

So we have that the y's are related to the X's by

$$y_t = \frac{1}{5} [L^{-5}-L^5][L^{-2}+L^{-1}+1+L+L^2]X_t$$

$$= A(L)B(L)X_t .$$

The spectrum of y is related to the spectrum of X by

$$(49) \qquad g_y(e^{-iw}) = A(e^{-iw})A(e^{iw})B(e^{-iw})B(e^{iw})g_x(e^{-iw}) .$$

We have

$$A(e^{-iw}) = \frac{1}{5} \sum_{j=-2}^{2} e^{-iwj} = \frac{1}{5} \frac{(e^{iw2}-e^{-iw3})}{(1-e^{-iw})} .$$

Thus,

$$A(e^{-iw})A(e^{iw}) = \frac{(\tfrac{1}{5})^2(e^{iw2}-e^{-iw3})(e^{-iw2}-e^{iw3})}{(1-e^{-iw})(i-e^{iw})}$$

$$= \frac{(\tfrac{1}{5})^2(2-(e^{iw5}+e^{-iw5}))}{(2-(e^{iw}+e^{-iw}))}$$

$$= \frac{(\tfrac{1}{5})^2 2(1-\cos 5w)}{2(1-\cos w)} = \frac{(\tfrac{1}{5})^2(1-\cos 5w)}{(1-\cos w)} \; .$$

Next, we have

$$B(e^{-iw}) = (e^{+iw5}-e^{-iw5})$$

so that

$$B(e^{-iw})B(e^{iw}) = (e^{iw5}-e^{-iw5})(e^{-iw5}-e^{iw5})$$

$$= (2-(e^{iw10}+e^{-iw10})) = 2(1-\cos 10w).$$

So it follows from (49) that

$$g_y(e^{-iw}) = [\frac{(\tfrac{1}{5})^2(1-\cos 5w)2}{(1-\cos w)} (1-\cos 10\ w)] \, g_x(e^{-iw})$$

$$= G(w)g_x(e^{-iw}).$$

where $G(w) = 2[(\tfrac{1}{5})^2(1-\cos 5w)(1-\cos 10w)/(1-\cos w)]$.

The term $G(w)$ is graphed in Figure __. It has zeroes at values where $\cos 5w = 1$ and where $\cos 10w = 1$. The first condition occurs on $[0,\pi]$ where

$$5w = 0, \; 2\pi, \; 4\pi,$$

or

$$w = 0, \; \tfrac{2}{5}\pi, \; \tfrac{4}{5}\pi.$$

The condition cos 10w = 1 on $[0, \pi]$ where

$$10w = 0, \quad 2\pi, \quad 4\pi, \quad 6\pi, \quad 8\pi, \quad 10\pi$$

or

$$w = 0, \quad \frac{1}{5}\pi, \quad \frac{2}{5}\pi, \quad \frac{4}{5}\pi, \quad \text{and} \quad \pi.$$

So $G(w)$ has zeroes at $w = 0, \pi/5, 2/5\pi, 3\pi/5, 4\pi/5$, and π.

From the graph of $G(w)$, it follows that even if X_t is a white noise, a y series generated by applying Kuznets' transformations will have a large peak at a low frequency, and hence will seem to be characterized by "long swings." These long swings are clearly a statistical artifact; that is, they are something induced in the data by the transformation applied and not really a characteristic of the economic system. With annual data, the biggest peak in Figure __ corresponds to a cycle of about 20 1/4 years which is close to the 20 year cycle found by Kuznets. Howrey's observations naturally raise questions about the authenticity of the long swings identified by studying the data used by Kuznets.

$$G(\omega) = \left(\frac{1}{5}\right)^2 \frac{(1-\cos 5\omega) \cdot 2(1-\cos 10\omega)}{1-\cos \omega}$$
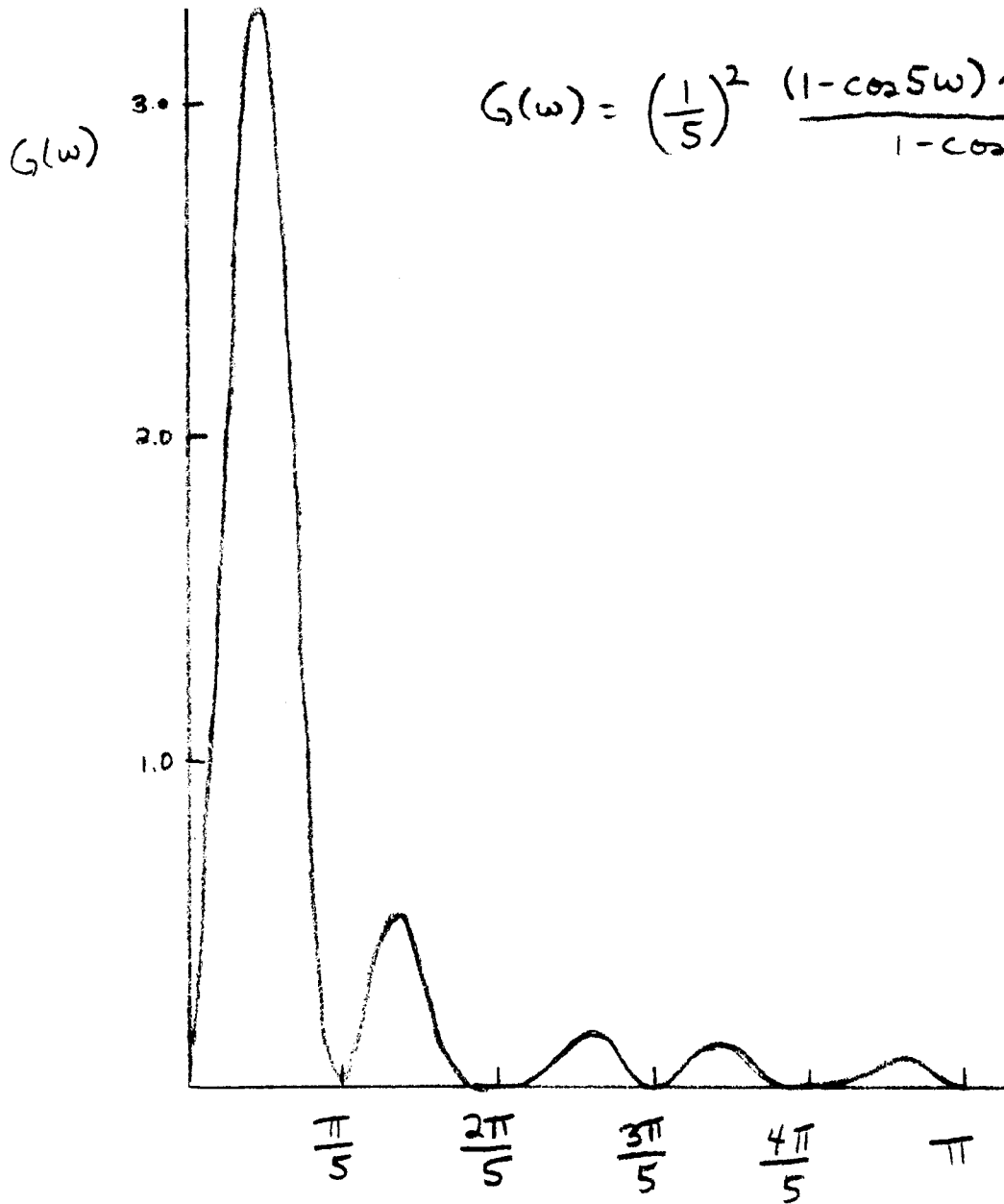
Exercises

1. Take a simple version of the Keynesian model described above, one
    with $\pi(M+B)/P$ excluded from the consumption function.

    A. Describe the behavior of the model (i.e., the response of all
        endogenous variables to jumps in each of the exogenous variables)
        where the endogenous variables are taken to be Y, N, C, I, r,
        and M while the exogenous variables are G, T, π, w, and P.

    B. Describe the behavior of the model where the endogenous variables
        are Y, N, C, I, P, and M and the exogenous variables are G, T,
        π, w, and r.

2. Take a simple version of the Classical Model. Describe the behavior
    of the model where the endogenous variables are Y, N, C, I, w, r,
    and M while the exogenous variables are G, T, π, and P.

3. Consider an economy described by the following equations.

$$(1) \quad \frac{Y}{k} = A\left(\frac{N}{K}\right)^{1.10} \qquad \text{(production function, Bodkin and Klein, Restat, 1967)}$$

$$(2) \quad \frac{N}{k} = \beta_o\left(\frac{w}{P}\right)^{\beta_1} \qquad \beta_1 < 0 \text{ (demand function for labor)}$$

$$(3) \quad I = I(r) \qquad I' < 0$$

$$(4) \quad C = C(Y-T) \qquad 0 < C' < 1$$

$$(5) \quad C + I + G = Y$$

$$(6) \quad \frac{M}{P} = m(r, Y, W), \qquad m_r < 0, \ m_Y > 0, \ m_W = 1$$

$$(7) \quad W = \frac{M+B}{P} + K$$

The endogenous variables are:

Y = GNP   C = consumption

N = employment   r = interest rate

P = the price level   W = real wealth held by the
                               public

I = investment

The exogenous variables are:

K = the capital stock   M = the money stock, a lia-
                               bility of the government

w = the money wage
                           B = stock of government interest
T = tax collections   bearing bonds

G = government expenditures,
    a flow

Government bonds are like savings deposits, having a variable

interest rate and fixed nominal value. Describe the effects on Y,

N, P, and r of:

A.   An increase in the money supply brought about by an open-

     market operation, i.e., $dM = -dB$.

B.   A once-and-for-all increase in the money supply caused by a

     one-time government money bonus to veterans, i.e., $dM > 0$,

     $dB = 0$.

C.   A once-and-for-all increase in the stock of outstanding

     government bonds, caused by a one-time gift of some new govern-

     ment bonds to veterans, i.e., $dB > 0$, $dM = 0$.

4.   Consider the following macroeconomic model:

$$Y = Y(K, N) \qquad Y_K, Y_N > 0; Y_{KK}, Y_{NN} < 0, Y_{KN} > 0$$

$$\frac{W}{P} = Y_N$$

$$N^S = N^S(\frac{W}{P}, r-\pi) \qquad N^S_{\frac{W}{P}} > 0, N^S_{r-\pi} > 0$$

$$N = N^S$$

$$I = I(r-\pi) \qquad\qquad I' < 0$$

$$C = C(Y-T) \qquad\qquad I > C' > 0$$

$$C + I + G = Y$$

$$\frac{M}{P} = m(r, Y) \qquad\qquad m_r < 0, \; m_Y > 0$$

where Y is GNP, K is the capital stock, N is employment, W is the money wage, N^S is the supply of labor, P is the price level, r is the interest rate, π is the anticipated inflation rate, I is investment, C is consumption, G is government expenditures, T is taxes net of transfers. The dependent variables are Y, N, N^S, W, P, r, C, and I. The exogenous variables are M, K, G, T, and π.

A.   Describe the effects on Y, P, r and N of: (1) an increase in M, (2) an increase in G with T constant, and (3) an increase in π. Does an increase in π leave the real rate of interest unchanged?

B.   Suppose that the government finances its deficit by printing bonds and/or money. At the equilibrium values of the variables that satisfy the above equations, are we assured that the public will be willing to accumulate new "bonds and/or money" at exactly the rate at which the government is creating it? Explain.

5.   The following passage is from Leijonhufvud's book <u>Keynesian Economics and the Economics of Keynes</u>.

Hansen's demonstration of the role of saving
and investment in the determination of the interest
rate involved the description of certain processes,
of which the following is an illustration: (1) an
increase in "Thrift" is interpreted as an upward
shift of the saving schedule; (2) with investment
"autonomous," this leads to a reduction in income;
(3) the money stock being given the reduction in
income is seen to create an excess supply of money,
which (4) spills over into demand for bonds and
drives down the interest rate.

The reasoning is false. A proper analysis would recognize that
while income is declining, there is an excess demand for money,
corresponding to the excess supply of commodities on which the
description explicitly focuses. The decline in income will be
halted when the excess demand for money and the excess supply of
commodities simultaneously reach zero. Keynes' obscure discussion
is to blame for the spread of the notion, implicitly accepted in
analyses of the type just paraphrased, that the rate of interest
will decline if and only if there has emerged an excess supply of
money. In fact, shifts in saving and investment will either have a
direct impact on securities markets and the rate of interest, or
the rate of interest will not be affected at all in the case con-
sidered, since the process does not generate an excess supply of
money. Instances of this type of analysis, most often characterized
by a mechanical manipulation of the IS-LM diagram or the correspond-
ing simultaneous equation system, are extremely common in the
income-expenditure literature.

With whose reasoning do you agree, Hansen's or Leijonhufvud's? In

answering, be sure to explain how one of them is incorrectly applying

Walras's law.

6. Consider an economy described by the following IS-LM diagram.
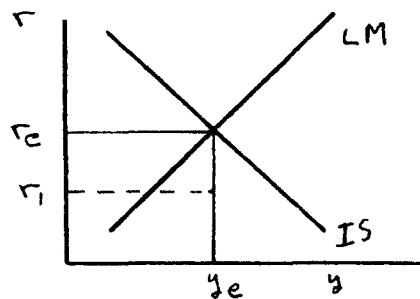
Here (Y_e, r_e) is an equilibrium output-interest rate pair. Consider

now the situation at (Y_e, r_1).

A.   Is there an excess supply or an excess demand for money at

(Y_e, r_1)?

B.   Is there an excess supply or an excess demand for output at

(Y_e, r_1)?

C.   Are your answers to (A) and (B) consistent with Walras's law?

7.   Assume an economy in which money matters; the monetary authority

can influence the interest rate, real output and employment and the

price level at any moment in time. Also, the firms of this economy

are price takers, at least in the labor market, so the marginal

product of labor and the real wage are equal at every moment. Now,

suppose that a constant purchasing power law is passed, a law which

requires all firms to compensate employees for any change in the

commodity price level by increasing money wages proportionately.

What is the significance of this new law for the monetary authority?

Is it still able to influence real output and employment, the

interest rate and the price level?

8.   There is a country whose government is financing all of its expenditures

by means of money creation. All money is the liability of the

government, there being no commercial banks. The government acts

to keep its rate of real expenditures, G, constant over time at the

real rate $G = \overline{G}$. Accordingly, money creation is governed by:

$$\frac{dM}{dt} \Big/ P(t) = \overline{G}$$

where M is the money supply and P the price index. Suppose that

real output is constant over time and that actual and expected

inflation at each moment both equal $\frac{dM}{dt} \Big/ M(t)$. Since $\overline{G}$ is a nonneg-

ligible proportion of the country's GNP, this country has been

experiencing a high rate of inflation. It also happens that the
government has prohibited residents of this country from holding
money and other assets of any foreign countries.

A. How would you find the maximum rate of real expenditures that
the government can finance by money creation?

B. Assume that $\overline{G}$ is initially lower than the maximum G described
in (A), and that the rate of money creation is correspondingly
lower than the rate needed to maximize G. Explain whether
abandoning the restriction on holding of foreign assets would
make it easier or harder for the government to finance its
expenditures at the same real rate $\overline{G}$ by money creation.

C. Assuming that it would still be possible to finance $\overline{G}$ solely
by money creation if the prohibition on holding foreign assets
were eliminated, would the equilibrium rate of inflation be
higher or lower than initially?

9. Consider the following model of a small country.

$$Y = F(N, K) \qquad F_N, \ F_K, \ F_{NK} > 0; \ F_{NN}, \ F_{KK} < 0$$

production function

$$\frac{w}{p} = F_N$$

marginal equality for
employment

$$I = I(r), \ I' < 0$$

investment schedule

$$X = X(t) \qquad X_t < 0$$

export schedule

$$t = \frac{e \cdot p}{p^*}$$

definition of terms
of trade

$$C = c(Y-T), \ 1 > c' > 0$$

consumption function

$$C + I + G + X = Y \qquad \text{Equilibrium condition in market for domes good}$$

$$M/P = m(r, Y) \qquad m_r < 0, \ m_Y > 0; \qquad \text{portfolio equilibrium condition}$$

Here Y is GNP, N is employment, K is capital, I investment, C consumption, X exports, M the money supply, r the interest rate, p the price level in this country (measured in dollars per unit domestic good), p* the price level in the rest-of-the-world (measured in pounds per unit rest-of-the-world), t the terms of trade (measured in rest-of-the-world good per unit domestic good), and e is the exchange rate measured in pounds per dollar (it is the price of dollars measured in foreign currency). Notice that exports decline when t rises. The domestic interest rate r must equal the world interest rate and hence is <u>exogenous</u>. The other exogenous variables are K, w, p*, G, T, and <u>either</u> M <u>or</u> e. The endogenous variables are Y, N, p, I, C, X, t, and <u>either</u> e <u>or</u> M.

A. Consider a regime in which the government pegs the exchange rate e at some arbitrary level, so that e is exogenous and M is endogenous. In this fixed exchange-rate regime, analyze the effects on the endogenous variables of:

   i) An increase in G.

   ii) An increase in the exchange rate e.

   iii) An increase in w.

B. Consider a flexible exchange rate regime in which the government sets M exogenously and permits e to be endogenous. Analyze the effects on the endogenous variables of:

i) An increase in G.

ii) An increase in M.

iii) An increase in w.

10. Consider an economy described by the following equations.

    (1) $Y = F(K, N)$ with F_K, F_N, $F_{KN} > 0$; and F_{NN}, $F_{KK} < 0$

                                   (production function)

    (2) $\dfrac{w}{p} = F_N$ (marginal equality for employment)

    (3) $r = F_K$ (marginal equality for capital)

    (4) $C = C(Y_D)$ with $0 < C' < 1$ (consumption function)

    (5) $Y_D = Y - T$

    (6) $C + I + G = Y$

    (7) $\dfrac{M}{p} = m(r, Y_D)$ with $m_r < 0$, $m_{Y_D} > 0$ (portfolio equilibrium condition)

The endogenous variables are:

Y = GNP   C = consumption

N = employment   r = interest rate

p = the price level Y_D = disposable income

I = investment

The exogenous variables are:

K = the capital stock   M = the money stock, a liability of the government

w = the money wage

                                   G = government expenditures

T = tax collections

Describe the effects on the endogenous variables of:

(A) An increase in the money wage.

(B) An increase in government expenditure.

(C) An increase in tax collections.

11. An economy is described by the following equations:

(A) $Y/K = f(N/K)$ with $f' > 0$, $f'' < 0$

(B) $w/p = f'(N/K)$

(C) $DK/K = .0004/(r-\pi)$

(D) $C = z(Y-T); \quad 0 < z < 1$

(E) $Y = C + DK + G$

(F) $M/(pK) = g(r, Y/K)$

(G) $N \equiv N^S$

(H) $DN^S/N^S \equiv n$

(I) $\pi \equiv Dp/p$

Here Y is GNP, K capital, N employment, N^S labor supply, C consumption, r the interest rate, π the expected rate of inflation, M the money supply, G government purchases, T taxes net of transfers, p the price level, and w the money wage; D is the (right-hand) time derivative operator; i.e., $Dx(t) \equiv dx(t)/dt$. The "givens" of the model are initial conditions for K and N^S, and time paths of the exogenous government variables M, G, and T. For paths of DM/M, G/K and T/K that are constant through time, the model possesses a steady-state equilibrium.

A. Derive a formula for the steady-state value of Y/K.

B. Suppose that $n = 0.02$ and that $DM/M = 0.05$. Compute the steady-state value of the nominal interest rate r.

C. Determine the effect of a once-and-for-all increase in G/K on the steady-state real rate of interest.

12. Consider an economy described by the following equations:

(A) $Y = F(K, N)$ with F_K, F_N, $F_{KN} > 0$; F_{NN}, $F_{KK} < 0$

(B) $w/p = F_N$

(C) $r = F_k$

(D) $C = C(Y-T)$ with $0 < C' < 1$

(E) $Y = C + I + G$

(F) $M/p = H(r, C+I)$ with $H_r < 0$, $H_{C+I} > 0$

Here Y is GNP, N employment, K capital, w the money wage, p the price level, C consumption, I investment, T taxes net of transfers, G government purchases, and M the money supply. The exogenous variables are K, M, G, T, and w. The endogenous variables are Y, N, C, I, r, and p.

Describe the effects on all six endogenous variables of:

A. An increase in M achieved via an open-market operation.

B. An increase in T.

C. An increase in G.

13. Consider an economy described by the following equations:

$$C_t = (\frac{1.80}{\sqrt{2}} - .81)Y_{t-1} + 100.$$

$$I_t = .81(Y_{t-1} - Y_{t-2}) + \varepsilon_t + 10.$$

$$Y_t = C_t + I_t + G_t.$$

where C, I, Y, and G are consumption, investment, GNP, and government purchases, respectively; and $\{\varepsilon_t\}$ is a "white noise" process with mean $E[\varepsilon_t] = 0$ and variance $\sigma^2 = E[\varepsilon_t^2]$. The units of time are quarters. Suppose that the government sets G_t at the constant level $\overline{G}$ every period.

A. Is the economy one with a business cycle? If so, what is the average period of the cycle as measured by the period of the cycles in the covariogram of GNP?

B. Suppose now that G_t is set so that fiscal policy leans against the wind. In particular, the government employs the "feedback rule:"

$$G_t = 10 - \lambda(Y_{t-1} - Y_{t-2}), \quad \lambda > 0.$$

Show how variations in λ will affect the behavior of GNP. Can λ be set so that there are no "business cycles," defined as cycles in the covariogram of GNP?

C. Suppose that the government uses the feedback rule:

$$G_t = 10 + \lambda_1 Y_{t-1} + \lambda_2 Y_{t-2}.$$

What values of λ_1 and λ_2 minimize the variance of GNP? (Hint: write Y_t in the form:

$$Y_t = \text{constant} + \sum_{i=0}^{\infty} w_i \varepsilon_{t-1}$$

and see how the w_i's depend on λ_1 and λ_2.)

14. Suppose that portfolio equilibrium is described by Cagan's equation:

$$m_t - p_t = -1(p_{t+1}^e - p_t) \qquad t=0, 1, 2, \ldots$$

Where m_t is the log of the money supply, p_t the log of the price level at t, and p_{t+1}^e the public's expectation of p_{t+1}, formed at time t. Suppose that expectations are "rational," so that

$$p_{t+1}^e = p_{t+1}.$$

A.   Suppose that $\{m_t\}$, t=0, 1, ... is given by

$$m_t = 10 \cdot \lambda^t \qquad t=0, 1, 2, \ldots$$

Compute the equilibrium value of p_t for t=0, 1, 2, and t=5, 6 for the following values of λ:

i) $\lambda = 1$.

ii) $\lambda = 1.5$

iii) $\lambda = 2.0$

B.   Suppose that m_t follows the path

$$m_t = 10\lambda^t \qquad t=0, 1, 2, 3, 4$$

$$m_t = 11\lambda^t \qquad t=5, 6, 7, 8, \ldots$$

Compute the equilibrium values of p_t for t=0, 1, 2, 3, 4, 5, 6 assuming that $\lambda = 1.5$. How does this time path for $\{p_t\}$ compare with that computed in A.ii? Graph the two paths.

15. An economy is described by the following equations:

$$C_t = c_1 Y_{t-1} + u_{1t} \qquad\qquad 0 < c_1 < 1$$

$$I_t = a_o + a_1(Y_{t-1} - Y_{t-2}) + a_2 r_t + u_{2t} \qquad a_1 > 0 > a_2$$

$$m_t = b_1 r_t + b_2 Y_{t-1} + v_t \qquad\qquad b_1 < 0, b_2 > 0$$

$$C_t + I_t = Y_t$$

where C_t is consumption at t, I_t is investment, Y_t is GNP, r_t is the interest rate, and m_t is the money supply; u_{1t}, u_{2t}, and v_t are serially uncorrelated random variables (white noises) that are

mutually independent (i.e., u_{1t}, u_{2t}, and v_t are pairwise uncorrelated), with variances σ_{u1}^2 σ_{u2}^2 and σ_v^2.

The monetary authority desires to minimize the mean-squared error:

$$E(Y_t - Y*)^2, \text{ where } Y* \text{ is the targent level of GNP.}$$

To achieve this end, the monetary authority considers two alternative strategies. The first is to peg the money supply via the feedback rule.

$$(1) \quad m_t = \Lambda_0 + \lambda_1 Y_{t-1} + \lambda_2 Y_{t-2} + \lambda_3 Y_{t-3}$$

where Λ_0, λ_1, λ_2, and λ_3 are parameters to be chosen. The second is to peg the interest rate via the feedback rule:

$$(2) \quad r_t = \Delta_0 + \delta_1 Y_{t-1} + \delta_2 Y_{t-2} + \delta_3 Y_{t-3}$$

A. Compute the optimal values of Λ_0, λ_1, λ_2 and λ_3. What is the mean-squared error attained under this rule?

B. Compute the optimal values of Δ_0, δ_1, δ_2, and δ_3 assuming that the interest rate rule (2) is used. What is the mean-squared error attained under this rule?

C. What would the monetary authority do, peg r or peg m? What feature of the above model is critical in accounting for this result?