

EDUCATION AND EXAMINATION COMMITTEE

OF THE

SOCIETY OF ACTUARIES

COURSE 2 STUDY NOTE

MACROECONOMICS

by

Paul Wachtel

Copyright 1997 by the Society of Actuaries

The Education and Examination Committee provides study notes to persons preparing for the examinations of the Society of Actuaries. They are intended to acquaint candidates with some of the theoretical and practical considerations involved in the various subjects. While varying opinions are presented where appropriate, limits on the length of the material and other considerations sometimes prevent the inclusion of all possible opinions. These study notes do not, however, represent any official opinion, interpretations, or endorsement of the Society of Actuaries or its Education and Examination Committee. The Society is grateful to the authors for their contributions in preparing the study notes.

**2-21-00
FOURTH PRINTING**

Printed in U.S.A.

TABLE OF CONTENTS

PREFACE	i
CHAPTER I—ECONOMIC MEASUREMENT	1
NATIONAL INCOME AND PRODUCT ACCOUNTS.....	1
Basic Definitions and Concepts	1
The Formal Accounts	2
The Savings/Investment Identity	4
Nominal and Real GDP.....	9
Index Number Theory	9
Index Numbers and the Measurement of GDP Growth and Inflation	11
OTHER MAJOR PRICE INDEXES	14
Data Sources	15
Data Presentation	16
BALANCE OF PAYMENTS.....	17
BUSINESS CYCLES.....	18
Business Cycle Defined.....	18
Cycle Dating	19
Cycle Indicators	22
INTEREST RATES AND EXCHANGE RATES	23
Interest Rates and the Yield Curve	23
Exchange Rates	25
CHAPTER II—GROWTH, PRODUCTIVITY AND LONG-RUN EQUILIBRIUM	29
ANALYZING GROWTH	31
PRODUCTIVITY.....	33
Productivity Trends.....	33
Sources of Economic Growth.....	34
How to Increase Productivity Growth.....	36
MODEL OF LONG-RUN EQUILIBRIUM.....	38
Output Equilibrium.....	39
SUMMARY.....	41
Distribution of Output and Interest Rates.....	41
Money & Prices.....	45
SUMMARY.....	48
Open Economy Equilibrium Conditions	48
CHAPTER III—UNDERSTANDING SHORT-RUN FLUCTUATIONS	51
QUANTITY ADJUSTMENT PARADIGM	53
THE PLANNED AGGREGATE DEMAND MODEL.....	55

THE SIMPLE KEYNESIAN MULTIPLIER.....	57
DETERMINANTS OF AGGREGATE DEMAND.....	58
Consumption.....	58
Investment.....	59
Foreign Sector.....	60
THE KEYNESIAN REAL SECTOR MODEL.....	60
Properties of the <i>IS</i> Curve.....	62
MONETARY SECTOR.....	63
Money Demand Function.....	63
Money Market Equilibrium.....	64
Properties of the <i>LM</i> Curve.....	65
Equilibrium Adjustments.....	67
SUMMARY OF THE <i>IS-LM</i> MODEL.....	68
<i>IS</i> Curve.....	68
<i>LM</i> Curve.....	68
ANALYSIS OF MONETARY AND FISCAL POLICY.....	69
Fiscal Policy.....	69
Monetary Policy.....	71
Special Cases.....	71
THE KEYNESIAN MODEL IN THE OPEN ECONOMY.....	72
Exchange Rates.....	72
Monetary Policy in an Open Economy.....	73
Fiscal Policy in an Open Economy.....	74
THE NEXT STEP: BACK TO LONG-RUN EQUILIBRIUM.....	74
CHAPTER IV—INFLATION AND THE COMPLETE MACROECONOMIC MODEL.....	75
TOTAL DEMAND AND SUPPLY ANALYSIS.....	75
Total Demand Curve.....	75
Total Supply Curve.....	77
Keynesian Supply Curve.....	77
Long-run or Classical Supply Curve.....	78
Some More Reasons Why Prices are Sticky.....	80
A DEMAND SHOCK.....	82
A Policy Expansion: Step by Step.....	82
UNDERSTANDING INFLATION MOMENTUM.....	84
Price Adjustment Function.....	85
Phillips Curve.....	85
Expectations of Inflation.....	86
Augmented Price Adjustment Equation.....	88
Summary.....	92
RATIONAL EXPECTATIONS.....	92
New Classical Macroeconomics.....	93
SOURCES OF INFLATION.....	94

Monetary Growth.....	94
Excess Demand	95
Relative Price Shocks.....	95
Wage Price Spiral.....	95
Inflation Expectations.....	96
CHAPTER V—BANKS, MONEY AND MACRO POLICY.....	97
BANKS AS CREATORS OF MONEY.....	97
Fractional Reserve Banking.....	98
The Money Multiplier.....	100
STRUCTURE OF THE FEDERAL RESERVE.....	103
TOOLS OF MONETARY POLICY.....	106
Open Market Operations	106
Discount Rate	108
Reserve Requirements	109
EFFECTS OF MONETARY POLICY.....	110
FISCAL POLICY.....	112
Fiscal Policy Activism: 1960s and 1970s	113
Policy Lags	114
Measuring Fiscal Policy.....	115
EFFECTS OF FISCAL POLICY.....	115

PREFACE

Modern textbooks in Macroeconomics are more often than not weighty tomes of 700 or more pages. Such encyclopedic treatments are hardly user friendly, particularly for students who are undertaking a program of self-study. My aim in preparing this study note has been two-fold. First, I have written a study note that individuals preparing for the actuarial exams can approach on their own without a course syllabus to direct them to the important parts. Second, I present a thorough, but concise, presentation of the major aspects of modern macroeconomic thought and provide enough relevant examples to bring macroeconomics to life in the mind of the reader. I hope that I have achieved these aims and that the readers will find the study note and their exam preparation to be a relevant learning experience as well.

I would like to express my thanks to the Society of Actuaries and Judy Strachan in particular for their patience as I have slowly re-worked the original 1991 study note. In addition, my appreciation to Arjun Jayaraman and Michelle Quinn for their assistance in the preparation of the manuscript. Finally, thanks to Irwin Vanderhoof who was kind enough to introduce me to the Society and to Richard Mattison who was the education actuary at that time.

Finally, I dedicate this monograph to my family—my wife Claire, my son Chaim, and my daughter Rachel. They make life both challenging and worth living and I appreciate their leaving me just enough time to finish this manuscript.

Paul Wachtel
New York
January 1997

CHAPTER I

ECONOMIC MEASUREMENT

NATIONAL INCOME AND PRODUCT ACCOUNTS

Basic Definitions and Concepts

The National Income and Product Accounts (NIPA) are a vast accounting scheme for aggregate economic activity. In the United States they are prepared by the Bureau of Economic Analysis (BEA) of the Department of Commerce. We will discuss the basic elements of the income and product accounts that are used to measure the overall, or aggregate, level of economic activity.

Economic activity gives rise to both output and income earned by the persons and machines involved in the productive activity. The overall level of economic activity can be measured by adding up either the value of output produced or the levels of income earned. The most common aggregate measure is the product side calculation of Gross Domestic Product (GDP). On the income side of the accounts, the measure of aggregate activity is National Income (NI). We will start with a conceptual definition of each measure. We will then show how the measures relate to one another and derive some important accounting relationships that utilize information from each.

On the product side, Gross Domestic Product (GDP) is defined as the market value of all final goods and services produced in a given time period by labor and property located within the U.S. Prior to the 1991 revisions to the NIPA, GNP (Gross National Product) was the aggregate measure that was most commonly used. GNP measures output produced by the labor and property of U.S. residents, regardless of where the labor and property are located. In 1990, GNP was about 0.2% greater than GDP. The aggregate emphasis was switched from GNP to GDP because (a) GDP is a more appropriate measure for tracking changes in economic activity and (b) most other countries use a GDP concept.

The key words in our definition are *product*, *final*, and *market value*. By product, we mean the consequences of a current act of production and we exclude the transfer of existing assets. By final product we mean output absorbed by the ultimate users of goods and services. This excludes intermediate products that are used as inputs in production processes. The concept is straightforward, but the measurement problems can be complex. A given product, say, sugar, can be used as a final product or as input into further production processes. Thus, BEA statisticians must have a mechanism for distinguishing between the sugar bought for household consumption and the sugar purchased by the local bakery for use in its production processes. In the latter case, the final product is the cake sold by the bakery.

The last key phrase is market value, which means simply that all of the product is valued at its market price. The dollar value of output is determined by its dollar market price. For virtually all output of the business sector, this is quite straightforward. However, there are instances in which there is no observed market price for output and the NI statisticians must impute a value. For example, there is no explicit market price for the financial services provided to demand (checking) deposit holders in lieu of interest earnings on their balances. Therefore, an imputation for the value of bank output is included in the NI accounts.

In one large and important sector it is not possible to value output at its market price because there is not a marketplace where these goods and services are sold. This sector is the government, which produces a broad array of services. The police force, for example, uses inputs such as labor services and automobiles to produce protection and law enforcement, an output that is very difficult to value. Production in the government sector is therefore valued at the costs of the inputs.

The standard measure on the income side of the accounts is *National Income (NI)*. It is a counterpart to GDP because the value of a product is equal to costs of production (which includes the income of providers of labor or other services) plus the profits earned in production. Thus, National Income is income earned in productive activity by all the factors of production: compensation of employees and the earnings of capital (profits, rental income, proprietors' income, and net interest earnings). There are also some technical differences between National Income and GDP that will be noted in the summary of the formal accounts below.

The key words in the definition of NI are *income earned in production*. This distinguishes income in a NIPA sense from accounting income or receipts that result from transfers or from the sale of assets. Thus, the \$100 that my mother gave me for my birthday is a transfer and not income earned in production. Similarly, the \$250 in cash that my neighbor gave me for my old car is receipts from an asset sale and not income earned in production.

The Formal Accounts

Table 1 shows a summary of the NIPA with many of the categories shown in the official accounts. Some of the major entries are explained below. Note that this overall income and product account is only the tip of the iceberg. A wealth of further detail and additional accounting statements are available. These include the value-added accounts, detailed accounts for specific sectors, and breakdowns of economic activity by industry. In addition to current data, the BEA maintains a vast historical record extending back to 1929 for annual data on the broad aggregates as well as quarterly data since 1947.

Personal Consumption Expenditure	4,698.7
Gross private domestic investment	1,014.4
Producer's durable equipment and nonresidential structures	667.2
Residential structures	287.7
Changes in business inventories	59.5
Net exports of goods and services	-96.4
Exports	722.0
Imports	818.4
Government purchases of goods and services	1,314.7
Federal	516.3
State and local	<u>798.4</u>
GROSS DOMESTIC PRODUCT	6,931.4
Plus: net receipts of factor income from rest of world	-9.0
Equals:	
GROSS NATIONAL PRODUCT	6,922.4

Minus: Consumption of fixed capital	818.8
Minus: Indirect business taxes, business transfer payments, statistical discrepancy, misc.	<u>608.6</u>
Equals:	
NATIONAL INCOME	5,495.1
Compensation of employees	4,008.3
Proprietors' income with adjustments	450.9
Rental income of persons with adjustments	116.6
Corporate profits with adjustments	526.5
Profits before tax	528.2
Profits tax liability	195.3
Dividends	211.0
Undistributed profits	121.9
Inventory valuation and capital consumption adjustments	-13.3
Net interest	392.8
<i>Note:</i> Data are billions of nominal dollars.	

Although a definition of all the accounting concepts is beyond the scope of our present discussion, we will examine some of the elements of Table 1. At the top of the table, the major components of GDP—the product account—are shown. Output of goods and services is allocated among the final absorbing sectors: households, business, government, and the foreign sector. About two-thirds of total output is absorbed by the consumption sector. The expenditures of the domestic sector include imports, goods, and services produced abroad. Imports, however, are not a part of GDP, which is the output produced by U.S. factors of production. Thus, imports are subtracted from exports to yield a net absorption of output by the foreign sector in the product account.

An important accounting convention in the product account is the way in which owner-occupied housing is treated. The purchase of a home is clearly not an expenditure of current consumption but rather an investment in a capital good that provides housing services. The entire output of the housing industry appears under the residential structures category of investment, and there are two additional entries, or imputations, which guarantee that the accounting treatment of owner-occupied housing is appropriate. The first imputation is for the rental services of owner-occupied housing. It is a part of the service expenditures component of personal consumption expenditures. The second is an imputation for rental income that appears on the income side of the accounts. This accounting convention means that rental housing and owner-occupied housing are treated symmetrically in the accounts. In the case of owner-occupied housing, the individual engages in a fictitious business which buys a productive asset—the house—rents it out (a service expenditure on rent), and receives rental income which accrues to the owner of the house.

In the middle of the table are three items which represent the wedge between the aggregates on the product and income sides of the accounts, GDP and NI, respectively. First among these is the net receipts of factor income from the rest of the world which relates GDP to GNP. Second is the allowance for the consumption of fixed capital that is subtracted because the capital income components of NI are reported net of depreciation. (Net of depreciation means that allowances for the wearing out of capital equipment has been subtracted.) The allowances are based primarily on depreciation estimates from tax returns. Third among these are the indirect business taxes (sales and excise taxes) which are part of GDP but not NI. They are part of GDP because they are included in the valuation of final output. However, they do not accrue to any factor of production as income does. There are other items in this wedge, such as transfer payments to individuals by the business sector and a statistical discrepancy. The discrepancy is quite small and arises because the income and product side measurements of overall economic activity are prepared independently.

At the bottom of the table the major components of NI are shown. About three-quarters of NI is compensation of employees, or remuneration from work. The next two categories represent income received by individuals as proprietors or partners and income from the rental of real property. Corporate profits are the next item, and a breakdown is shown that consists of before-tax profits plus the accounting adjustments for the value of inventories and the adjustment to depreciation allowances. Before-tax profits consist of profit taxes paid dividends, and undistributed profits. The final element of NI is net interest paid by business, which consists of interest paid minus interest received plus net interest received from abroad. Thus, net interest represents another form of the return to capital.

The Savings/Investment Identity

The formal accounts include more detail than is necessary for macroeconomic analysis. Since the NIPA provide the backbone for macroeconomic analysis as it will be presented, a simplified accounting framework which retains all the basic concepts used in macroeconomic theory will be useful.

A simplified set of accounting identities will be presented here. We will derive the identity between aggregate investment and saving which is an important building block for later analyses. The identity relates the income and product sides of the accounts since investment is part of output and saving is the part of income that is not spent.

We begin with a definition of GDP, which corresponds to the top half of Table 1. GDP is the sum of product absorbed by the consumption (C), gross private domestic investment (I), government (G), and foreign sectors (X – M):

$$GDP = C + I + G + (X - M)$$

The net absorption by the foreign sector is exports (X) less imports (M), which are included in the expenditures of the domestic sectors.

GDP and NI are equivalent measures of aggregate activity except for the wedge that consists of net receipts of factor income from abroad (NR), depreciation allowances (CC), and indirect business taxes (IBT). That is:

$$GDP + NR - CC - IBT = NI$$

We now introduce the concept of Disposable Income (DI) which consists of resources available for spending by individuals. First, add income received by individuals, which is not included as income earned in production:

- (i) transfer payments from the government (TR) and
- (ii) interest paid on the government debt (INT).

Second, subtract those parts of NI which are not received by individuals:

- (i) retained earnings (RE); note that corporate profits are equal to retained earnings (RE) plus corporate tax payments (Tc) plus dividends,
- (ii) personal and corporate tax payments (Tp and Tc, respectively), and
- (iii) interest on the government debt paid to foreigners (INF).

Thus, we have

$$DI = NI + TR + INT - RE - Tp - Tc - INF$$

To complete the derivation, three additional substitutions are needed. First, substitute in the above identity for national income, $NI = GDP + NR - CC - IBT$. Second, for disposable income substitute its components—consumption and personal savings, $DI = C + PS$. Finally, substitute for GDP the sum of the expenditure categories, $GDP = C + GI + G + (X - M)$. This yields

$$C + PS = C + I + G + (X - M) - CC - (IBT + Tp + Tc) - RE + TR + INT - INF + NR$$

Rearranging terms, we arrive at the identity between gross investment on the left and national saving on the right:

$$I + (X - M - INF + NR) = PS + (RE + CC) + (Tp + Tc + IBT - G - TR - INT)$$

or

$$\begin{aligned} &\text{Gross private domestic investment} + \text{Net foreign investment} \\ &= \text{Personal saving} + \text{Business saving} + \text{Government saving} \end{aligned}$$

- *Net foreign investment*, $NFI = (X - M - INF + NR)$, is the excess of receipts from foreigners (from exports, X, and income from factors of production abroad, NR) over payments to foreigners (for imports, M, and interest payments, INF). The flow of payments results in a change in the investment position of the country. Thus, NFI is also the increase in U.S. claims on foreigners minus the increase of foreign claims on the United States. These two ways of looking at NFI—the current flow of payments or the change in capital holdings—correspond to the current account and capital account of the Balance of Payments (see below).

If $NFI > 0$, then the U.S. receipts from abroad exceed payment to foreigners and the U.S. is accumulating assets abroad. As the term implies, there is investment and asset accumulation abroad.

- *(Gross) Business saving* is $BS = RE + CC$
- *Government saving (the government surplus, GS)* is equal to tax receipts ($Tp + Tc + IBT$) minus government outlays ($G + TR + INT$). It has, of course, been negative in recent years as the government has run a deficit or “dissaved.”

In summary, the saving-investment identity is:

$$I + NFI = PS + BS + GS$$

A summary of the derivation is shown in Table 2 below:

TABLE 2The NIPA Saving Investment Identity

Gross Domestic Product

$$GDP = C + I + G + (X - M)$$

$$\text{National Income} = GDP + NR - CC - IBT = NI$$

$$NR = \text{Net factor income from abroad}$$

$$CC = \text{Capital consumption}$$

$$IBT = \text{Indirect business taxes}$$

$$\text{Disposable Personal Income} = DI = NI + (TR + INT) - (Tp + RE + Tc + INF)$$

$$TR = \text{transfer payments by government}$$

$$INT = \text{interest paid on Government debt}$$

$$Tp, Tc = \text{personal and corporate tax payments}$$

$$RE = \text{retained corporate earnings}$$

$$INF = \text{interest on govt. debt paid to foreigners}$$

$$DI = C + PS$$

$$PS = \text{Personal Saving}$$

$$I + (X - M - INF + NR) = PS + (RE + CC) + (Tp + Tc + IBT - G - TR - INT)$$

$$I + NFI = PS + BS = GS$$

$$I = \text{Gross private domestic investment}$$

$$NFI = X - M - INF + NR \text{ Net foreign investment}$$

$$BS = RE + CC \text{ Business saving}$$

$$GS = \text{Taxes} - \text{Government outlays} = \text{Government saving}$$

$$\text{Taxes} = Tp + Tc + IBT$$

$$\text{Government outlays} = G + TR + INT$$

Trends in saving and its components (as ratios to GDP) are shown in Figure 1 on the next page. In the Figure, GPS is gross private saving or personal plus business saving ($GPS = PS + BS$). GPS is larger than national saving ($GPS + GS$) because government saving is more often than not negative. The discrepancy is large when recessions lead to large deficits (e.g., 1974–75) and after the tax cuts of the early 1980s. National Saving and Gross Private Domestic Investment differ due to Net Foreign Investment. The two track closely until 1982 when investment becomes significantly larger. The data are summarized in Table 3, which shows recent historical trends in U.S. national savings and investment and its major components.

Figure 1
US Savings and Investment (as a fraction of GDP)

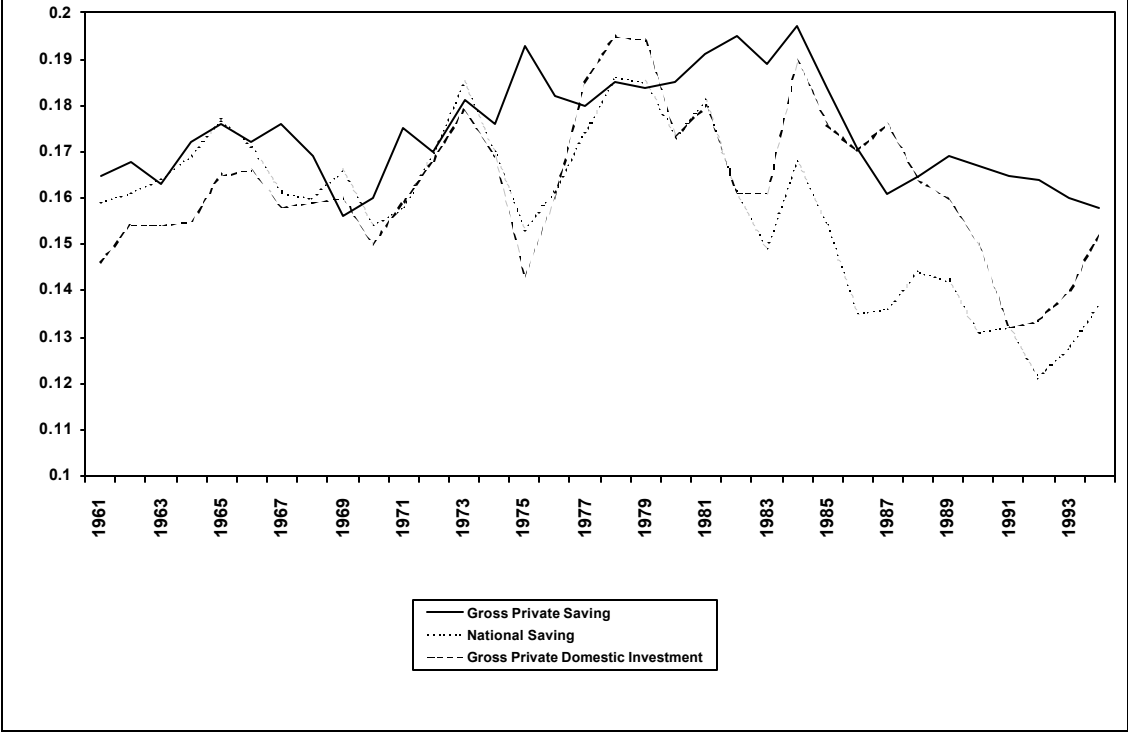


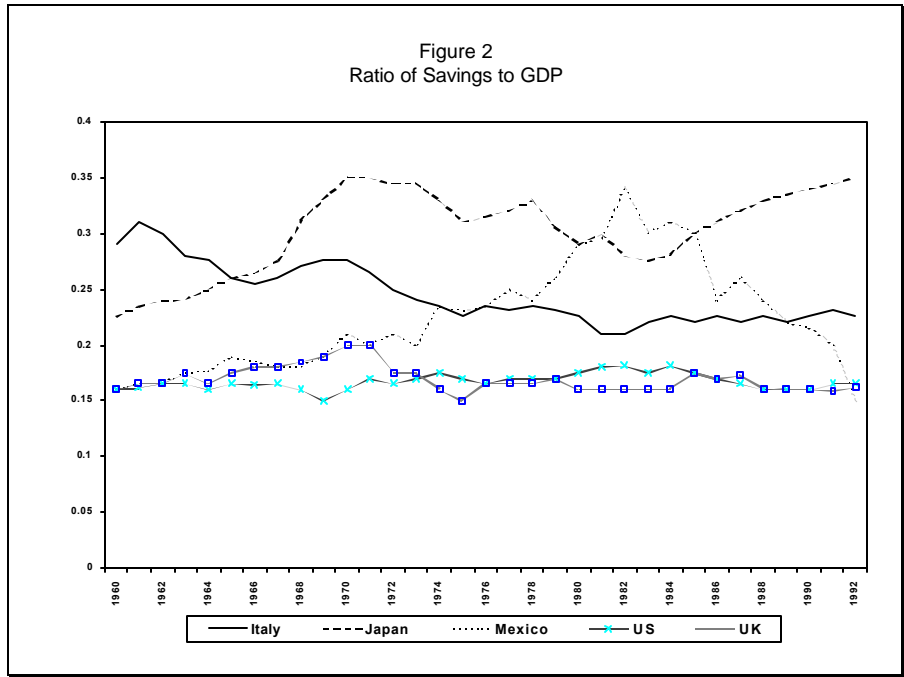
TABLE 3
SAVINGS AND INVESTMENT TRENDS
AS A PERCENTAGE OF GDP

	<u>1961–65</u>	<u>1966–70</u>	<u>1971–75</u>	<u>1976–80</u>	<u>1981–85</u>	<u>1986–90</u>	<u>1991–94</u>
Gross Private Domestic Investment, I	15.6	15.9	16.4	18.1	17.4	16.0	13.9
• Business investment	9.5	10.6	10.8	12.3	12.8	11.4	10.5
• Residential investment	6.1	4.9	5.6	5.8	4.6	4.6	3.4
Net foreign investment, NFI	0.9	0.4	0.4	0.0	-1.2	-2.5	-1.1
National saving	16.7	16.2	16.7	17.6	16.2	13.7	12.8
• Business saving, BS	12.3	11.6	11.9	13.3	13.5	12.8	12.5
• Personal saving, PS	4.6	5.1	6.0	5.0	5.6	3.3	3.5
• Government saving,	-0.2	-0.4	-1.2	-0.8	-2.9	-2.4	-3.2
Federal	-0.2	-0.5	-1.8	-1.9	-4.1	-3.2	-3.6
State and local	0.0	0.1	0.6	1.1	1.2	0.8	0.4
Net saving, PS+BS+GS – CC	8.2	7.8	7.3	6.9	4.3	2.7	2.0
Statistical discrepancy, I + NFI – (PS + BS + GS)	-0.2	0.1	0.1	0.5	0.0	-0.2	0.0

From Figure 1 and Table 3, take note of the following important changes:

- (i) the trend in savings, particularly net saving
- (ii) the trend in gross investment
- (iii) the growth of NFI and GS
- (iv) changes in PS

Finally, there are big differences in the savings and investment behavior of different economies. The national savings rates for five important economies for the last 30 years are shown in Figure 2. The saving rate in Japan has been over 30% for most of the period while the savings rates in both the U.S. and the U.K. are between 15 and 20%.



Nominal and Real GDP

The NIPA accounts defined above measured income and output at current prices. Since prices as well as quantities change over time, nominal GDP is not a useful measure of the level of economic activity or output. We need a *real* measure that holds price change constant.

If the quantities of final goods and services produced in period t are denoted by Q_{it} and their prices are P_{it} , then (nominal or current dollar) GDP is simply $\sum_i Q_{it}P_{it}$. Real GDP in period 0 prices would be calculated by making the same calculation with period 0 prices. Thus, real GDP is $\sum_i Q_{it}P_{i0}$. The ratio of nominal GDP to real GDP is called the *implicit price deflator*.

If life were simple, we would measure real growth rates by simply taking the percentage change in real GDP, and we would measure inflation rates by simply looking at the percentage change in the implicit price deflator. Alas, life is sometimes more complicated, and neither percentage change is a satisfactory measure of the desired concept. In fact, the standard measures of real growth rates and inflation rates for GDP published by the Commerce Department and used by forecasters and analysts are somewhat different. Because measuring the growth rate real GDP and inflation are complicated issues, we will first discuss index number theory generally.

Index Number Theory

A price index is generally defined as a weighted average of price relatives for the individual components of the index. The price relative is simply the ratio of the current price of the i -th item (P_{it}) to its price in the base period (P_{i0}). Thus the value of the index in period t is

$$I_t = \sum_i w_{it} \frac{P_{it}}{P_{i0}}$$

where w_{it} is the weight applied to the i -th item for the t -th period. The index obviously has a value of one in the base period.

Index numbers differ primarily in regard to the choice of the weights applied to the price relatives. There are two types of weights in common use, named after the nineteenth-century statisticians who introduced them, Laspeyres and Paasche.

A Paasche index uses weights based on current period quantities. The weight for the i -th good is its shares of current (t -th period) expenditures:

$$w_{it} = \frac{Q_{it}P_{i0}}{\sum Q_{it}P_{i0}}$$

The Paasche index can then be written as:
$$I_t = \frac{\sum Q_{it}P_{i0}}{\sum Q_{it}P_{i0}}$$

It is the ratio of the cost of the current period market basket to its cost in the base period. For the whole economy, the numerator would be GDP and the denominator would be current output valued at period 0 prices. In other words, the Paasche index for GDP is the implicit price deflator described above.

A Laspeyres index uses weights, which are constant over time and are equal to the i -th good's share in total expenditure in the base period. Let Q_{it} be the quantity of the i -th good purchased in period t . Thus, $Q_{i0}P_{i0}$ is the base period expenditure on the i -th good and the Laspeyres weights are given by:

$$w_i = \frac{Q_{i0}P_{i0}}{\sum_i Q_{i0}P_{i0}}$$

Substituting for the weights in the general formula for an index number yields the definition of a Laspeyres index:

$$I_t = \frac{\sum Q_{i0}P_{it}}{\sum Q_{i0}P_{i0}}$$

The index is the ratio of the current cost of the base period market basket to its cost in the base period.

The difference between the two calculations, Laspeyres and Paasche, is simply a matter of whether the market basket is fixed at base period expenditure patterns or changes continuously to reflect changes in expenditure patterns. The choice between the two schemes is not arbitrary but has important implications.

The fixed-weight or Laspeyres index calculation applies the same weights to the prices of an unchanging basket of goods, and thus overlooks the major way people cope with inflation. Consumers search for substitutes for goods whose prices increase more rapidly than the prices of other goods. If they substitute away from such goods, then the share in total expenditures of goods with large price increases may decline. Thus, the fixed-weight index, which maintains the base period expenditure weights, will overstate the true rate of inflation.

The opposite argument can be made with respect to the Paasche index calculation. The changing weights fully reflect any substitution away from goods which have become relatively more expensive. Thus, the

current period expenditure scheme may understate the true inflation rate. Neither index is perfect, but for some purposes, one may be more appropriate than the other.

Chain linked price indexes were developed as a mix of the Lasperyes and Paasche calculations. The level of the index is defined by:

$$I_t = \frac{I_{t-1}(\sum Q_{i,t-1}P_{it})}{(\sum Q_{i,t-1}P_{i,t-1})}$$

In each period, the index compares price changes with quantity weights from the immediately past period.

The rate of change of this index, $(\frac{I_t}{I_{t-1}}) - 1$, depends on price changes between t and $t-1$ and the quantity weights used are constant and relevant to contemporaneous experience. The chain linked index approach solves some of the problems encountered in measuring growth rates and inflation.

Index Numbers and the Measurement of GDP Growth and Inflation

Consider first the inflation rate implied by the implicit price deflator, a Paasche index. The implicit price deflator for period t is the ratio of current GDP to current output valued at base period prices:

$$\frac{\sum_i Q_{it}P_{it}}{\sum_i Q_{it}P_{i0}}$$

The change in this measure from period $t-1$ to t would be influenced by changes in the Q_{it} 's as well as changes in prices, P_{it} 's. Thus, the rate of inflation calculated this way would reflect changes in the composition of product as well as price change.

Thus, a better measure of inflation is calculated from a fixed weight deflator (a Laspeyres index). The fixed weight deflator, using period 0 as the base period, is defined by:

$$\frac{\sum_i Q_{i0}P_{it}}{\sum_i Q_{i0}P_{i0}}$$

The fixed weight deflator compares the value of base period production valued at current prices to base period GDP (the denominator). The percentage change in this fixed weight deflator depends only on changes in prices and changes in the composition of output. For many years the rate of change in the fixed weight deflator was used as the standard measure of inflation in the overall economy.

The drawback of the fixed weight deflator is that after a few years, the base period weights (i.e. the quantities Q_{i0}) may not be a good reflection of current expenditure patterns. Thus, it is necessary to change the weights about once a decade so that they are relevant to the patterns of expenditures and output in the economy.

In the 1990s government economists began to realize that there were similar problems with the definition of the real growth rate. If real GDP is defined as $\sum_i Q_{it}P_{i0}$, then the rate of growth in real GDP between $t-1$ and t is the ratio of real GDP $_t$ to real GDP $_{t-1}$ less one. Looking at the ratio, we see that real GDP growth uses base period price weights to compare quantities in $t-1$ to those in t .

$$\frac{\sum_i Q_{it} P_{i0}}{\sum_i Q_{it-1} P_{i0}}$$

This may not be an accurate measure of growth if there have been large changes in relative prices from the base period.

The single most dramatic relative price change in recent decades has been the rapid decline in the price of computer equipment. Computer prices in 1995 were only about 35 percent of their 1987 level. Over the same time period, other prices increased about 30 percent. A fixed (1987) weight calculation of real GDP overstates the importance of computer production because the quantity of computers produced has increased enormously and the weight given is the (high) base period (1987) price. Thus, the rate of change of real GDP overstates the amount of real growth that has occurred because the increase in computer production is being given too much weight. The effect of this particular anomaly (the decline in the relative price of computers accompanied by the enormous increase in output) on real growth measures has been so dramatic that the national income statisticians have introduced a new growth measure for real GDP.

Since the start of 1996, the preferred measure of real GDP growth is a *chain-linked measure*, which uses contemporaneous prices to weight quantity changes. Thus, the price weights will change over time and reflect the relative importance of different products in current production. For example, the chain-linked measure of the rate of growth in real GDP from $t-1$ to t can be defined as

$$\frac{\sum_i Q_{it} P_{it-1}}{\sum_i Q_{it-1} P_{it-1}} - 1.$$

(The actual calculation in use is slightly more complex because it uses an average of prices in t and $t-1$.) Between 1990 and 1994 real growth averaged 2.2% using the old measure and 1.8 percent using the new chain-type measure.

The chain-type calculation has another advantage over the fixed base period calculation of real growth. With fixed base period weights, the historical picture of real growth was revised whenever the base period was changed. The base period has been changed seven times in the post-war period in order to keep up with the changing patterns of production and expenditure. Repeated revisions have made recessions appear less severe than when first reported. For example, the decline in real output in 1974 was originally 1.4 percent with 1972 base year prices, but only 0.6 percent with 1987 base year prices. These revisions do not occur with chain-type calculations where the relevant contemporaneous prices are retained for all calculations.

GDP GROWTH AND COMPUTERS

Consider an economy that produces two goods—potato chips and computer chips: Notice that computer output has increased rapidly while computer prices have fallen down dramatically. Potato chip production and prices have increased moderately.

year	<u>Potatoes</u>		<u>Computers</u>	
	Bags	Price	Number	Price
1	10000	2.00	1	10000
2	11000	2.50	2	5000
3	12000	3.00	4	2500
4	13000	3.50	8	1250

* The calculation of nominal GDP for each year is simple:

year 1	10000 (2)	+	1(10000)	=	\$ 30000
year 2	11000 (2.50)	+	2 (5000)	=	\$ 37500
year 3	12000 (3.0)	+	4 (2500)	=	\$ 46000
year 4	13000 (3.50)	+	8 (1250)	=	\$ 55500

* Real GDP can be calculated using any year as the base year. The left column uses year 1 prices while the right column uses year 4 prices.

Real GDP—Fixed Base Year Weights

	Year 1 Prices	Year 4 Prices
1	10000 (2) + 1 (10000) = 30000	10000 (3.50) + 1 (1250) = 36250
2	11000 (2) + 2 (10000) = 42000	11000 (3.50) + 2 (1250) = 41000
3	64000	47000
4	106000	55500

Growth Rate of Real GDP—Fixed Base Year Weights

2	40.0%	13.1 %
3	52.4%	14.6%
4	65.6%	18.1%

Clearly the choice of base year weights (year 1 prices on the left and year 4 prices on the right) has a dramatic effect in the calculation of the real GDP growth rate. The growth rate depends on the relative importance of computer chip production in total output.

* A better measure of the GDP growth rate is given by a chain weighted growth rate calculation. For example, for real growth between any two years, say A and B, use real GDP calculated with year A prices.

Growth Rate of Real GDP—Chain Weighted

Growth between years 1 and 2; use real GDP calculated with year 1 prices:

$$(42000 - 30000) / 30000 = \mathbf{40.0\%}$$

Growth between years 2 and 3; use year 2 prices:

$$\text{Year 3 GDP with year 2 prices} = 12000 (2.5) + 4 (5000) = 50000$$

$$\text{Year 2 GDP (nominal)} = 37500$$

$$(50000 - 37500) / 37500 = \mathbf{33.3\%}$$

Growth between years 3 and 4; use real GDP calculated with year 3 prices:

$$\text{Year 4 GDP with year 3 prices} = 59000$$

$$\text{Year 3 GDP (normal)} = 46000$$

$$(59000 - 46000) / 46000 = \mathbf{28.3\%}$$

OTHER MAJOR PRICE INDEXES

Consumer Price Index The Consumer Price Index (CPI) is probably the most commonly cited index. It measures the price of a representative basket of goods purchased by consumers. The purpose of the CPI is to track the price of a given basket of goods and therefore it utilizes a Laspeyres, fixed weight calculation.

The CPI is calculated monthly by the Bureau of Labor Statistics, which collects a vast amount of price data from surveys of all types of selling establishments. The fixed weights reflect the importance of each good in the typical consumer's budget. The market basket is determined from periodic Consumer Expenditure Surveys, which are used to update the weighting scheme about once a decade.

The CPI is widely used to measure price change and is carefully watched. However, in recent years there has been considerable discussion that implies that the CPI overstates the "true" increase in the cost of living. In fact, there have been suggestions that the use of the CPI to index (or make automatic inflation corrections) to payments like social security pensions be reduced. If the CPI overstates inflation then social security payments have been rising too rapidly and that is something that a deficit-ridden budget can ill afford. An important reason why the CPI overstates inflation is the treatment of quality improvements. Consider an example such as washing machines. Imagine that a manufacturer introduces a new version that is the same size as its predecessor and looks the same as well. The price of the new model is \$50 or 10 percent higher than the earlier model. On first glance, the statisticians collecting price data for the calculation of the CPI will record a price increase of 10 percent. However, a closer look indicates that some modifications to the motor assembly makes the machine consume less electricity which can lead to substantial lifetime savings in energy costs. Furthermore, a re-design of the drum and plastic parts makes this machine less likely to tear clothing. This is clearly a better machine. Now does the \$50 price increase reflect the costs of a higher quality machine or does it reflect a higher price of washing machines? In all likelihood, the increase reflects both and the statistician must determine how much of the \$50 to attribute to price change and how much to attribute to quality improvements. It is clear now that conservative government statisticians have been systematically underestimating the importance of quality improvements.

Another issue that influences the CPI is the way statisticians handle new products. Many new products (computers, calculators, etc.) decline in price for a period of time after their initial introduction. If the index number calculations are slow to introduce new products, then these periods of often rapid price decline are systematically being overlooked. Again, the measured CPI will overstate the inflation rate.

The possibility of bias in the CPI is not just a statistical nuance; it has significant political ramifications as well. In fact, the Senate Finance Commission set up a commission of economists chaired by Michael Boskin to study the issue. The commission reported in December 1996 that the CPI probably exaggerates true increases in the cost of living by about 1.1% per year. A correction to the CPI of this order of magnitude has enormous ramifications for the government budget since about a third of Federal spending (mostly retirement programs) is indexed to changes in the CPI and revenue are affected as well through the indexation of tax brackets. The Boskin commission estimated that correcting for a 1.1% per year overstatement in the CPI would save the government around a trillion dollars over the next 12 years. Clearly many politicians would like to rely on such a statistical correction to balance the budget.

The commission report says that about a third of the estimated bias is due to the fact that the CPI (a fixed weight Laspeyres calculation) does not take account of important changes in spending patterns. The Bureau of Labor Statistics, the agency that calculates the index, notes that the CPI is not a cost of living index. However, the remaining biases are due to the fact that Americans do more and more shopping at discount stores and benefit very rapidly from an expanding array of new and better products.

Producers Price Index The Producers Price Index (PPI) is another fixed weight monthly index. It measures goods and commodities purchased by producers. It is important as a measure of changes in the costs of production.

Data Sources

Now that we understand the definition of GDP, as well as technical aspects of the calculation, we can turn to a very practical and mundane question. Where does the data come from? Continuous measurement of all economic activity would be prohibitively expensive. The NIPA data are valuable but not sufficiently valuable to warrant the expense of monitoring every economic transaction. The product side of the account relies on data from various surveys and samples which are combined with benchmarks constructed from periodic economic censuses. For the income side of the account, there is a mechanism that collects virtually complete data on income. The mechanism is the tax collection system, which generates many data that are used to calculate national income.

It is, of course, important for both forecasting and policy purposes to have timely estimates of the data. Quarterly estimates of the basic aggregates are prepared by the BEA, and the initial public releases of the information are carefully watched indicators of the state of the economy. However, the initial estimates are based on incomplete information and are therefore revised, often substantially, as additional information becomes available. There is a trade-off between speed of preparation and accuracy of the data. In recent years government statisticians have been criticized because the data originally released have often been subject to substantial revision. The problem is sometimes so severe that the early data may bear little information about trends that are actually emerging.

The first available data (the “advance”) are released during the month after the quarter’s end. The “preliminary” estimate is released a month later and the “final” estimate a month after that. The data are then unchanged until complete information for the calendar year from tax and other sources can be used to make revisions. These annual revisions, which are usually released in midsummer, provide a complete reworking of the data for the previous calendar year and often include revisions for 2 prior years as well. Finally, the BEA undertakes a comprehensive revision project about every 5 years which introduces conceptual changes, uses new data sources and methods, and incorporates new benchmarks from the

economic censuses. The historical data are revised as far back as necessary when the major revision takes place.

Table 4 shows the advance, preliminary and final data for real GDP for some recent quarters. The revisions in these examples are not unusually large or small. The data shown are growth rates for the quarter, seasonally adjusted at annual rates (SAAR), as explained in the next section. The latest data (as of the end of 1996) reflect benchmark revisions and the shift to a “chained” calculation of growth rates.

TABLE 4

Rate of Growth of Real GDP (SAAR) and Announcement Dates

	<u>Advance</u>	<u>Preliminary</u>	<u>Final</u>	<u>Latest</u>
94I	2.6 (Apr. 28)	3.0 (May 27)	3.4 (June 29)	2.5
94II	3.7 (July 29)	3.8 (Aug 26)	4.1 (Sept 29)	4.9
94III	3.4 (Oct 28)	3.9 (Nov 30)	4.0 (Dec 22)	3.5
94IV	4.5 (Jan 27)	4.6 (Mar 1)	5.1 (Mar 31)	3.0

Data Presentation

There are several conventions that are used when GDP data are presented. First, the quarterly data are almost always presented as annual rates. That is, output for the quarter is shown as if it were maintained for a whole year. Second, the data are almost always shown after seasonal adjustment. That is, the influence of seasonal variations in activity is removed using the common procedures developed by the Census Bureau to estimate seasonal patterns from past data. Thus, GDP data are usually presented as SAAR or seasonally adjusted at annual rates.

Here are GDP data for two recent quarters:

	<u>1995-I</u>	<u>1995-II</u>
GDP (billions of dollars, SAAR)	7147.8	7196.5

The numbers do not mean that over \$7000 billion of final output was produced in each quarter. Instead, they tell us that if production in the quarter proceeded at the same rate for a whole year then GDP for the year would exceed seven trillion dollars. The growth rate for GDP in 1995-II (the percentage change from the previous period) is also an annualized growth rate, calculated by:

$$\left[\left(\frac{7196.5}{7147.8} \right)^4 - 1 \right] * 100 = 2.8$$

Generally, the convention used to calculate the SAAR percentage rate of change in any series, say X, is:

$$\left[\left(\frac{X_t}{X_{t-1}} \right)^4 - 1 \right] * 100.$$

Finally, for many discussions the quarterly growth rates are not the best measure of output or inflation trends. Growth rates can bounce around from quarter to quarter, so a broader picture can be obtained

from annual growth rates. Thus, it is common to use annual changes or rates of growth. For this purpose, it is common to examine the growth from the same quarter in the previous year. Furthermore, discussions of GDP growth in the year will use fourth quarter to fourth quarter growth. Thus, the preferred measure of GDP growth in 1995 is the percentage change between GDP in 1995-IV and GDP in 1994-IV. (In the past, analysts would calculate the annual growth rate as the percentage change in total output for the year.) Fourth quarter to fourth quarter growth rates provide a good measure of the growth that takes place over the calendar year.

BALANCE OF PAYMENTS

The balance of international payments is an important measure of international economic and financial activity. However, balance of payments data are notoriously difficult to measure. In the good old days, government statisticians had an easier time—they stood on the dock, observed every boat entering or leaving the harbor, and examined its cargo of goods and money (gold). Things are not so simple these days when goods enter in thousands of sealed containers and money gets transferred electronically. Nevertheless, the statisticians are able to chart the international position of the economy.

The balance of payments accounts are shown in Table 5 with data for 1994. Receipts of U.S. residents from abroad (e.g., for export sales and from income on overseas investments) is a positive entry and a payment made abroad (e.g., to pay for imports) is a negative entry.

TABLE 5

BALANCE OF PAYMENTS, 1994

CURRENT ACCOUNT

Receipts of U.S. residents from rest of world (ROW)	852
• Exports	715
- Merchandise	502
- Services	213
• Income received	137
(Interest, dividends and reinvested earnings of foreign affiliates of U.S. corporations)	
Payments to ROW	-967
• Imports	-820
- Merchandise	-669
- Services	-151
• Income paid	-147
Net unilateral transfers payments	-36
<u>CURRENT ACCOUNT BALANCE (Receipts + Payments + Net transfers)</u>	-151
MERCHANDISE TRADE BALANCE	-166

CAPITAL ACCOUNT

Capital outflow

• <i>Increase in U.S. assets abroad / Payment for foreign assets purchased</i>	
- Official reserve assets	5
- Other	-131

Capital inflow

•	<i>Increase in Foreign assets in U.S. / Receipt for U.S. assets purchased by foreigners</i>	
-	Foreign official	39
-	Other	252

CAPITAL ACCOUNT BALANCE 165

Statistical Discrepancy =

- (Capital account balance + Current account balance)

The current account balance is conceptually the same as net foreign investment in the NIPA. Recall from above that $NFI = X - M - INF + NR$ or net foreign investment is net exports plus net income received. (The NIPA presentation above ignored foreign transfers for the sake of simplicity.) The balance of payments accounts are calculated separately from the NIPA so there are differences between reported current account balances and NFI due to differences in coverage, definitions, and timing. Thus, the current account balance in 1994 was \$ -151 billion while the NIPA estimate of NFI was \$ -140 billion. (What's \$11 billion among friendly government statisticians?)

If payments to the rest of world exceed receipts (the current account is in deficit), then foreigners accumulate assets in the U.S. The change in the asset position of the U.S. is summarized by the capital account in Table 5. On the capital account a payment for a foreign asset purchased is a negative entry and accumulations of U.S. assets by foreign residents are positive entries. Since the 1994 current account balance was negative, the capital account balance (net purchases of U.S. assets by foreign residents) should be positive and of the same amount. We see that there is a large discrepancy between the current and capital account balances. This reflects the impossibility of the task at hand. There is simply no way of accounting for all cross-border transactions.

BUSINESS CYCLES

Economists have always followed the periodic changes that occur in level of business activity. These fluctuations are called business cycles. In this section, we will briefly define the term business cycle and describe some of the data that are used to monitor business cycle developments.

Business Cycle Defined

In a free enterprise economy, plans and decisions are made independently by a large number of economic agents. From time to time, imbalances between supply and demand will emerge and agents will not always be able to make the necessary adjustments to remove the imbalance. The inability to foresee and plan for all contingencies leads, on occasion, to an accumulation of imbalances throughout the economy. Such aggregate fluctuations are called business cycles.

Business cycles are recurring changes in economic activity. They do not follow any fixed periodic pattern such as seasonal cycles. Furthermore, each cyclical episode can differ with respect to the duration, depth, and diffusion of the cycle.

Although each business cycle is unique, there are enough similarities to make some general statements about the performance of the economy over the course of a typical cycle. Let us enter the process with an economic expansion getting underway.

The process of expansion can be fueled by a number of forces including an anti-recessionary macroeconomic policy, foreign demand, underlying growth expectations, and the inherent forces that bring a contraction to an end. The latter can include the influence of low mortgage interest rates on housing demand. Furthermore, low interest rates and the ready availability of skilled labor, plant capacity, material inputs and credit may lead entrepreneurs with underlying confidence in the economy to seize the opportunity and start new projects. Such responses can go a long way to starting an expansion phase.

The expansion impulse will spread quickly through the economy. Increases in earnings in the expanding sectors will generally lead to increased retail sales throughout the economy. New orders will expand in all sectors and bring the recession to an end. In the early phase of the expansion, output can be increased with only small increases in employment. Thus, productivity growth (the growth of output per labor hour) is likely to be rapid. As a consequence, profits are likely to respond quickly.

As the expansion goes on, capacity constraints loom in the not-too-distant future and delivery lags begin to lengthen. At the same time, interest costs and equipment prices are favorable. Thus, contracts and orders for investment goods begin to rise, and, with some lags, investment expenditures increase as well.

At this point, the expansion is well under way. GDP quickly surpasses its previous peak and a mood of optimism spreads over the economy. There is a willingness to undertake new activities and the expansion is self-reinforcing. Although there may be pauses in growth due to brief inventory adjustments, the strength of consumer demand and investor confidence can maintain an expansion for a long period of time.

Nevertheless, forces that can generate a recession do appear, and if a few of them come together, the expansion can reach its peak. First, as the expansion eliminates all slack in the economy, shortages of key resources might create physical barriers to further expansion. Second, as capacity is reached, costs of production will go up, profit margins decline and productivity growth slows. Growth expectations and the mood of optimism may begin to erode. Third, unless the monetary authority accommodates all inflationary pressures, the expansion is likely to lead to increased interest rates. Home building is often the first sector affected by interest rate pressures. Fourth, a vigorous expansion may lead to a too rapid buildup of inventories and capital goods.

At this stage, the economy is perched precariously between a slowdown in the expansion and a recession. If the balance of contractionary forces grows the economy will enter a recession.

After the peak of economic activity, many firms will find their inventories expanding rapidly as demand falls. There are pressures on profits and many firms might be experiencing financial difficulties. Cash flow might be insufficient to service debt incurred during the expansion and the number of bankruptcies is likely to increase. Once the contraction is clearly underway, unemployment will increase. The forces of recession will also be self-reinforcing.

However, there are a number of forces that make recessions rather short phases. First, consumption and investment plans are to a large extent determined by long-run expectations. Second, competitive pressures lead firms to take advantage of the recession to improve efficiency and increase market share. Third, the depreciation of capital leads to new investment demand. Finally, policy-makers are likely to take action to shorten a recession and this drives expectations. Thus, most recessions come to an end within a year and before the toll of unemployment has a major impact on the well being of society.

Cycle Dating

Business cycles are commonly separated into expansion and contraction or recession phases. Virtually everyone relies upon the dating or classification of cyclical episodes prepared by the National Bureau of Economic Research. The NBER is a private think tank whose Cycle Dating Committee is viewed as the

official arbiter of what constitutes a recession and when it began. In December 1990, the recession became almost official when the committee stated that “it appears likely” that the committee will decide that a recession began somewhere between June and September. In April 1991, before the recession had ended, the committee announced that the official start of the recession was July 1990. The start of a recession is called a business cycle peak since it is the peak of the ending expansion, and the end of a recession is called the cycle trough. The expansion began with many fits and starts and the committee hesitated for some time before reaching the conclusion that the recession had ended. It was not until December 1992 that the committee concluded that the cycle trough had been in March 1991.

Table 6 shows the dating and some important characteristics of post-war business cycles. There have been 9 recessions in the post-war period, ranging in length from 6 to 16 months with a median length of 10 months. The average decline in real Gross Domestic Product (GDP) was 1.9%, and the average high for the unemployment rate was 7.7%. Post war expansions have been typically longer, ranging from 12 to 106 months. The median length of the post war expansions is 42 months (the mean is 52 months).

TABLE 6**Post War Recessions**

<u>Peak</u>	<u>Trough</u>	<u>Length</u>	<u>Change in</u> ¹	
			<u>Real GDP</u>	<u>Non-Farm Employment</u>
Nov. 48	Oct. 49	11	-1.1	-5.2
July 53	May 54	10	-2.2	-3.5
Aug. 57	April 58	8	-3.3	-4.3
April 60	Feb. 61	10	-0.8	-2.3
Dec. 69	Nov. 70	11	-0.4	-1.2
Nov. 73	March 75	16	-4.1	-2.2
Jan. 80	July 80	6	-2.5	-2.2
July 81	Nov. 82	16	-2.9	-3.0
July 90	March 91	8	-1.5	-1.7

Post War Expansions**Growth in First Three Years²**

<u>Trough</u>	<u>Length</u>	<u>Real GDP</u>	<u>Non-Farm Employment</u>
Oct.49	45	8.4%	4.9%
May 54	39	3.2	2.8
April 58	24	4.5*	3.7*
Feb. 61	106	5.3	2.6
Nov. 70	36	4.6	3.5
March 75	58	4.4	3.7
July 80	12	3.5*	1.9*
Nov. 82	92	4.8	3.5
March 91	?	2.8	1.2

¹ Percentage change from peak to trough.² Growth from reference trough for first three years of expansion at annual rate.

* To end of expansion which was less than three years.

Cycle Indicators

Business cycle developments are followed by looking at large numbers of economic indicators that move over the course of a cycle. The study of cyclical indicators was introduced in the 1920s by economists at the National Bureau of Economic Research and elsewhere. In the post-war period, the Commerce Department analyzed and published data on the cyclical behavior of hundreds of measures of all types of economic activity—employment, production, orders, investment, inventories, sales, prices, costs, money stock, interest rates, credit difficulties. At the end of 1995, a few of the government's statistical efforts were privatized. Now, the Conference Board is publishing regular reports on cyclical indicators and compiles the widely followed indexes of leading, lagging and coincident indexes that had been developed by the government.

Cycle indicators are classified by their relationships to the reference cycles. That is, a measure of economic activity is termed a leading indicator if it has systematically and consistently turned down before the peaks in the reference cycles and turned up before the troughs. Similarly, lagging indicators lag the reference cycle turns and coincident indicators (usually measures of aggregate activity) follow the overall cycle.

Of particular interest are the leading cyclical indicators. Eleven such series are included in an index of leading indicators which is a widely followed monthly measure of business cycle activity. The components of the index are leading series usually because they reflect some intentions or plans for future economic activity, are an initial response to changing economic conditions or are a cause of future activity.

The components of the leading indicator index are:

- Average weekly hours, manufacturing
- Average weekly initial claims for unemployment insurance
- Manufacturer's new orders in 1987 dollars, consumer goods and materials industries
- Vendor performance, slower deliveries diffusion index
- Contracts and orders for plant and equipment, 1987 dollars
- New private housing units authorized by local building permits
- Change in manufacturer's unfilled orders in 1987 dollars, durable goods industries
- Change in sensitive materials prices
- Stock prices, S&P 500
- Money supply, M2 in 1987 dollars
- University of Michigan Index of Consumer Expectations

It is easy to see why these series are likely to lead. For example, employers initial responses to changes in demand will be to change the hours worked. Thus, the average workweek will rise or fall before the number of jobs changes. New orders, building permits, contracts and unfilled orders are clearly harbingers of future production. The real money supply is a measure of the monetary policy stance. The stock market might measure investor confidence and the index of consumer expectations is a survey measure of consumer sentiment and psychology.

The index of coincident indicators consists of aggregate measures of overall economic activity:

- Employees on non-agricultural payrolls
- Personal income less transfer payments in constant dollars
- Industrial production index
- Manufacturing and trade sales in 1987 dollars

Finally, the components of the index of lagging indicators are:

- Average duration of unemployment in weeks
- Ratio of manufacturing and trade inventories to sales (in 1987 dollars)
- Percentage change in index of labor cost per unit of output
- Average prime interest rate charged by banks
- Commercial and industrial loans outstanding in 1987 dollars
- Ratio, consumer installment credit outstanding to personal income
- Percentage change in Consumer Price Index for services

INTEREST RATES AND EXCHANGE RATES

Interest Rates and the Yield Curve

Since interest rates and exchange rates play an important role in all discussions of macroeconomics, it will be helpful to briefly present some important definitions.

To begin, consider a bond that pays the holder \$100 at maturity in N years and also pays the holder annual coupon payments of $100i$ where i is the coupon interest rate. The relationship between the yield to maturity, r , (which we will call the interest rate on an N year bond) and the price of the bond is:

$$B = \sum_{d=1}^N \frac{100i}{(1+r)^d} + \frac{100}{(1+r)^N}$$

where

B	=	price of bond
i	=	coupon interest rate
r	=	yield to maturity
N	=	time to maturity

The yield to maturity is the interest rate that discounts the bonds payments to its price.

If there are no coupon interest payments—a zero coupon bond (strip or a Treasury bill)—then the relationship is simply:

$$100 = B (1+r)^N$$

This illustrates the inverse relationship between the bond price— B —and the yield to maturity (market interest rate)— r . For example, consider a ten-year zero coupon bond (the promise to pay \$100 in ten years time). As the interest rate increases, the value (price) of this promise goes down:

$r = .03$	$B = 100/(1.03)^{10} = 74.4$
$r = .05$	$B = 61.4$
$r = .07$	$B = 50.8$

Now consider a two year bond ($N=2$), with a 6% coupon rate that has just been issued at par ($B=100$). The yield to maturity is also 6% because:

$$100 = \frac{6}{1.06} + \frac{6}{(1.06)^2} + \frac{100}{(1.06)^2}$$

If one year later the price of the bond increases to $B = 102$, what is the yield to maturity now? Remember that the coupon rate is still 6% and that now maturity occurs in just one year:

$$102 = \frac{(6+100)}{(1+r)} \text{ or } r = 0.039$$

The **yield curve** is the relationship between time to maturity (N) and yield to maturity (r) for bonds that have the same characteristics otherwise (e.g., risk of default). The most common yield curve is for U.S. government bond rates.

The relationships between yields at different maturities are often explained with the **expectations hypothesis**. We can illustrate the expectations approach to the term structure by considering again a bond with two years to maturity. An investor with a two-year horizon has two options:

- (i) buy a two year bond with yield to maturity r_2 , or
- (ii) buy a one-year bond with yield to maturity r_1 now and reinvest in another one-year bond when it matures. The investor does not know what the one year bond yield will be one year from now but the expected one-year yield is Er_1 .

If the investor is indifferent between the two options, then the expected yield will be equal. The expectations hypothesis is simply that the two options are equivalent:

$$(1 + r_2)^2 = (1 + r_1)(1 + Er_1)$$

If the two year rate, r_2 , is higher than the one year rate (an upward sloping yield curve), then the expected one-year rate one year from now is higher than the current one year rate, i.e., $Er_1 > r_1$. That is, an upward sloping yield curve implies that investors expect shorter-term rates to increase in the future.

The expectations hypothesis can be used to infer future short-term rates or expected changes in interest rates from the yield curve. For example, if I observe that the two-year rate is 4½% and that the one-year rate is 3¾%, then the expected one year rate is 5¼%.

The simple expectations hypothesis presented here ignores several important factors that generally make shorter-term bonds more attractive to investors. As a result, shorter-term yields tend to be lower and the yield curve is on average positively sloped:

- Shorter-term bonds are less subject to price variation when market rates change.
- There is more trading activity in shorter-term bonds, so there is more liquidity.
- Many investors have explicit maturity preferences (preferred habitats).

Inflation alters the purchasing value of financial assets. Therefore, it is often important to think in terms of the increase in purchasing power when holding dollar denominated financial assets—**the real rate of interest**. An example illustrates the relationship between the nominal and real rates.

Consider a one-year bond that is issued at 100 with a coupon interest rate of 5%. At the end of one year the bondholder receives \$105. However, the purchasing power of the amount returned is different than its purchasing power at the time the investment was made if there has been any price change during the year. Say that the inflation rate has been 3%. In this case, the purchasing power (in terms of dollars values at the start of the investment period) of the amount returned at the end of the year is:

$$\frac{105}{1.03} = 101.94$$

In real terms, the return to the bondholder is $1.94 = 2\%$.

Generally, the above calculation can be written as:

$$\frac{(1+r)}{(1+p)} = (1+r^{real})$$

where r is the nominal rate of interest, p is the inflation rate and r^{real} is the real rate of interest.

The real rate of interest can be related to the nominal rate with the **Fisher relationship** (named after a famous pre-war American economist):

$$\text{Nominal rate} = \text{Real rate} + \text{Inflation rate}$$

In terms of our example, the nominal bond rate of 5% is composed of a real return of 2% plus an inflation premium of 3%. The Fisher relationship is an approximation because the cross product term between the real rate and the inflation is ignored. (Unless inflation or interest rates are very high, it is quite small).

Exchange Rates

The exchange rate is defined as the number of units of foreign currency per dollar. A glance at the newspaper indicates various dollar exchange rates (as of May 22, 1996):

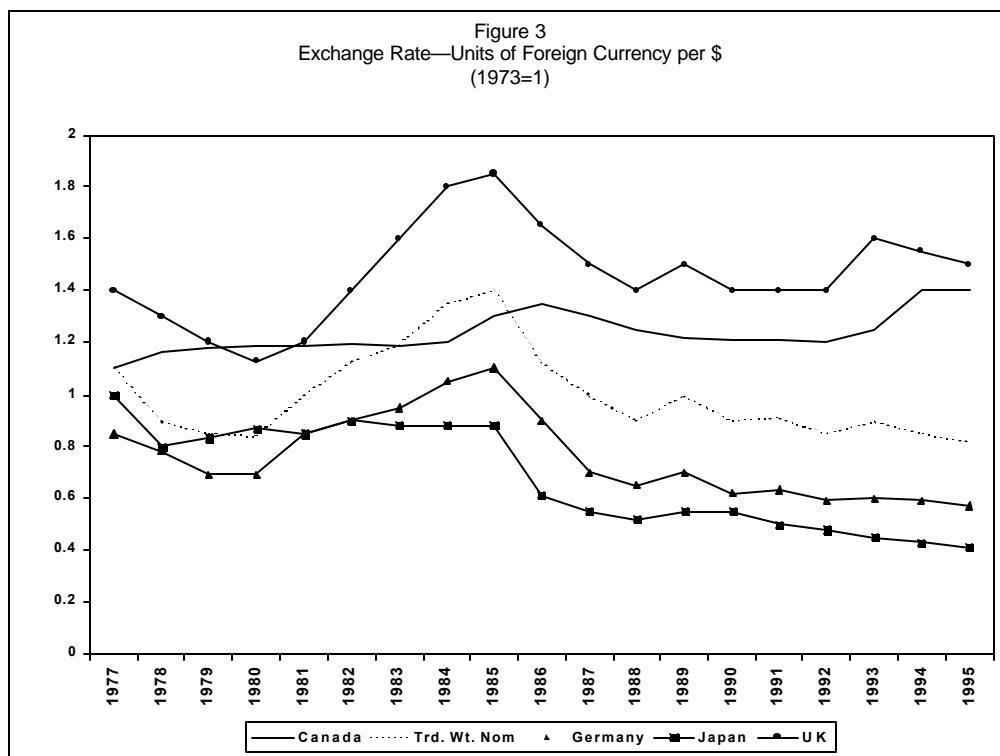
Japanese Yen	=	107.00
German Mark	=	1.54
Mexican Peso	=	7.38
French Franc	=	5.21
Canadian Dollar	=	1.37
British Pound	=	\$ 1.52 per pound (an exception from the usual quotation)

TABLE 7

Dollar Exchange Rates and Percentage Changes

	Exchange Rate					Percentage Change			
	<u>1975</u>	<u>1980</u>	<u>1985</u>	<u>1990</u>	<u>1994</u>	<u>1975-80</u>	<u>1980-85</u>	<u>1985-90</u>	<u>1990-94</u>
Yen	297	227	238	149	98	-23.6	+4.9	-37.4	-34.2
Mark	2.46	1.82	2.94	1.64	1.57	-26.0	+61.5	-44.2	-4.3
Can \$	1.02	1.17	1.37	1.16	1.38	+14.7	+17.1	-15.3	19.0
French Franc	4.29	4.23	8.98	5.49	5.37	-1.4	+112.3	-38.9	-2.2

The interesting thing about exchange rates is that they change a great deal and that they do not change together. Table 7 illustrates the magnitude and variability of exchange rate changes over the last 20 years. The exchange rates for the dollar against several major currencies are shown in Figure 3. Since there is a dollar exchange rate for every currency, it is useful to have a single measure of the dollar exchange rate. It is common to use an index of exchange rates for the dollar where every individual exchange rate is weighted by the proportion of U.S. trade that takes place with that country. For example, if 20% of U.S. trade is with Canada and 10% with Japan, then the trade weighted exchange rate Index assigns corresponding weights to the Canadian dollar and Yen exchange rates. The trade weighted exchange rate for the dollar is also shown in Figure 3.



If the exchange rate increases, a dollar purchases more units of foreign currency and the dollar has appreciated in value.

APPRECIATION $e \uparrow$

DEPRECIATION $e \downarrow$

Figure 3 shows that the dollar appreciated from 1980 to 1985, by over 40% using the trade weighted exchange rate. It depreciated by almost as much between 1985 and 1989. It has been fairly stable since that time although there have been some large changes in specific bilateral exchange rates in the 1990s.

An increase in the nominal exchange rate—appreciation of the dollar—does not necessarily imply that the dollar will buy more foreign goods. Consider a situation where the dollar appreciates from 4 to 5 Mexican pesos. The American visitor can buy 25% more pesos with each dollar; but his ability to buy Mexican goods depends on what has happened to the Mexican price level. If inflation in Mexico was just

25%, there has been no real change in what the dollar can buy in Mexico. For this reason, we need to introduce the concept of the **real exchange rate**.

The real exchange rate is defined as: $e_{real} = e_{nominal} \frac{P}{P_{foreign}}$

where P is the domestic and $P_{foreign}$ is the foreign price level.

An example will illustrate the definition.

Say that a price of an auto is \$2000 in the U.S. ($P=2000$) and 3000 Swiss Francs in Switzerland ($P_{foreign}=3000\text{SFr}$) in Switzerland. Also, the nominal exchange rate (SFr per dollar) is $e_{nominal} = 1.5$ which means that the auto costs the same in both countries. The real exchange rate is $e_{real} = 1.0$ because prices are the 'same' in the U.S. and Switzerland.

If e increases (the dollar appreciates) to 3.0 AND, ALSO, auto prices increase so $P = 2000$ and $P_{foreign} = 6000$, then $e_{real} = 1$. The real exchange rate is unchanged because prices and the exchange rate have all changed by the same proportion.

However, if $e = 3$ and the ratio of prices remains the same then e_{real} goes up, there is a real appreciation of the dollar.

If we take the rate of change of the definition of the real exchange rate, we have:

$$\% \Delta e_{real} = \% \Delta e_{nominal} + \% \Delta P - \% \Delta P_{foreign}$$

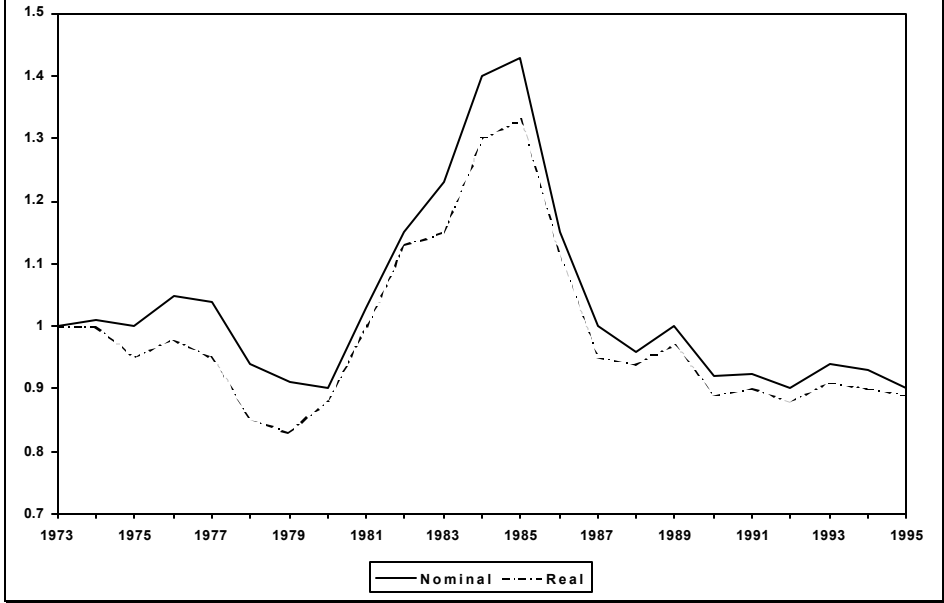
The real exchange rate changes when nominal exchange rate changes do not offset the differential between the domestic and foreign inflation rates.

Consider some examples.

- If the domestic inflation rate exceeds the foreign inflation rate AND the nominal exchange rate is unchanged \rightarrow the real exchange rate appreciates.
- If the domestic inflation rate exceeds the foreign inflation rate AND the real exchange rate is unchanged \rightarrow the nominal exchange rate depreciates.

Figure 4 shows the trade-weighted nominal and real exchange rate for the dollar since 1973. It is interesting to note that all of the major fluctuations in the nominal exchange rate are mirrored by changes in the real exchange rate. That is, aggregate exchange rates movements are not closely related to inflation differentials. The story would be somewhat different for bilateral exchange rates. If we examine, for example, the exchange rate for the Mexican peso, the nominal exchange rate changes in order to offset the inflation differential and the real exchange rate is less volatile than the nominal exchange rate.

Figure 4
Trade-Weighted Exchange Rates
(1973=1)



CHAPTER II

GROWTH, PRODUCTIVITY AND LONG-RUN EQUILIBRIUM

An examination of the wealth of nations reveals two facts—one well known and one little known. The well-known fact is that there are enormous differences in the value of output per capita as we travel around the world. In 1993, GDP per capita was almost \$25,000 in the United States, which is about fourteen times larger than GDP per capita in China. Some countries are very wealthy and some are very poor. The less well-known fact is that the relative position of different countries can change dramatically over relatively short periods of time. In 1965, real GDP per capita in the U.S. was 17 times larger than that in Korea; twenty-five years later the ratio was less than 5.

The comparison of incomes in different countries is sometimes a complicated matter. When GNP per capita is converted to dollars using the exchange rates, we find that in 1993 GNP per capita in Japan (\$31,490) and Switzerland (\$35,760) are much higher than the level in the U.S. (\$24,740). The comparison is misleading because prices in Japan and Switzerland are much higher than comparable prices in the U.S. The goods and services produced in different countries should be valued consistently so that the differences in GNP reflect real differences in the volume of production. A better comparison is made when GNP is converted to dollars with purchasing power parities instead of exchange rates. PPP conversions show how many units of currency are needed in one country to buy the same amount of goods and services which one currency unit will buy in the other country.

Even when PPPs are used to make GNP in different countries comparable in terms of purchasing power, there are enormous differences in per capita real output around the world:

GNP Per capita in 1993 on a PPP Basis

United States	\$24,740
Japan	20,850
Singapore	19,510
France	19,000
Germany	16,850
Spain	13,510
Korea	9,630
Hungary	6,050
Brazil	5,370
Russia	5,050
Bulgaria	4,100
Egypt	3,780
China	2,330
India	1,220

Source: World Bank, *World Development Report*, 1995.

PPP calculations for long periods of time are not readily available, so to see how relative income positions have changed over time, we will look at the growth of real GDP per capita from 1965 to 1989.

	<u>GDP Per capita, \$</u>		
	<u>1965</u>	<u>1989</u>	<u>Average annual growth rate</u>
China	88	350	5.7%
Philippines	477	710	1.6%
Korea	811	4400	7.0%
Brazil	1075	2540	3.5%
Uruguay	1944	2620	1.2%
Singapore	1925	10450	7.0%
New Zealand	9890	12070	0.8%
United States	14060	20910	1.6%

In 1965, Korea and the Philippines were undeveloped economies with real GDP less than 10 per cent of that in the U.S. Sustained growth in Korea has led to a situation where real per capita GDP in 1989 was over 20 per cent of the American level while slow growth in the Philippines means that the ratio is less than 5 per cent of the American level. Similarly, look at Brazil and Korea; real GDP was almost over one-fourth larger in Brazil in 1965. By 1989, because of a growth rate in Korea that was twice as large, real GDP in Korea was bigger than that in Brazil by three-quarters. Another interesting comparison is between Uruguay and Singapore, two small countries with the same levels of real per capita GDP in 1965. By 1989, Singapore's real per capita output was almost four times larger than Uruguay. Rapid growth can occur in economies (e.g., Korea) which start with little physical infrastructure and the ability to improve quality of human capital through education. In addition, technology transfers and foreign investment will lead to convergence among developed economies (e.g., Germany and the U.S.). Nevertheless, growth rates matter; they determine the relative wealth of nations. Differences in growth rates make significant differences in just one generation.

Why do countries grow? This is a simple question with a complex and controversial answer. First and foremost, growth is due to the increased availability of the resources or the inputs into the production process—capital and labor. Second, growth is due to improvements in the technology of production which will take place from both the transfer of technology from country to country and the development of new technology and the fruits of invention. Finally, there are a number of other issues that might effect the growth process, such as:

- Macroeconomic policy—monetary and fiscal policy, as well as trade and industrial policies, may or may not create a growth oriented environment; this includes the openness of the economy to trade and competition that helps it utilize its resources to maximum advantage.
- Political stability and a legal structure that securely defines and protects property rights are essential to the development of a growth-oriented environment.
- Social and cultural attitudes can also effect the ability of a given economy to exploit the available resource base.

The controversy emerges when we examine countries that have experienced unusual growth spurts, such as Singapore and Korea in the above table—two Asian tigers. Is their rapid growth simply due to an increase in available capital and labor resources? These countries have high savings rates and started with a very small capital base in the 1960's. Thus, the growth rate of the capital stock was very rapid. Similarly, the labor force grew rapidly from both population growth and the increased participation of the population in the active labor force. Nevertheless, the sustained rapid growth in these countries suggests that something else was going on as well. Perhaps, it was the government direction of resources to key industries. However, few observers would attribute a generation of rapid growth to the success of

government planning processes. If government planners could generate rapid growth, then miserable economic growth record would not have led to the disappearance of communist regimes throughout Europe and Asia. Instead, perhaps there are growth synergies that enable expanding economies to learn from their growth experience and rapidly increase their ability to use resources. This endogenous growth phenomenon—that growth feeds on itself and teaches the economy to grow some more—is difficult to understand but surely plays a role in the sustained growth spurts which we see in some countries.

ANALYZING GROWTH

The analysis of the sources of economic growth starts with the production function:

$$Y = A * F(K, N)$$

The level of output (really the potential or the natural rate of output, since it is assumed that all resources are in fact utilized), Y , is a function of capital, K , and labor N , inputs. Furthermore, the whole production frontier increases or shifts out as the level of technology, A , increases. Growth in A reflects the ability to produce more with a given level of inputs because of improvements in production technique and knowledge, improved management, the diffusion of technological progress and the synergies that create growth rates (see above). In addition, A may shift in the short-run to reflect shocks to the production process from, for example, the weather (droughts, storms), the impact of government regulation, supply shocks (e.g., an oil embargo) that inhibit production.

The production function can be rewritten to show how the growth rate of output depends on the growth rates of the inputs. To show this, take the derivative of the production function with respect to time:

$$\frac{dY}{dt} = A \frac{dF(K, N)}{dt} + F(K, N) \frac{dA}{dt}$$

$$\frac{dY}{dt} = A \frac{\partial F}{\partial K} \frac{dK}{dt} + A \frac{\partial F}{\partial N} \frac{dN}{dt} + F(K, N) \frac{dA}{dt}$$

Now, define the elasticities of Y with respect to K and N respectively as:

$$\mathbf{a}_K = \frac{dY}{dK} \frac{K}{Y} = A \frac{\partial F}{\partial K} \frac{K}{Y}$$

$$\mathbf{a}_N = \frac{dY}{dN} \frac{N}{Y} = A \frac{\partial F}{\partial N} \frac{N}{Y}$$

Divide each term on both sides of the last equation for dY/dt , by Y , yielding $(dY/dt) / Y$, the growth rate in output on the left and on the right side substitute \mathbf{a}_K and \mathbf{a}_N as defined above:

$$\frac{dY}{Ydt} = \mathbf{a}_K \frac{dK}{dt} \frac{1}{K} + \mathbf{a}_N \frac{dN}{dt} \frac{1}{N} + \frac{1}{A} \frac{dA}{dt}$$

$$\frac{\Delta Y}{Y} = \mathbf{a}_K \frac{\Delta K}{K} + \mathbf{a}_N \frac{\Delta N}{N} + \frac{\Delta A}{A}$$

The above equation for the decomposition of output growth is the framework for analyzing the sources of growth. It is used to determine to what extent growth can be attributed to growth in the capital stock, $(\frac{\Delta K}{K})$, and growth in labor inputs, $(\frac{\Delta N}{N})$.

The residual, $\Delta A / A$, or that part of output growth which is not explained by the growth in the factors of production, is called technological progress or total factor productivity growth. It is the extent to which both the capital and labor inputs are becoming more useful in production.

$$\text{Total Factor Productivity Growth} = \frac{\Delta A}{A}$$

A yet more important concept is the rate of growth of labor productivity or the percentage change in output per unit of labor (which is the rate of growth of output less the rate of growth of labor inputs):

$$\text{Labor Productivity Growth} = \frac{\Delta Y}{Y} - \frac{\Delta N}{N}$$

Note that if the elasticities sum to one, we can rewrite the growth equation and the definition of labor productivity growth becomes:

$$\text{Assume } \mathbf{a}_K + \mathbf{a}_N = 1$$

$$\left(\frac{\Delta Y}{Y} - \frac{\Delta N}{N}\right) = \mathbf{a}_K \left(\frac{\Delta K}{K} - \frac{\Delta N}{N}\right) + \frac{\Delta A}{A}$$

This equation shows that labor productivity growth can be attributed to two causes:

- technological progress $\frac{\Delta A}{A}$
- capital deepening $\frac{\Delta K}{K} - \frac{\Delta N}{N}$

Capital deepening is the rate of growth of capital per unit of labor input (rate of growth of K/N). When the capital-labor ratio is increasing, each worker can increase his or her production because the amount of capital available is growing.

There is an additional interpretation of the weights or elasticities in the equation for the decomposition of output growth. Recall that in a competitive economy the real wage, w/p , is equal to the marginal product of labor and that the return on capital, c , is the marginal product of capital. If capital and labor are paid their marginal products, then the total return to the owners of capital is cK and the total return to labor is $(w/p)N$. Finally, by substituting for the marginal products in the definitions of the elasticities, it is easy to see that the elasticities are the total factor returns as shares of output:

$$\mathbf{a}_K = \frac{cK}{Y} \quad \mathbf{a}_N = \frac{(\frac{w}{p})N}{Y}$$

Specifically, \mathbf{a}_N is wage and salary payments as a proportion of total income and \mathbf{a}_K is the returns to capital (profits, rents, etc.) as a proportion of total income.

Output growth is a weighted average of growth in the capital and labor inputs where the weights are the proportions of national income which are paid to the respective factors of production plus the rate of technological progress.

PRODUCTIVITY

Labor productivity growth, the growth of real output per unit of labor input (e.g., hours worked) is the primary source of growth in real wages. To see this, look at the labor share definition of a_N in rate of growth terms and rearrange:

$$\frac{\Delta \frac{w}{p}}{\frac{w}{p}} = \frac{\Delta a_N}{a_N} + \left(\frac{\Delta Y}{Y} - \frac{\Delta N}{N} \right) \quad \frac{\Delta \frac{w}{p}}{\frac{w}{p}} = \frac{\Delta w}{w} - \frac{\Delta p}{p}$$

Real wages can grow only when there is labor productivity growth or labor's share of product increases. Although labor's share does vary cyclically, it has remained quite constant for long periods of time (except for changes due to increases in the size of the government sector). Thus, productivity growth is the source of wage growth. The real wage can only increase when technological improvements or increases in the capital stock augment labor productivity. An important point can be brought out by remembering that real wage growth is nominal wage growth less inflation. Thus, if nominal wages increase more rapidly than productivity (with labor's share constant), the result will be inflation rather than real wage growth.

Productivity Trends

A major problem of the post-war era is that the average, or trend, rate of labor productivity growth has declined. From 1950 to 1969, labor productivity growth in the U.S. business sector averaged 3.2 per cent per year. Since that time, that figure has only been exceeded in three years. The productivity growth slowdown started in the early 1970s and many economists originally thought that it was principally a temporary phenomenon associated with the supply shocks of the 1970s—the 1973 and 1979 oil shocks, the inflationary shocks from the agricultural sector, deep recessions in 1981–82 and 1990–91. However, the following data for the average rate of growth of labor force productivity for the non-farm business sector shows that the slowdown has persisted for over two decades:

1960–73	2.85%
1974–81	0.94
1982–90	1.11
1991–95	1.07

The exceptional experience for productivity growth is apparently the first 25 years after the Second World War when productivity in the U.S. grew at a rate that was more than twice the average rate for the next quarter century. A number of things contributed to this slowdown such as the increased share of service industries in the economy (which have notoriously low and often mismeasured rates of productivity growth) and the maturing of the manufacturing sector which reduced the contributions of new inventors and discoveries.

The productivity growth slowdown is not restricted to the U.S. Table 1 shows OECD (Organization for Economic Cooperation and Development) estimates of trend rates of labor productivity growth (with estimates of the cyclical influence removed). Although the slowdown is most pronounced in the U.S., it occurred in all the major economies.

TABLE 1Trend Growth in Labor Productivity (Business Sector) in Percent per Year

	Early '60s–1973	1974–79	1980–86	1987–92
U.S.	2.0	0.0	0.5	0.9
Japan	8.3	3.4	2.8	2.9
Germany	4.5	3.1	1.5	1.9
U.K.	3.3	2.0	2.6	1.8

Source: OECD US Economic Survey, 1993

The decline in the productivity growth rate has serious implications. It means that real-wage growth is necessarily less than it was a generation ago. Individuals who have entered the work force in the post-Vietnam era experience smaller real-wage gains than did their parents in the post-World War II era. Moreover, their expectations are likely to be that their experiences should match that of their parents. That is, they expect their real wages to grow from year to year. Thus, their expectations of wage gains generate inflationary pressures and a failure to realize such expectations contributed to the economic malaise of recent times.

Sources of Economic Growth

There are empirical studies that decompose the growth of output into its various sources. Given the capital and labor shares, we can attribute the observed growth of real product into portions due to capital and labor inputs. Total factor productivity is the residual or that portion of output growth not explained by growth of the inputs. However, if we underestimate $\frac{\Delta K}{K}$ and $\frac{\Delta N}{N}$, then $\frac{\Delta A}{A}$ might be very large.

For example, the labor input can be a simple count of hours worked or, better yet, hours corrected for the level of education of individuals in the labor force (the quality of the labor input). Thus, measurement of total factor productivity growth is somewhat sensitive to the measures of K and N used. Nevertheless, most studies suggest that the contribution of total factor productivity to output growth is substantial.

Table 2 presents a breakdown of the sources of economic growth in the United States since World War II. Data are shown for each of the postwar decades—the 50's to 80's. The table starts with the average annual growth rate in non-farm business output and the growth of the net stock of fixed capital hours worked. Total factor productivity is derived as the residual from:

$$\frac{\Delta A}{A} = \frac{\Delta Y}{Y} - a_N \frac{\Delta N}{N} - (1 - a_N) \frac{\Delta K}{K}$$

To illustrate, the calculation for the first subperiod is:

$$1.9 = 4.0 - (0.67)1.3 - (1 - 0.67)3.7$$

Table 2Growth in the U.S. Economy

	1950–59	1960–69	1970–79	1980–89	1990–93
Output DY/Y	4.0	4.2	3.0	2.9	1.5
Capital DK/K	3.7	4.4	4.4	3.7	3.5
Labor hours DN/N	1.3	1.7	1.8	1.8	-0.0
Total Factor Productivity DA/A	1.9	1.6	0.5	0.6	0.6
• Due to government capital	0.4	0.4	0.2	0.0	0.0
• Due to education	0.4	0.1	0.0	0.4	0.4
Labor Productivity	2.7	2.5	1.2	1.1	1.5
Capital Deepening	2.4	2.7	2.6	1.9	3.5
Labor share a_N	.67	.67	.73	.74	.74

Average annual percentage changes for Private non-farm business sector in 1982 \$

Sources: OECD Economic Survey of the U.S., 1991–2, Economic Report of President, 1995, Citibase

Total factor productivity growth is that part of output growth which is not accounted for by growth in capital and labor inputs. It can also be viewed as the rate at which the production functions shift out or the rate of technological progress. It is quite large in the first two decades where it accounts for almost half of total growth. The residual includes the effect of improvements in the quality and allocation of inputs, as well as advances in knowledge. The table also includes some estimates of the contributions towards total factor productivity growth from two important sources. First, the productive contribution of government owned capital (e.g., roads, airports) that is not included in the private capital stock was substantial in the 1950s and 60's when there was significant infrastructure investments by the public sector. Second, the increasing educational level of the work force increases the productivity of each hour of labor input. Changes in both of these factors explain part, but by no means, all of the slowdown in total factor productivity growth.

The slowdown in the growth of labor productivity is of particular interest. It is of serious proportions. The slowdown started in the 1960s, even while the capital stock was growing rapidly. Since the mid-1970s productivity growth had declined even further. To a large extent this might be a result of the effects of the two large recessions in the period and the changes in energy prices. On the other hand, it is possible that there has been a decline in the rate of technological improvements. It should also be noted that the early post-war period was an atypical experience. The post-war and post-depression era allowed for unusually high levels of productivity growth for almost 20 years.

How to Increase Productivity Growth

We showed above that the two sources of growth in labor productivity are capital deepening and technological progress. The above data indicate that the rate of technological progress since the mid-1960s is only about half as large as it was in the first two post-war decades. Capital deepening has also declined. In the middle subperiod, capital deepening was less than it was earlier, because even though the capital stock grew rapidly, there was a large increase in the rate of growth of labor. In the last subperiod, the rate of growth of the capital stock slowed down considerably.

An interesting conjecture would be to calculate what labor productivity growth would have been in the 1980s if the capital stock had grown as rapidly as it had in earlier decades. Suppose that $\frac{\Delta K}{K} = 4.4$ percent; then the calculation of labor productivity growth would be:

$$0.6 + 0.26(4.4 - 1.8) = 1.3$$

Even with exceptionally rapid capital stock growth, labor productivity growth would have been modest.

The reason for this is that $\frac{\Delta K}{K}$ has a small weight, $a_K = 0.26$. Thus tax policy, which promotes investment, is likely to have only a small effect on productivity growth. Moreover, large changes in investment have only a small effect on the capital stock in the short-run. This can be seen from the identity that relates investment to the capital stock:

$$K_t = K_{t-1} + I_t - \text{Depreciation}$$

The value of the capital stock (K) in the U.S. is almost \$10,000 billion while gross investment (I) in 1993 was around \$900 billion. Thus, changes in the capital stock are always gradual.

The key to understanding the productivity slowdown is then the growth of total factor productivity, $\frac{\Delta A}{A}$.

We will look at some of the standard explanations found in the historical studies. However, it is also important to mention that episodes of rapid growth are often characterized by both capital deepening and rapid total factor productivity growth. This can occur when new capital equipment embodies new technology and the new technology has benefits to society that extend beyond the capital. This is an illustration of endogenous or self-expanding growth as new technology effects exiting production processes and growth takes off. The synergy between new and existing technology implies that we may all learn from new technology although it is only directly used in certain activities. Endogenous growth phenomena help explain the remarkable growth spurts of some Asian countries in particular that we noted at the start of our discussion.

Table 3 summarizes historical studies by John Kendrick and Edward Denison on the sources of growth for the whole economy. To a large extent, the decline in total factor productivity growth is due to structural and demographic changes in the economy. Less than a third of the post-war decline in $\Delta A/A$, according to Kendrick's calculations, is due to a slowdown in advances in knowledge.

TABLE 3Sources of Total Factor Productivity Growth

	1929–1948	1948–1966	1966–1973	1973–1978
$\Delta A/A$	2.3	2.8	1.6	0.8
Labor quality	0.8	0.7	0.8	0.9
Age-sex composition	0.0	-0.1	-0.4	-0.2
Resource reallocations	0.4	0.8	0.7	0.3
Volume of production	0.4	0.4	0.2	-0.1
Impact of government	0.1	0.0	-0.1	-0.3
Advances in knowledge	0.7	1.4	1.1	0.8
Other	-0.1	-0.4	-0.7	-0.6

To begin, the rate of growth of labor quality has contributed to total factor productivity for over 50 years. The sources of this growth category are the steady improvement in the level of education of the labor force and the health of workers (i.e., the decline in days lost owing to ill health). Second, changes in the age-sex composition of the labor force has hindered productivity growth in the post-war period, particularly after the mid-1960s. This occurred because there was rapid growth in the labor force in that period as the baby boom generation entered the labor force and female labor force participation increased. These new entrants were inexperienced, low paid, and generally less productive than older members of the labor force. Of course, this phenomenon is reversing itself as the growth of new entrants slows and the baby boomers gain work experience. The reversal may well boost productivity growth throughout the 1980s and 1990s.

The economy is composed of different types of economic activity with different levels of productivity. One source of productivity growth over time is the movement of economic resources from parts of the economy with relatively low productivity to parts where there is high productivity. This started at least 100 years ago with the movement of the labor force off the farm to the manufacturing sector. The increase of output per unit input was simply a consequence of the reallocation. We cannot expect this particular movement to be a continued source of productivity growth, because there is only a small part of the labor force left on the farm and labor productivity in agriculture is now very high.

Another kind of resource allocation is related to the growing importance of the service sector, wherein there has been less rapid productivity growth. This is a consequence of prosperity; we can afford to have a larger fraction of our labor force working in the service sector. However, such a situation dilutes the productivity growth taking place in the manufacturing sector. One problem with understanding trends in productivity growth is that service sector productivity is difficult to measure. It is probable that our current measures understate productivity in that sector. In the government and the financial services sectors (both of which have grown rapidly—government in the 1960s and 70s and financial services more recently), output is measured by the size of the labor input so productivity is often assumed away. Furthermore, it is possible that productivity growth in the service sector will rival that of manufacturing in

the future when the service sector begins to utilize new technologies. For example, the long-run productivity effects of the spread of PCs throughout industry has yet to be measured.

The efficiency of resource use varies over the business cycle. Inefficient use of resources emerges when there is a lot of excess capacity. Thus, the position in the business cycle affects the measured rate of growth of productivity. Particularly in the last subperiod, which includes a major recession, productivity growth was reduced by volume effects.

The impact of government on productivity growth is twofold. First, the government provides services to the business sector, which promotes productivity. Since the standard capital stock data does not include government capital (e.g., bridges, school buildings, etc.), the contribution of such capital to productivity growth is not measured. Second, government regulation (e.g., pollution controls) uses resources without increasing measured output (e.g., clean air and water is not measured as output) and thus reduces productivity growth. The net impact of government was virtually zero until the mid-1960s, when increased business regulation began to have a considerable negative effect. Thus, the costs of maintaining the quality of life (e.g., costs of environmental regulation) have a negative effect on measured productivity.

Finally, technological progress, or that part of $\frac{\Delta A}{A}$ not otherwise explained, is the residual effect. There has been substantial decline in progress since the early post-war period. This could be because there are simply fewer opportunities for technological progress in the contemporary economy. Alternatively, it could be a result of a reduction in the resources devoted to developing new technology. For example, the proportion of GDP devoted to research and development expenditures started to decline in the 1960s.

Thus, the decline in productivity growth in the post-war period can be largely attributed to resource allocations, the impact of government policies, and the cyclical problems of the past decade. None of these would be much affected by changes in capital formation policy.

There is good reason to support the belief that the productivity growth slump will end. This is likely to occur because some of the factors that have had a negative impact on productivity growth are going to be reversed. For example:

1. Research and development expenditures as a percentage of GDP have risen from the lows of the late 1970s.
2. Tax changes have reduced the taxation of business income and spurred capital expenditures.
3. The expansion increased capacity utilization, which generally leads to economies of scale and more efficient production.
4. The baby boom generation is now passing into its prime working years, where added experience will aid productivity growth.
5. Deregulation of industries such as telecommunications, banking, and air transport increases competitive incentives to innovate and increase productivity. Similarly, the burden of complying with government social regulation (anti-pollution, health, etc.) has begun to gradually lessen.

The relationship between investment and growth in potential output is often complex. This can be seen if we consider that a given investment expenditure may have one or more of several purposes: it may replace existing capital, upgrade the technology of existing capital, be a “non-productive” addition to capital (e.g., pollution abatement equipment), or augment the productive capital stock.

MODEL OF LONG-RUN EQUILIBRIUM

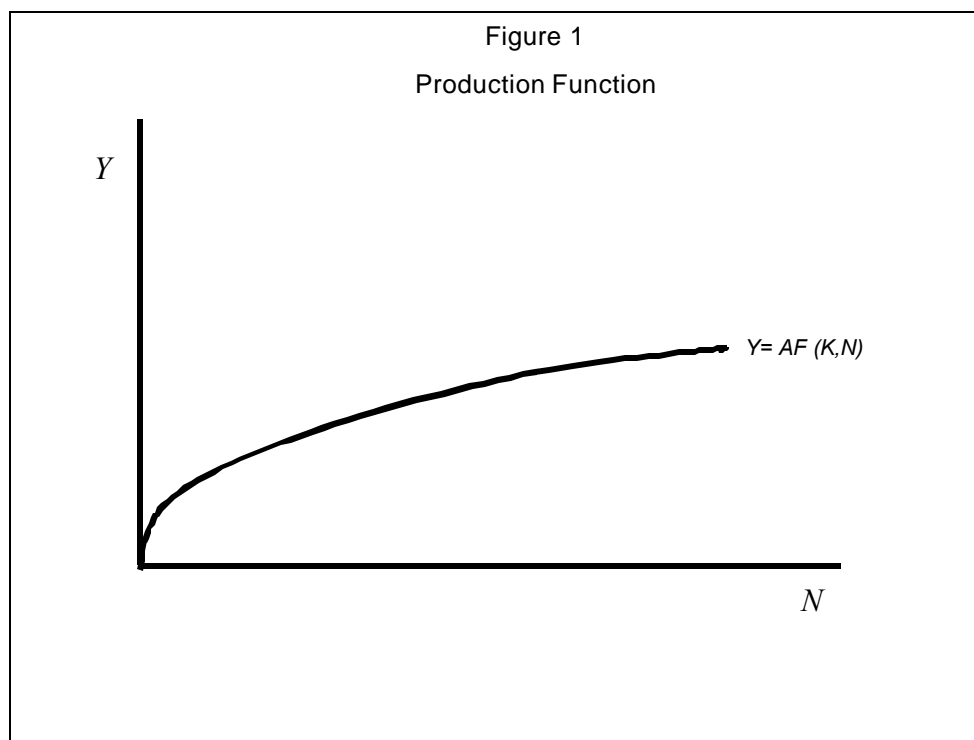
We now turn to the development of a model of long-run macroeconomic equilibrium. The model will have the following components:

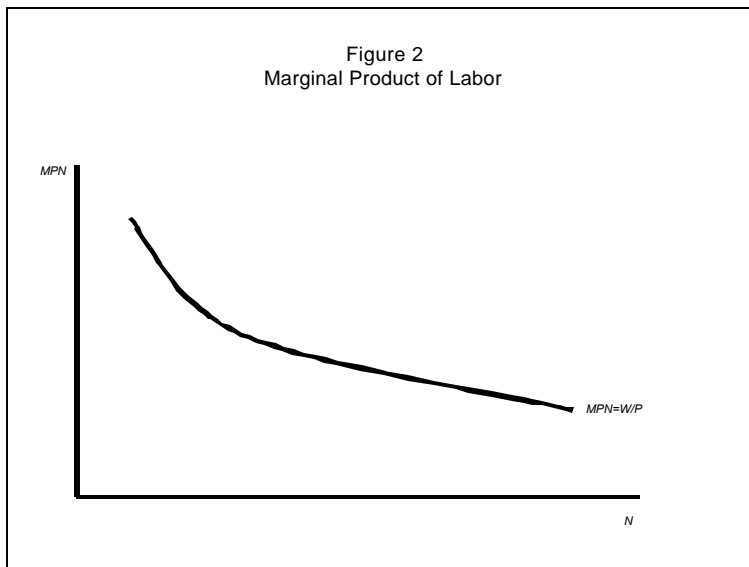
- The equilibrium level of output which is determined by the production function and the labor market. The supply and demand for labor determine the equilibrium real wage and the level of employment.
- A distribution model that shows how the equilibrium output is divided between consumption goods and investment goods. This part of the model explains the determination of the real interest rate and also the level of investment.
- A monetary sector that relates the money stock to the price level.
- An international or open economy sector that relates the equilibrium to exchange rates.

Output Equilibrium

The production function— $Y = AF(K, N)$ —can be used to derive the long-run equilibrium level of output. To begin, we will take the level of technology (A) and the capital stock (K) to be given. The equilibrium level of output then depends on the size of the labor input into production. The labor input is determined by a labor market equilibrium.

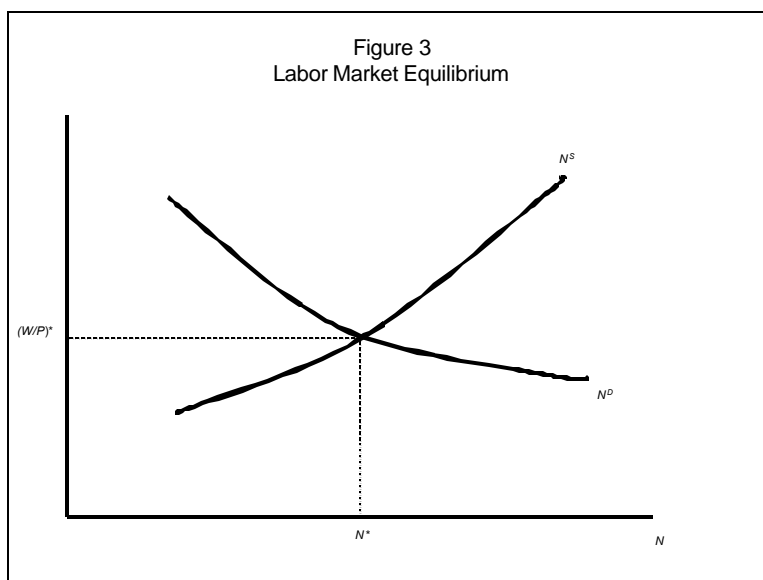
Figure 1 shows the production function for a given level of technology A and the capital stock, K . It shows that equilibrium or potential output increases with N and that there are diminishing returns to additional labor inputs. If the level of technology or the capital stock increases, then the production function in Figure 1 shifts upwards. This production function can be used to derive the marginal product of labor as shown in Figure 2 where MPN stands for the marginal physical product of labor (the increment to output from an additional unit of labor).





Profit maximization dictates that the firm should hire labor until the value of the incremental output ($P \cdot MPN$) equals the wage cost (w). The labor demand by firms for any real wage will be the amount that equates the real wage, $\frac{W}{P}$, to the marginal product of labor. Thus, for any given real wage, the MPN curve tells us the profit-maximizing amount of labor demanded. The labor demand curve will shift out if A increases which will increase MPN for all levels of N .

The labor demand curve is shown in Figure 3, along with a labor supply curve. The labor supply curve is drawn with a positive slope, indicating that an increase in real wages leads to a greater supply of labor hours. It is possible that individuals will do just the opposite, decrease their work hours as the real wage increases. A higher real wage rate means that an individual can attain a target income with fewer hours of work. Thus, an increase in the real wage may lead to an increase in leisure. The empirical evidence on labor supply indicates that there is a small positive effect of real wages on labor supply and thus we draw the supply function with a positive slope. Finally, the equilibrium level of labor, N^* , and the equilibrium real wage, $(\frac{W}{P})^*$, are shown in Figure 3.



Equilibrium in the labor market does not mean that there is no unemployment. Time spent in normal job search activities and movements in and out of the labor force (by school leavers, persons involved in housekeeping, etc.) lead to a **natural rate of unemployment** which is currently estimated at approximately a five percent unemployment rate in the U.S. and perhaps as much as 8 percent in Western Europe. This natural rate may well be unsatisfactorily high and we can consider why this is the case and what may be done to reduce it.

If the costs of job search are reduced by better job training or improved placement services, then the time spent searching would fall and the natural rate of unemployment would be lower. Differences in the incentive to stay unemployed can explain part, but by no means all, of the international differences in unemployment rates. In the U.S., unemployment benefits provide about one-third of average wage for one-half of a year. By contrast, in Holland, unemployment benefits provide about three-quarters of the average wage for three years, a substantial incentive to stay unemployed and to keep job searching at a very leisurely pace, which results in a higher natural rate.

The labor market equilibrium determines N^* and the level of technology and the capital stock are given. Thus, returning to the production function, we can determine the equilibrium level of output, Y^* .

SUMMARY

The initial elements of the long-run macro equilibrium are:

1. Labor market equilibrium
 - Demand for labor
 - Marginal product of labor = real wage
 - $MPN = W/P$
 - Supply of labor

⇒ Equilibrium Labor hours N^* and real wage $(W/P)^*$

2. Production Function
 - $Y = A F (K, N)$

⇒ Long-run Equilibrium or Potential or Natural Output – Y^*

Distribution of Output and Interest Rates

Recall that in Chapter I, we introduced the NIPA identity, which separated the level of output into its components:

$$Y = C + I + G + (X - M)$$

and also derived the investment-saving identity:

$$I + NFI = PS + BS + GS$$

Combine personal and business saving into private saving and reorganize the identity to:

$$Private\ savings = I - GS + NFI$$

From a NIPA perspective, this is an identity that must always hold. If anticipated flows of income and goods do not result in a balance, there will be unanticipated accumulation of investment goods (e.g. inventories) or unanticipated savings or dissavings. Alternatively, if the flows do not balance, there will be adjustments in the markets for investment and savings flows that lead to an adjustment process.

Our distribution model examines the adjustment process in the market for investment and savings—the demand for investment goods and the supply of savings—that keeps the savings and investment equal. We begin by looking at the determinants of savings and investment. Generally, savings and investment decisions are quite different.

Private savings depends on:

- (i) disposable income,
- (ii) future income and taxes,
- (iii) wealth, and
- (iv) real interest rates.

Investment is determined by:

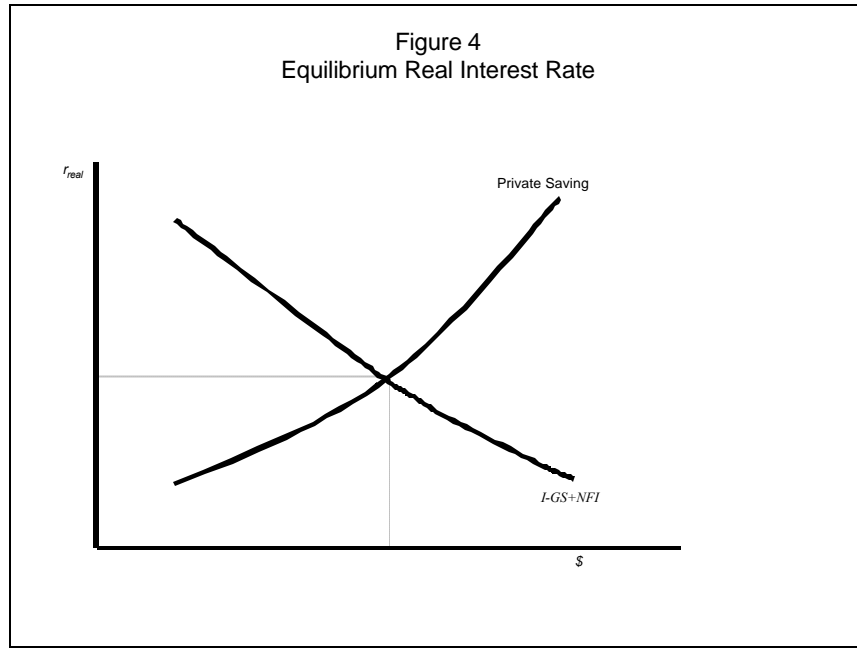
- (i) the productivity of the capital stock,
- (ii) current and future income, and
- (iii) the real rate of interest.

The market mechanism that keeps anticipated savings and investment in balance is the adjustment of the real rate of interest.

The flow of private saving increases with real interest rates; thus the supply of savings is a increasing function of real interest rates. That is, as the real return to saving increases, individuals and businesses will defer consumption and save more of their current income.¹

The demand for investment goods decreases with real interest rates, which represent the cost of financing investment expenditures or, alternatively, the opportunity cost of holding physical capital. Thus, the demand for investment is a decreasing function of real interest rates. Figure 4, shows the relationships between the real interest rate and both the demand for finance (investment) and the supply of finance or savings. In each case the function is drawn for a given level of the output equilibrium, Y^* . In addition, we are taking the level of the government saving and net foreign investment as given in this discussion. The other determinants of investment and saving are held constant as well. The flow of private savings and the demand for investment financing through the financial markets determines the level of real interest rates — r_{real} —and keep investment and saving in balance.

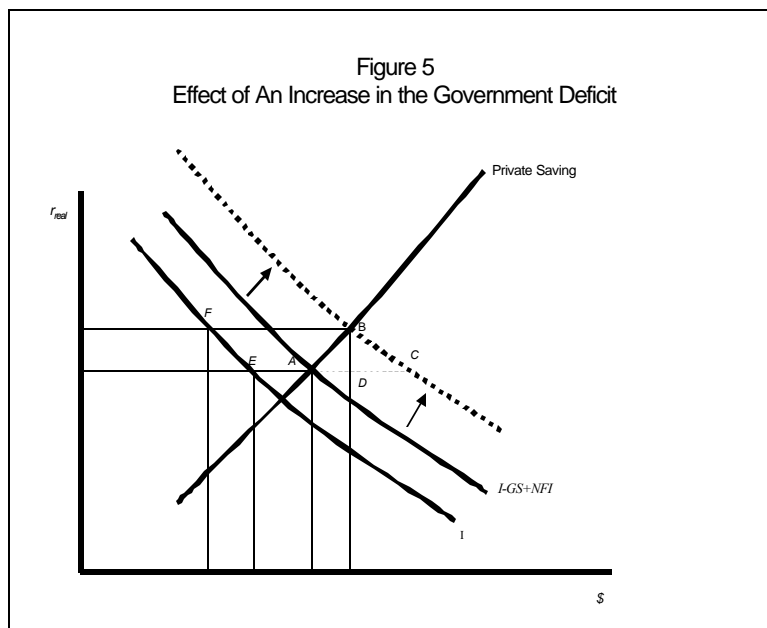
¹ The effect of real interest rates on savings can also be negative. If individuals save in order to attain some future wealth target (e.g., for retirement or in order to purchase a car or a house), then a higher rate of interest makes it easier to attain the target in the future and current saving may decline. However, empirical evidence indicates that there is some positive interest elasticity of saving.



The savings and investment schedules in Figure 4 will shift when any of the other determinants of saving and investment change. Such a change will alter the investment saving equilibrium, interest rates and the distribution of Y^* between consumption and investment goods. We will examine one extremely important change:

The long-run effect of an increase in the government deficit on interest rates and output

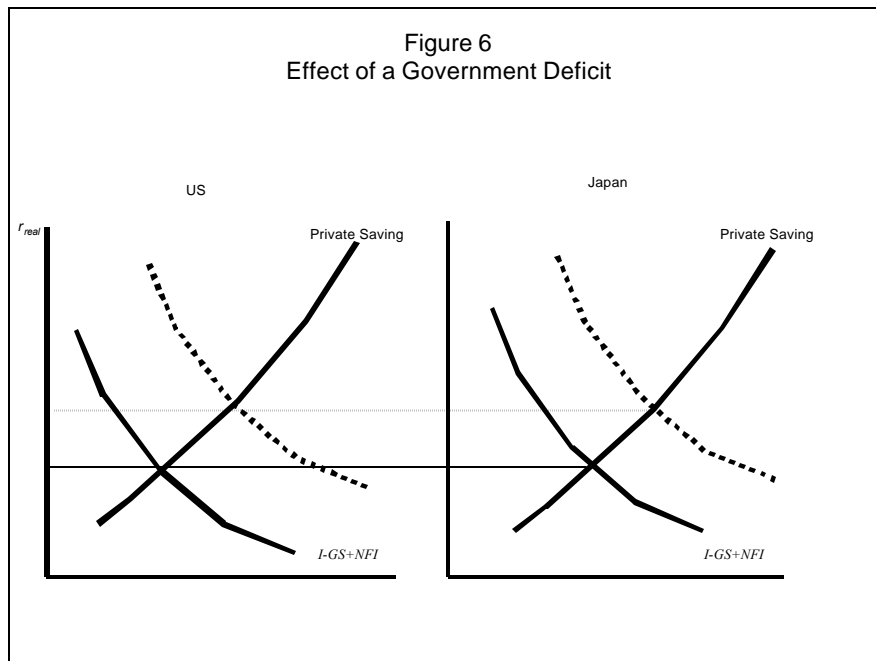
Consider an increase in government expenditure that increases the deficit. Figure 5 shows that Government saving, GS , declines and the investment schedule ($I - GS + NFI$) shifts out to the right. The investment-saving equilibrium shifts from A to B in figure 5 and the real rate of interest increases. Figure 5 also shows the investment schedule (I) without the government saving and foreign investment parts. In the initial equilibrium, the level of investment is indicated by E and in the new equilibrium it falls to F . In the new equilibrium, there is more private saving, less investment and a higher real rate of interest.



This example is important because the increased government deficit **crowds out** investment. The deficit increases by AC in Figure 5 and the total investment and saving increases by a smaller amount, AD . Thus, the deficit is crowding out private investment expenditure—some of Y^* is being shifted from investment goods to government goods and services.

These changes have significant long-run implications because the increased government deficit has reduced the amount of investment. In the new equilibrium, the economy is accumulating fewer investment goods and the capital stock will not be growing as quickly. The government deficit reduced capital accumulation and the long-run potential growth rate for the economy is reduced as well. This, in a nutshell, is the problem with government deficits sustained for any period of time—they have long-run consequences for growth.¹

The relationship between an increase in the government deficit and the current account in the balance of payments is shown in Figure 6. Consider a situation where there are two countries in the world (call them the U.S. and Japan) and the capital flows freely around from country to country. International capital flows lead through arbitrage to a common world real interest rate in both countries. Thus, the investment-savings equilibria in both the U.S. and Japan are initially established at the same real interest rates. When the government deficit increases in the U.S. the $I - GS + NFI$ curve shifts out because GS falls. The increase in real interest rates creates a current account surplus in Japan, private saving at the higher real rate exceeds $I - GS + NFI$. The Japanese economy finances the U.S. deficit through purchases of U.S. bonds and its net foreign investment, NFI , increases.



¹ It is possible that the increase in government expenditures will be devoted to productivity enhancing activities like research and development or the infrastructure of the economy. However, few observers are confident that the government's activities have the same growth potential as private investment.

3. Distribution Model

$$\textit{Private Saving} = I - GS + NFI$$

⇒ Equilibrium real interest rate r_{real}

4. The Fisher relation that relates the real rate of interest to the nominal rate was explained in Chapter I:

$$\textit{Nominal rate} = \textit{real rate} + \textit{inflation}$$

⇒ Nominal interest rate

$$i_{\text{nom}} = r_{\text{real}} + \%DP$$

Money and Prices

The next part of the model is the financial sector that relates the money stock to the price level. We will begin by introducing an asset called money. We define money and examine the monetary sector equilibrium.

Money is the stock of assets used to conduct transactions. In addition, holdings of money represent a store of value because they can always be exchanged for goods. Finally, the money unit of account (e.g., dollar, franc, or yen) provides a common reference unit for quoting prices (called a numeraire). Although there is general agreement about this conceptual definition, it is often difficult to implement it specifically.

What is money? The answer is obvious, even to a toddler who has not problem recognizing bills and coins—currency. However, this definition is much too restrictive. Many objects have served various societies as money (including wampum and even cigarettes). In addition money has changed its physical characteristics over time. In the classical world, money consisted of gold and silver coins. Paper money that was convertible into specie dates to the seventeenth century. It is only recently that paper money that is not convertible into a precious metal has been widely used. Furthermore, in the modern world the role of money is often played by computer entries that need never take on any tangible form.

It is best to define money in terms of the roles that it plays in economic society, rather than in terms of its physical attributes. The most important function of money is that it act as a **medium of exchange**. People use money for purchases and sales of goods and services. Money is then whatever is accepted and used for transactions.

This transactions role of money is extremely important. Imagine the difficulties involved with living in a society without money. If people could not use money in exchange for goods and services, they would need to resort to barter. That is, they would need to exchange goods and services directly. A farmer would need to take her produce and find a cloth maker who not only had the type of cloth the farmer needed for her clothing but who also wanted the kind of produce the farmer had to sell. Next, the farmer would need to find someone who had the type of parts that fit her tractor and who wanted exactly the produce the farmer offered, and so on. The time and effort expended would be enormous.

The savings of time, effort and resources afforded by the use of a uniform money for all transactions provides a great benefit to society. It is no wonder that one of the principal goals of government is to see to it that society has a viable money asset.

Money also serves two other functions. It acts as a standard of value; it is the unit in which prices are expressed. It would be enormously difficult to keep track of prices in the absence of a common reference unit. Finally, money can also act as a store of value. If we hold money, we are able to transfer our purchasing power over time.

What then, are the assets in our economy that serve these purposes and can be called money? In practice, it is not often easy to draw a hard-and-fast line between money and other financial assets. In the United States, the Federal Reserve Board provides several formal definitions for the money stock. A summary of the formal definitions of the money supply is presented in Table 4. The narrow definition $M1$, consists of the assets which are most clearly held for transactions purposes. The broader definitions $M2$ and $M3$ include other financial assets that can be used for transactions with minimum difficulty or can be readily converted into a transaction asset. In addition to the three definitions of the money supply in common use, the Federal Reserve also prepares and monitors two broadly defined money aggregates, L (for liquidity) and *total debt*.

The narrowly defined money stock, $M1$, consists of coins, currency, demand deposits, and other checkable deposits (including travelers checks). **Coins and currency** (paper money) are clearly used for transactions and should be included in even the most narrow measure of money. In addition, **checkable deposits** (deposits at banks and other financial institutions that are subject to payment upon demand) are used and accepted for most transactions. In fact, in the U.S. most transactions use checks.

Some financial assets or instruments are not used directly for transactions but are easily and readily converted into a form that can be used for transactions. These **near-money** assets are included in the broader measures of the money supply. The broadly defined money supply, $M2$, includes all of the items in $M1$ and some instruments that are very easily converted into a transactions balance or can themselves be used for transactions with some restrictions. Thus, in addition to coins, currency, and checkable deposits, $M2$ includes savings deposits (including money market deposit accounts—MMDA), small time deposits (under \$100,000) with a specific maturity, money market mutual funds (MMMF) and overnight repurchase agreements.

There are measures of the money supply that are still broader than $M2$. These measures include financial instruments that are somewhat less easy to use for transactions than are the instruments in $M2$. For example, $M3$ includes all items in $M2$ along with time deposits in excess of \$100,000 term repurchase agreements, and some Eurodollar deposits held by U.S. residents.

TABLE 4**MONETARY AGGREGATES, DECEMBER 1994**

	Billions of \$	Percentage Change over 1993
Currency	353.6	10.0%
Demand deposits	383.3	-0.4
OCDs	402.3	-2.9
Travelers checks	8.4	6.3
	-----	-----
<i>M1</i>	1147.6	1.7
Savings deposits and MMDA	1145.5	-5.8
Small time deposits	818.1	4.1
MMMF	374.5	7.4
Overnight RPS and Euro\$	117.2	27.0
	-----	-----
<i>M2</i>	3600.0	0.9
Large time deposits	363.6	7.3
Term RPs and Euro\$	103.6	7.0
Institutional MMMF	176.6	-10.4
	-----	-----
<i>M3</i>	4282.4	1.2
Other liquid assets		
Savings bonds, short-term Treasuries, Bankers Acceptances, Commercial Paper		
<i>L</i>	5269.9	2.4
<i>Total Debt of Domestic Nonfinancial sectors</i>	12961.0	5.0

NOTE: Totals do not add up due to omitted balancing items.

In the 1960s and 1970s the monetarist economists emphasized the importance of the stock of the transactions asset. As a result, policy-makers began to follow the *M1* aggregate closely. This culminated in a 1979 decision by the Federal Reserve to conduct policy by targeting the growth in *M1*. However, the deregulation of financial institutions and the technological advances in banking since the 1970s has reduced the differences between assets in *M1* and in *M2*. In particular, some checkable deposits earn interest and deposit assets can be instantaneously moved into a transactions asset (at your corner cash machine). Thus, the favored definition of the money stock for policy and for examining the effects of money on the economy is now *M2*.

Some economists feel that the appropriate way of gauging the influence of the monetary sector and monetary policy on the economy is to consider the overall availability of spending power or credit rather than the quantity of transactions assets. Thus, the Fed reports data on total liquid assets, the aggregate labeled *L*, which adds other liquid assets to *M3*. An even broader measure is the total amount of credit or borrowing in the economy, so the federal Reserve also monitors the total debt of the domestic nonfinancial sectors. Note that the growth rates in the different monetary aggregates can differ

substantially; this makes the conduct of monetary policy a complex and confusing matter as we will see later on.

For the present we will assume that there is a uniquely defined set of assets used for transactions purposes which we call money and label M .

The quantity equation relates the stock money to the level of nominal income. The stock of the transactions assets turns over at a rate—called **velocity**, V —in order to generate the transactions that underlie the level of nominal income— PY .

The next element of the long-run equilibrium model is the **quantity theory of money**. It states that the stock of money, M , times the rate of turn over is equal to the level of nominal income, PY :

$$MV = PY$$

Furthermore, the rate of turnover, velocity, will in the long run be determined by the technology of the payments system, so we will take it as given. In addition, the level of real output, Y , is determined by the real sector equilibrium, Y^* . So, if V is given and Y is fixed at Y^* , then the quantity equation tells us that in the long-run the price level, P , is determined by the money stock, M .

The quantity theory of money implies that change in the monetary sector affect the price level and have no other effects. The enormous emphasis on monetary policy suggests that there are some short-run effects of changes in monetary policy and the money stock that we will explore later. For the present, we take a long-run view that states that in the long-run money determines the price level and the rate of money growth determines the inflation rate:

$$\%DM + \%DV = \%D(PY) = \%DP + \%DY$$

SUMMARY: The long-run equilibrium view of the money sector is given by the quantity equation

5. Quantity Equation $MV = PY$

⇒ The money stock determines the price level.

Open Economy Equilibrium Conditions

We will not develop a full model of the open economy at the present, but instead examine some of the long-run equilibrium conditions of an open economy.

The first open economy equilibrium condition was already presented in Chapter I. Recall that the real exchange rate was defined as

$$e_{real} = \frac{p}{p^f} e_{nom}$$

If the real exchange rate is constant, the purchasing power parity is maintained. That is, goods have the same price in the home and foreign countries. Now if all goods are tradable and if the costs of transportation were small, then we would expect purchasing power parity (PPP) to be maintained. I would always buy goods where they are cheapest and so changes in demand would effect prices and exchange rates so that prices would be the same in all places. However, there can be substantial deviations from PPP in the short-run because goods are not always tradable.

I buy Chinese-made shirts because they are less expensive than American-made goods; but my Chinese laundry is not in China even though laundry charges are surely less expensive there than in the U.S.

With a constant real exchange rate, changes in the nominal exchange rate can be explained by relative inflation rates in the home and foreign countries. From the definition of the real exchange rate, we have the **purchasing power parity** condition:

$$\Rightarrow \quad De_{nom} = \%Dp^f - \%Dp$$

Exchange rates do not always change in order to maintain PPP for two reasons. As noted above, goods are often non-tradable. In addition, there are other influences on exchange rates, most importantly interest rates. Our second open economy equilibrium condition relates exchange rates to interest rates and is called appropriately **interest rate parity**.

With the free flow of capital among countries, the expected returns from holding financial assets denominated in different currencies should be the same. That is, the flow of capital should change interest rates and/or exchange rates so that the expected returns are equalized.

The interest rate parity (IRP) condition is given by :

$$(1+i) = (1+i^f) \frac{e_{nom}}{e^{exp}}$$

where i and i^f are the domestic and foreign nominal interest rates respectively; e_{nom} is the nominal exchange rate which will simply be called e below; and e^{exp} is the expected nominal exchange rate (one year hence, assuming that the interest rates are for a one-year period as well). The interest rate parity can best be explained by an example.

Consider a situation where one-year U.S. bonds yield 8% and one-year German bonds yield 6%. The current nominal exchange rate is 2 and the nominal exchange rate expected one year from now is 1.94.

$$i = .08 \quad i^f = 0.06 \quad e = 2.00 \quad e^{exp} = 1.94$$

Are investments in U.S. and German bonds equivalent?

If I buy \$10,000 in U.S. bonds, I will have \$10,800 in one year. Alternatively, I can use my \$10,000 to buy DM today at the exchange rate of 2 and invest the 20,000DM in a German bond that will be worth 21,200DM in one year ($1.06 \times 20,000$). Recall that the expected exchange rate is 1.94 so that the DM investment will be worth $\$10928 = 21200/1.94$. The expected return on the DM bonds was 9.28%, more than the expected return on the U.S. bonds.

I still might decide to hold U.S. bonds. Why? What important difference remains between holding of \$ and DM bonds?

The expected return on DM bonds is calculated from:

$$(1 + .0928) = (1 + .06) (2/1.94)$$

$$\text{Expected return} = \left[(1 + i^f) \frac{e}{e^{exp}} \right] - 1$$

$$\approx i^f - \frac{\Delta e^{\text{exp}}}{e}$$

The expected return is approximately (an approximation that is close enough to be useful) equal to the foreign interest rate less the expected rate of change of the exchange rate. In our example the \$ was expected to appreciate by 3% (from 2DM to 1.94DM); the percentage change in e^{exp} was -0.3 and the expected return was approximately $.06 - (-0.3) = .09$.

The interest rate parity condition is that the expected return on foreign bond holdings shown above is equal to domestic bond returns, i .

A comparison of domestic and foreign interest rates tells us something about the exchange rate expectations of representative market participants. For example, 3-month CD rates in the U.S. and Japan were 3.35% and 2.31% respectively in November 1993. The exchange rate at that time was 107.88 Yen per dollar. The interest rate parity condition can be used to derive the expected exchange rate at that time.

Was the Yen expected to appreciate or depreciate and by how much?

SUMMARY: The open economy equilibrium conditions are:

6. Purchasing Power Parity

$$\%De_{nom} = \%Dp^f - \%Dp$$

7. Interest rate parity

Domestic interest rate is equal to foreign interest rate less the expected rate of exchange rate depreciation.

$$i = i^f - \frac{\Delta e^{\text{exp}}}{e}$$

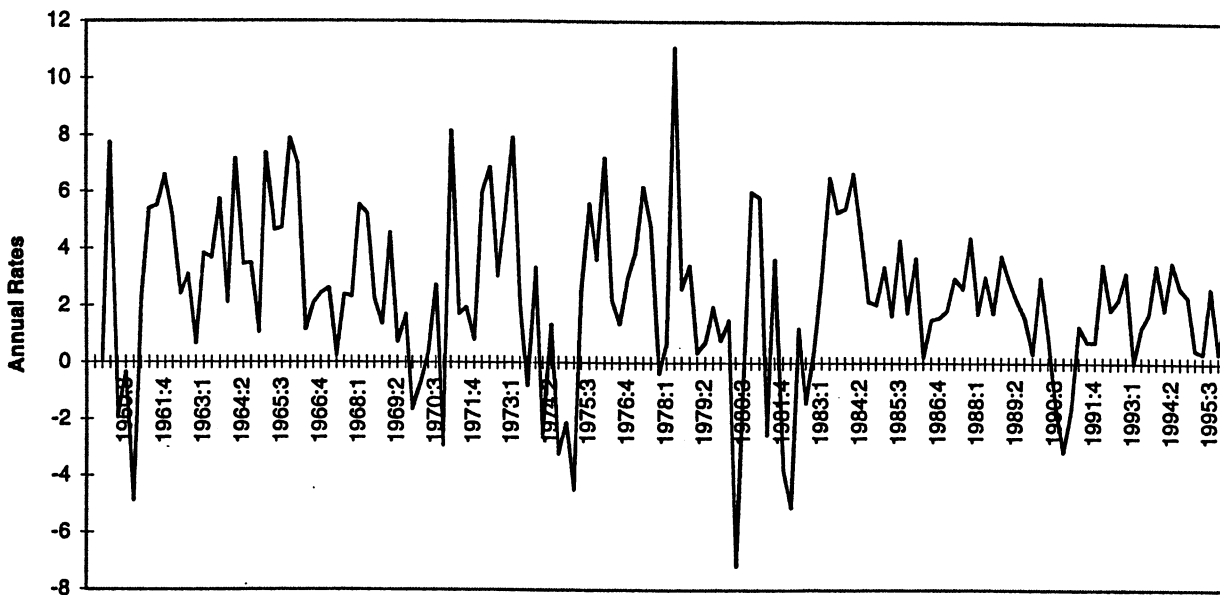
CHAPTER III

UNDERSTANDING SHORT-RUN FLUCTUATIONS

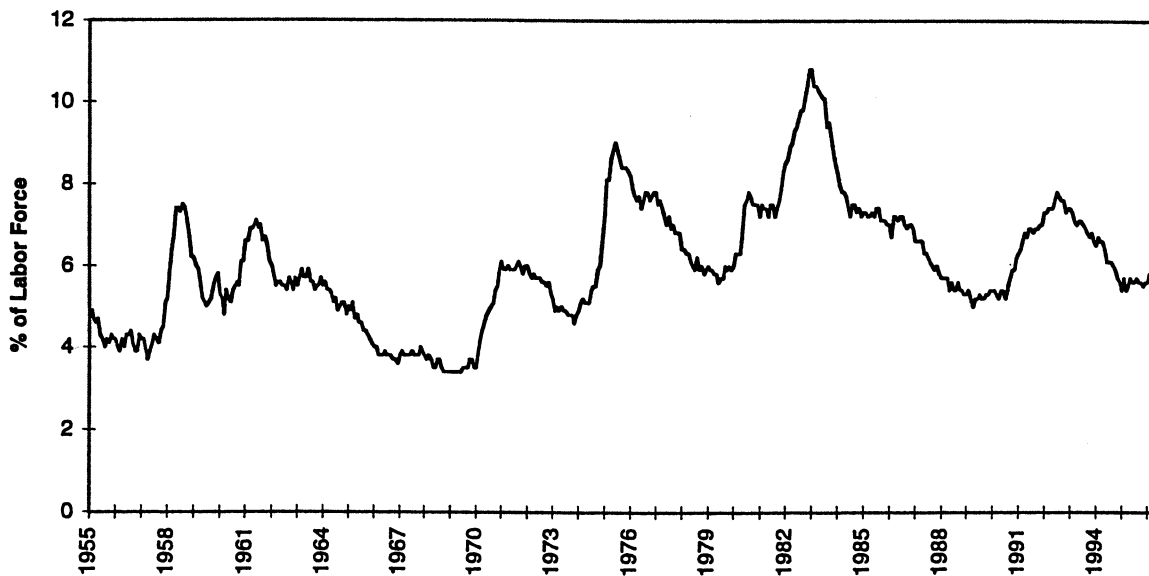
The business cycle fluctuations that are so pervasive in developed economies suggest that the economy is often not at its long run or natural equilibrium. In Chapter I, we defined a business cycle and observed that there have been nine business cycle contractions or recessions in the post-World War II period. Figure 1 shows the importance of short-run macroeconomic fluctuations in the U.S. Furthermore, the conditions that hold in the long-run equilibrium—the Fisher relationship, purchasing power parity, etc.—do not always hold true in the short-run. In this Chapter we will provide a theoretical framework to explain and interpret the rather strong and persistent short-run deviations from long-run equilibrium.

Figure 1

Quarterly % Change in Real GDP



Unemployment Rate in U.S.



This part of macro theory is often called the Keynesian model. It is so-called because its origins are found in the work of John Maynard Keynes, a prominent British economist of the interwar period. Keynes was dissatisfied with the inability of the classical equilibrium approach to provide an adequate understanding of the macro phenomena of his time—massive and persistent unemployment. In the U.K. from the 1920s on, and in the U.S. and elsewhere during the Depression of the 1930s, it was clear that market mechanisms were not leading to an equilibrium that utilized all available resources. The excruciating slowness of adjustments in modern macroeconomics needed some explanation.

Keynes developed a sometimes puzzling ‘general theory’ (as he called it) that did not provide all of the answers but it nevertheless had an overwhelming influence on post-war macroeconomic thinking. In fact, for a time Keynesian modeling became synonymous with macroeconomics.

The Keynesian approach examines the aggregate demand for goods in real terms and describes the adjustments of the economy towards equilibrium. The simplest Keynesian model abstracts entirely from price changes. As unreasonable as that sounds, the model is useful for understanding how short-run output changes occur. The broader Keynesian model sheds light on the interactions between the demand for goods and services and the financial sector. It will provide the basis for examining the short-run effects of monetary and fiscal policy. The Keynesian model is useful for understanding how the economy adjusts towards its long-run equilibrium.

QUANTITY ADJUSTMENT PARADIGM

The principal theme of classical price theory that underlies the study of microeconomics is:

If there is an imbalance between supply and demand in a competitive market, then prices change to clear the market or establish equilibrium.

The emphasis in price theory is on the balancing role of prices. Similarly, the long-run equilibrium model in Chapter II utilizes the equilibrating role of prices. The prices that maintain equilibrium in that model are the real wage, the real interest rate and the exchange rate—the prices of labor, time and foreign currency. Real wage flexibility brings about labor sector equilibrium and the real interest rate balances the allocation of output between investment goods and consumption goods. Together, the equilibrium labor supply and capital stock (in turn determined by the amount of investment) determine the equilibrium output level.

The Keynesian real sector model takes a different approach. The underlying theme of this approach is that output or quantity adjustments occur when there is an imbalance between supply and demand. That is, in the short-run, disequilibrium leads to changes in the quantity of output. In order to explore this approach we will assume that prices remain constant and examine the macroeconomic consequences of a model where quantity adjustments occur. Clearly, both price and quantity adjustments take place in actuality; so we will relax the assumption shortly.

The simple Keynesian model is an application of the quantity adjustment paradigm. That is:

If there is an imbalance between supply (output or production) and demand (expenditure), then producers will change the quantity of output produced.

We will start by showing why quantity adjustments are the most common responses, particularly in short-run periods and when output is non-perishable. The argument can best be made by citing an example of a manufacturing industry where adjustments in the quantity of output are, at least in the short-run, more important and more pervasive than adjustments in prices. We will then provide a more general discussion of price changes and the reasons why they do not occur so frequently.

Consider the case of an automobile manufacturer that monitors the level of inventories in order to judge whether there is an imbalance between sales and production. If sales fall off, inventories will begin to rise about the desired level. Once the imbalance is recognized, some type of adjustment is necessary. The producer could reduce the price of output and/or cut the level of production.

A cut in price, perhaps in the form of price rebates that are often used in the auto industry, would encourage sales and solve the inventory buildup problem. However, a price change takes time to put in place and the exact response of the consumers is often hard to judge. In the short-run, a quicker and more certain solution is found by making changes in the level of output. Automobile manufacturers often respond to excess inventory accumulation by reducing the quantity of automobiles produced. Price changes—often in the form of price rebates—occur but less frequently.

The automobile manufacturer chooses a quantity, or output, adjustment instead of a price adjustment because the outcome is more certain. If new cars are mounting up in the manufacturer's storage lots, a cut in output will almost certainly clear this unwanted inventory accumulation (although it could be frustrated by an unanticipated decline in sales). The firm knows the costs of shutting down an assembly line and laying off workers and will generally be able to plan the adjustment process. Nevertheless, price rebates and other forms of price adjustments have become more common in the automobile industry in recent years. The reason for this is that the costs of quantity adjustments have risen because the production work force is on a guaranteed annual wage.

Generally, price reduction is a very problematic way of removing the inventory imbalance. In order to use a price response, the manufacturer would have to have a precise estimate of the sensitivity of sales to price changes (the elasticity). Furthermore, a price reduction must be advertised, and it can take weeks or months before sales respond to a price change. Nevertheless, price adjustments are used in the automobile industry. Price rebates are offered when inventories are extremely large because in such instances the needed quantity adjustments (closing down production altogether) can be too costly.

All in all, in the short run, a change in price is often a costly and unreliable means of equilibrating sales and production, particularly when the imbalance may well be a temporary phenomenon and a rapid response is desired. Thus, quantity adjustments are the primary adjustment mechanism used to maintain sales-production equilibrium in the short run. This conclusion relies upon two characteristics of the product market. First, the argument presumes that goods can be held in inventory (which is by and large true for manufactured goods). If output were perishable, price adjustments would dominate. If we go to the wholesale produce market, we will find that the market for strawberries is cleared daily by price adjustments. Second, we assume that there is some degree of product differentiation. If all output were homogeneous and markets perfectly competitive then price adjustments would clear the market.

Some readers may find the argument for price adjustments too strong to be believable. In a modern economy with very rapid information flows and less emphasis on manufacturing and greater emphasis on services, price changes do occur. The example that comes to mind is airplane fares, which seem to be in a constant state of flux. Thus, we will present a less stringent argument. That is, price changes are often made infrequently and therefore the principal adjustments to supply and demand imbalances in the short-run are quantity changes. We will make this argument by showing (i) that most price changes are infrequent, and (ii) the reasons why this is rational economic behavior.

As an enterprising Economics professor, Alan Blinder (who has also served as vice-chairman of the Federal Reserve Board) interviewed corporations about their price change behavior.¹ He simply asked:

“How often do the prices of your most important products change in a typical year?”

¹ See Alan Binder, “Why Are Prices Sticky? Preliminary Results From an Interview Study,” *American Economic Review Papers and Proceedings*, May 1991.

Less than 15 percent of the firms reported that prices are changed more frequently than quarterly. The median response was annually and Blinder concluded that annual price change seems to describe behavior. Firms also report delays of three or four months after a significant change in costs or demand before price adjustments are made. Indeed, prices in the American economy are quite *sticky*.

There are many reasons why a rational decision-maker might keep prices constant in the face of changing costs or demand. Blinder asked corporate executives to indicate which reasons were most important to them. Among the reasons cited, in order of importance are:

- firms respond to shocks by changing delivery lags or altering the quality of the auxiliary services provided;
- firms are reluctant to change prices because they do not know whether other firms will follow suit;
- firms base their prices on cost;
- firms have implicit contractual arrangements (an “invisible handshake”) with customers which proscribe price increases when markets are tight and thereby preserve long-term customer relationships;
- firms have explicit contractual relationships which fix prices for a period of time;
- there are costs of price change (so-called menu costs—the cost of printing new menus) that inhibit price changes.

A characteristic of the modern macroeconomy is that price changes tend to occur gradually. Thus, we will develop the Keynesian approach to macroeconomics where, to begin, we take prices as fixed and concentrate on the implications of output adjustments.

THE PLANNED AGGREGATE DEMAND MODEL

In this section we will construct the Keynesian model of aggregate demand that shows how the quantity adjustment paradigm works. The model simply examines whether actual output— Y^s —and the desire to absorb output—planned aggregate demand, Y^d —are equal. If *actual output* differs from *planned aggregate demand*, then quantity adjustments bring the economy back to balance.

To see how this works, start with the accounting definitions from Chapter I. Recall that real GDP has four principal components:

C	=	consumption expenditures
I	=	investment or capital formation
G	=	government expenditure on goods and services
NX	=	$(X-M)$ net exports

Total real GDP is equal to the sum of these components. For convenience, we ignore the accounting differences between Gross Domestic Product and national Income and call the sum simply real output or income. That is, total expenditure on all final goods and services is total real output or income, which we call Y :

$$Y = C + I + G + NX \quad (1)$$

Since the value of output produced is equivalent to income earned in production, Y can be called either real output or real income. Equation (1) is an identity that shows how actual output is allocated among the various expenditure categories and that the sum of expenditures is equal to income earned.

Our point of departure from the accounting for actual output described by the identity (1) is to consider the underlying supply and demand behavior. First, we will call the actual output decision or the supply of output Y^s . Second, the amount of final output that the households, business enterprises, and governments want to absorb or demand is called planned aggregate demand— Y^D . Y^s and particularly Y^D will vary with economic circumstances.

Real sector equilibrium is defined as a situation where the overall desire to absorb product— Y^D —is exactly equal to the output produced — Y^s :

$$Y^s = Y^D \quad (2)$$

Equilibrium in this context is a short-run equilibrium. We have, for the time being, dropped the assumptions that defined the long-run equilibrium level of output. In the long run, equilibrium output is Y^* which is determined by technology and resource availability (A , K and N). We will not allow output to vary from Y^* and will analyze supply and demand forces that influence output in the short-run. Furthermore, we will assume that in the short-run wage and price adjustments do not occur, an assumption that we will relax later.

We will consider the forces that determine planned aggregate demand— Y^D —and then consider what happens when actual output and the sum of expenditure plans are not equal. That is, we will specify the determinants of consumption and investor behavior as well as the level of government expenditure and net exports. There are of course numerous factors that can influence the demand behavior of consumers or investors. We will specify a simple model that is restricted to just the principle determinants.

A key contribution of the Keynesian approach to the real sector is that the major determinant of planned aggregate demand is the level of income itself. This relationship is the strongest for consumption expenditures that are the biggest—about two-thirds—part of output.

Keynes hypothesized that planned consumption expenditures increase with income and, importantly, a \$1 increase in income earned leads to a less than \$1 increase in planned consumption. He termed the ratio of the change in consumption to the change in income the *marginal propensity to consume*. We will retain Keynes' assumption because it is supported by the bulk of empirical evidence. The consumption relationship shown below relates consumption to after tax income, $(Y - T)$, where T is the level of taxes.

The next determinant of Y^D to consider is the level of interest rates, R . Our discussion of the effect of interest rates on investment and saving from Chapter II is relevant here. Investment demand is a decreasing function of interest rates. Saving is a positive function of interest rates and therefore consumption is a negative function of rates.

Net exports will also depend on the level of output and income because imports increase with the level of income. Net exports will also depend on the exchange rate. An exchange rate appreciation—increase in e —leads to fewer exports and more imports. Thus, net exports is a decreasing function of exchange rates.

To summarize, the relationship between output and planned aggregate demand can be written as:

$$Y^D = C^d + I^d + G + NX^d$$

$$Y^D = C^d([Y - T], R) + I^d(R) + G + NX^d(Y, e)$$

where, for example, $C^d(\dots)$ means that consumption is a function of the items in parentheses and the signs of the affects are shown below. The level of government expenditure is taken to be exogenous—or determined independent of the economic structure.

Now, consider what happens if $Y^D > Y^S$. If expenditure or the demand for output exceeds the amount being produced, then there will be an unplanned or undesired depletion of inventories. Producers will adjust the quantity of output and so Y^S will rise to Y^D . In this simple model we assume that capacity constraints do not create any difficulties and that prices are unaffected. There is still one complication to note in this simple quantity adjustment process. As output increases, planned aggregate demand will increase as well, given our specifications above for C^d and NX^d . There will be a stable adjustment to a new equilibrium as long as an increase of output or income leads to a somewhat smaller increase in planned aggregate demand. This is why Keynes' assumption about the marginal propensity to consume is important. Generally, we assume (and as stated earlier, it is consistent with the facts) that an increase in income leads to a somewhat smaller increment to planned aggregate demand.

THE SIMPLE KEYNESIAN MULTIPLIER

The planned aggregate demand model can be used to demonstrate the Keynesian multiplier relationship. The model is highly simplified—there is no worry about price changes or capacity constraints. The multiplier relationship is the effect on the quantity equilibrium of some outside or exogenous change that affects planned aggregate demand.

We can use a specific algebraic version of the multiplier relationship to illustrate the effects. The algebraic model consists of two parts: an identity that defines Keynesian equilibrium and a specification of planned aggregate demand behavior.

The equilibrium condition (2) is that actual output is equal to planned aggregate demand: $Y = Y^D$. The components of planned aggregate demand are the same as in the previous section, consumption (C), investment (I), government expenditure (G) and net exports (NX).

Consumption is a function of income (or output, Y ; they are the same in this simple framework) after taxes, T , and interest rates:

$$C = C([Y - T], R)$$

where $C()$ indicates the functional relationship. The marginal propensity to consume, $\frac{dC}{dY}$, will be indicated by C_Y and is assumed to be less than one. To add a further step of realism, let taxes be a function of income:

$$T = T(Y)$$

Similarly, the investment and net exports functions are:

$$I = I(R) \text{ and } NX = NX(Y, e)$$

Thus, planned aggregate demand is the sum of its components:

$$Y^D = C[Y - T(Y)] + I(R) + G + NX(Y, e) \quad (3)$$

The equilibrium level of income, Y , is the solution to equation 3. The equilibrium condition is:

$$Y = Y^S = Y^D$$

We can see how the equilibrium value changes when there is an autonomous change in one of the determinants of demand by substituting for PAD from (2) in (3) and taking the total differential of equation (3):

$$dY = C_Y(dY - T'dY) + I_R dR + dG + NX_Y dY = NX_e de$$

We can quickly see the effect on the equilibrium level of income of a government expenditure change. The effect on income—the government expenditure multiplier—is seen by setting the changes in the other exogenous variables to zero ($dR = 0$ and $de = 0$) and solving for dY/dG :

$$\frac{dY}{dG} = \frac{1}{1 - C^Y(1 - T') - NY_Y}$$

The size of the government expenditure multiplier in our model depends on several parameters—the effects on C , T and NX of a change in income or the extent to which PAD changes when Y changes.

We can estimate the size of the multiplier by examining aggregate data to get reasonable estimates of the two parameters. The average ratio of consumption to GDP after taxes is about .8, and taxes are about one-fourth of GDP ; the effect of income changes on net exports are so small that this term can be ignored. These rough estimates give a multiplier of $2 \frac{1}{2}$ with our model.

Although this simple model ignores many important economic relationships, it can be useful. For example, the expansion of output during the Vietnam war years can be viewed as an application of the multiplier process described in this section. As shown below, the major change in PAD between 1965 and 1968 was a very large increase (almost one-fourth) in the government's real demand for goods and services. The much smaller changes in I and NX canceled each other out, so we can take autonomous expenditure as unchanged. The increase in income is just about $2 \frac{1}{2}$ times the increase in G .

The mid-1960s episode is well described by the simple multiplier model. In those years, the primary influence on the economy was the war-related expansion in demand and its impact through the multiplier process and quantity adjustments on total output. Of course, if we look into developments in subsequent years, we would have to consider effects on interest rates, inflation rates and exchange rates as well. But that will have to wait the development of a more general macroeconomic model.

	GNP in constant (1982) dollars				
	Y =	C +	I +	G +	NX
1965	2087.6	1236.4	367.0	487.0	-2.7
1968	2365.6	1405.9	391.8	597.6	-29.7
	-----	-----	-----	-----	-----
	278.0	169.5	24.8	110.6	-27.0

DETERMINANTS OF AGGREGATE DEMAND

The different components of planned aggregate demand— C^d , I^d , NX^d —are affected by many different things. The functional specifications shown above are simplifications that focus on the most important determinants of demand. In this section, we will provide a somewhat more general discussion of the determinants of aggregate demand.

Consumption

Modern theories of consumption demand emphasize that consumption expenditures are not only related to current income but are also affected by the household's overall command over resources. That is, the wealth of the financial sector will influence its consumption plans. Wealth can be defined broadly to include actual assets held (financial assets, real estate, etc.) and also expected future income.

The role of wealth and expected income in determining consumption is very well illustrated by an episode that occurred in the late 1960s, during the war in Vietnam. The multiplier expansion described above led to a rapid expansion in output. By late 1968, the unemployment rate dropped to a historically low level of 3.4%. Policy-makers were concerned that the economy was operating at full capacity and that any further expansion to government expenditure would simply lead to the emergence of inflationary pressures. Once labor and productive capacity were fully utilized, quantity adjustments should no longer be feasible and price adjustments to excess demand pressures would be inevitable.

The policy-makers had a ready solution. If taxes were increased, then disposable income would fall and consumption demand would decline, as described by our simple multiplier model above. This would set in motion a multiplier process in reverse and prevent the macroeconomy from overheating. Because of the multiplier process, only a small tax increase was deemed necessary to constrain the economy.

However, as we well know tax increases are not politically popular. To mute the pain in an election year, Congress passed a temporary tax surcharge and every election campaign emphasized the temporary nature of this war surcharge. As a result, the tax increase, although it reduced current income, had little effect on consumption demand. The public made the reasonable and correct inference that a temporary surcharge has little effect on overall lifetime resources and maintained its customary level of consumption demand by temporarily reducing savings. The policy-makers were using our simple model in forecasting the effect of a tax increase on output through the multiplier. But the public was smarter; the tax increase did not hold back an overheating economy and inflationary pressures did begin to emerge.

Investment

As noted earlier, investment expenditures are sensitive to the rate of interest. We will begin by showing why this is the case and then discuss other determinants of investment.

Consider how a firm might evaluate an investment project that is *expected to generate a stream of returns for some time into the future which we can denote as: $RET_1, RET_2, RET_3, \dots$* . We define the internal rate of return, IRR , on the project as the discount rate that equates the stream of future returns to the project's cost:

$$\text{Cost} = \sum_t \frac{RET}{(1 + IRR)^t}$$

As long as the cost of capital (from borrowing and/or obtaining equity capital) is less than the internal rate of return, the project should be undertaken. When the cost of capital (or the interest rate for financing) is less than IRR , the project is expected to be profitable. The lower the interest rate relative to the internal rate of return, the larger the number of investment projects that will be attractive. Thus, we can write a function for investment demand as:

$$I = f(IRR - R)$$

where R is the interest rate. The IRR will depend on the size of the capital stock already in place, which does not change very much over time. Thus, for short run analysis we can consider investment to be a decreasing function of interest rates:

$$I = I(R)$$

Although investment is only about 16% of GDP, it is very important for two reasons. First, it varies a great deal and affects the behavior of GDP growth in the short run. Second, the size of investment determines the growth in the stock of capital, which in turn determines productivity growth, and the extent to which real living standards increase in the long run.

The interest rate is used here as a general measure of the cost of capital. However, this is a simplification that ignores the vast number of different ways of financing investment projects. The relevant interest rate should represent the costs of equity financing, bank loans, private placements, bond financing, etc., and, also, financing with short-term obligations and long-term obligations. In empirical macroeconomic studies, it is common to relate investment to a cost of capital that is an average of different financing costs.

Two other empirically significant determinants of investment are corporate cash flow and the change in output. Output increases are a good indication that capacity constraints may soon be reached and therefore, large output increases increase the demand for investment goods. Firms are anxious to plan their production capacity needs in advance. Much business investment is financed internally by firms that use their own profits to finance investment. Of course, firms will still compare the internal rate of return on a project to other alternatives such as putting the accumulated profits in the bank. Nevertheless, cash flow will influence investment activity because the use of internal funds saves any costs of obtaining financing. Fees and the necessity to provide information to outside investors or lenders may make financing from external sources less attractive. Thus, when cash flow increases, there is likely to be a rapid increase in investment demand.

Foreign Sector

The net exports component of planned aggregate demand depends first on the domestic and foreign income and second on the relative prices of foreign and domestic goods.

As Y increases, the demand for imported goods will rise and therefore NX falls. Furthermore, the export component of NX will depend on economic conditions abroad. Expanding foreign economies will demand more American goods.

NX also depends on the relative price of foreign and domestic goods. If the price of foreign goods rises relative to domestic goods, then NX increases. First of all, Americans will shift toward domestic production and imports fall. Secondly, foreigners will find American goods more attractive and exports will increase. In Chapter I we defined the real exchange rate which measures the relative price of foreign and domestic goods. In most of our discussions of the simple Keynesian model, we assume that prices are constant in the short-run and therefore the nominal exchange rate determines NX .

As a reminder: The exchange rate is defined as the units of foreign currency per dollar. Thus, when the newspaper reports the Yen exchange rate at 130, it means that 130 Yen trade for one dollar. If the exchange rate **declines**, then the dollar has **depreciated** in value.

An exchange rate depreciation makes foreign goods more expensive. It takes more dollars to buy a 10000Yen item and import demand will decline. In addition, the Yen price of an American \$100 item will fall which makes it cheaper to the Japanese consumer. So, an exchange rate depreciation increases exports and decreases imports. It causes NX and planned aggregate demand to increase.

THE KEYNESIAN REAL SECTOR MODEL

As we can see, the Keynesian real sector equilibrium can be affected by a large number of economic, political and other phenomena. For the moment we will concentrate on the influence of output and interest rates— Y and R . The reason for this is simply that we will be building some tools that will make it easy to expand the model to include the financial sector.

Interest rates are important because they are the linkage between the real sector, the market for goods, and the financial sector, the market for money. In fact, there are two building blocks of the Keynesian approach—the quantity adjustment paradigm that underlies our discussion of Keynesian real sector

equilibrium and the Keynesian emphasis on the short-run role of interest rates in the monetary sector. Thus, our emphasis here will be on the joint determination of output and interest rates and the short-run interactions between the real sector and the monetary sector.

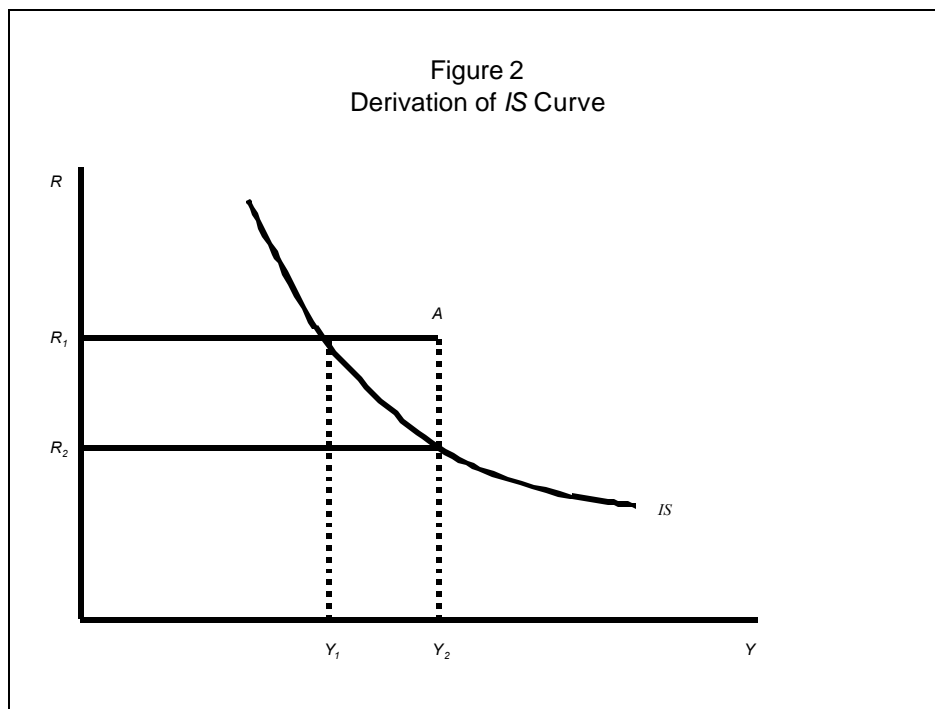
We will call the model that examines both of these markets—goods and money markets—the full Keynesian model and will show how the markets interact and determine equilibrium values for both output (Y) and interest rates (R). We will start by developing a summary presentation for the relationship between interest rates and the aggregate demand equilibrium. We will then develop the Keynesian approach to the monetary sector of the economy. Finally, we combine the two sectors to show how the full Keynesian model works.

In the last section we showed how changes in interest rates affect planned aggregate demand. An increase in the rate of interest increases the cost of financing investment projects and thus, makes fewer investment projects profitable—planned aggregate demand is reduced.

There will be different possible levels of output and interest rates— Y and R —for which the real sector will be in short-run equilibrium. That is, for some values the level of planned aggregate demand will be exactly equal to the level of output produced:

$$Y = C^d(Y - T, R) + I^d(R) + G + NX^d(Y, e)$$

A particularly useful way of looking at the different possible output equilibrium is found in Figure 2. The interest rate is measured on the vertical axis and output is on the horizontal axis. Given the level of the other variables, called exogenous or determined elsewhere— G , T and e —there are different possible equilibria where planned aggregated demand (the right hand side of the equation) is equal to Y . This line, labeled IS , is a locus of points that represent real sector equilibrium. For each interest rate, it gives the level of equilibrium output. Of course, for a given IS curve, all the other exogenous factors that might affect PAD are held constant (i.e., G , T , e , and anything else that might affect demand).



The curve is called an *IS* curve because along this curve planned Investment is equal to planned Saving. More generally, there is a real sector equilibrium where planned aggregate demand is equal to output or income.

Points off of the *IS* curve represent disequilibria. For example, examine the point labeled *A* in Figure 2. For the interest rate at *A*, R_1 , the output equilibrium, Y_1 , is less than current output, Y_2 . At this point, output exceeds planned aggregate demand and producers will be accumulating undesired or unplanned inventories of goods. Their appropriate quantity response is to reduce output. Generally, points to the right of the *IS* curve will represent excess supply and points to the left will represent excess demand.

Properties of the *IS* Curve

IS curve slope. The *IS* curve shown in Figure 2 has a negative slope; it can be readily shown why it is drawn in this way. Consider again the above model where the consumption component of *PAD* depends on after tax income— $C(Y - T)$ —and the investment component depends on interest rates— $I(R)$ —and net exports depends on output and the exchange rate— $NX(Y, e)$. Furthermore, assume that taxes, government expenditure and the exchange rate are exogenous. The real sector equilibrium is given by equation (3) above where output is set equal to the sum of the components of *PAD*.

The slope of the *IS* curve is obtained by taking the total differential with the values of G , T and e held constant:

$$dY = C_Y(dY - T' dY) + C_R dR + I_R dR + NX_Y dY$$

where C_Y is the marginal propensity to consume, NX_Y is minus the marginal propensity to import and I_R is the effect of a change in interest rates on investment. The slope of the *IS* curve is:

$$\frac{dR}{dY} = \frac{(1 - C_Y[1 - T']) - NX_Y}{C_R + I_R}$$

The slope of the *IS* curve is negative because $I_R < 0$ and, as an empirical matter, the marginal propensity to consume is less than one and the marginal propensity to import is small. If investment is unaffected by interest rates, $I_R = 0$ and the slope is infinite; the *IS* curve is vertical.

The size of the slope of the *IS* curve is important as well because it reflects the extent to which planned aggregate demand responds to changes in interest rates. If investment is relatively interest elastic, then a change in interest rates has a large effect on investment demand. In this instance the *IS* curve will be relatively flat. If *PAD* is interest inelastic then the *IS* curve is relatively steep. An extreme case is when *PAD* is completely unaffected by changes in interest rates. In this case the *IS* curve is vertical.

IS curve shifts. The position of the *IS* curve is determined by the variables other than the interest rate which shift the *PAD* curve. That is, the position of the *IS* curve is determined by the level of the exogenous or autonomous components of *PAD*. By autonomous we mean those parts of *PAD* which are not influenced by the level of income.

For example, imagine that the dollar exchange rate depreciates which leads to an improvement in the balance of trade; NX increases. Thus, planned aggregate demand for any given interest rate increase and the level of output that equates Y and planned aggregate demand is greater. The *IS* curve shifts out or to the right.

This concludes the presentation of the *IS* curve, a tool that summarizes the real sector model equilibrium concept and the role of interest rates. It will be a useful apparatus once we add on a model for the monetary sector.

MONETARY SECTOR

In this section we will introduce the role of money. A definition of money was presented in Chapter II. We will now show how movements in interest rates maintain equilibrium between the supply and demand for money. The monetary sector equilibrium will be summarized by the LM curve. When the IS and LM curves are combined, we will be able to examine the interaction between the real and monetary sectors.

Money Demand Function

In our presentation of macroeconomics there are two major Keynesian contributions. We have already examined the first, the quantity adjustment paradigm as a way of looking at real sector adjustments. The second Keynesian innovation is the money demand function. It was innovative because classical thinking did not investigate the desire to hold money balances. However, Keynes' own explanation of money demand turned out to be inadequate. Thus, we will use some widely accepted post-Keynesian ideas to specify the money demand function.

Our discussion of the real or goods sector focused on the real or constant dollar level of demand or output, Y . To begin we will examine money demand by specifying the demand for real money balances or money holding measured in terms of its purchasing power. Since the stock of money assets, M , is a nominal quantity, real money balances are given by (M/P) where P is the aggregate price level.

Our specification of the money demand function or the demand for real money balances is:

$$\left(\frac{M}{P}\right)^D = L(Y, R)$$

That is, the demand for real money balances is a function of **liquidity preference** to use Keynes' own terminology. Thus, we use the letter L to denote the money demand function. We will describe below that liquidity preference depends on income and interest rates. Thus, the arguments of the function, $L(\cdot)$, are the level of real income or output, Y , and the interest rate, R . Our first task is to show how and why real income and the interest rate affect the demand for money.

Money demand is positively related to the level of income because the volume of transactions increases with Y . Thus, as Y increases the demand for money holding will increase as well.

Next, we turn to the relationship between money demand and the interest rate. Why does the demand for money—the transactions asset—depend on the interest rate? This was an intriguing and controversial issue in the early years of Keynesian economics. The answer, which is now standard, is not the one that Keynes suggested.

We will provide a post-Keynesian argument that relates the demand for transactions balances (money) to interest rates that was developed in the early 1950s by James Tobin and William Baumol. Consider an individual or business which will be making a certain amount of real expenditures (transactions) per year. Should it plan to start the year with sufficient holdings of money to cover all transactions, or should it plan to make financial transactions during the year to replenish its money stock holdings? In the first instance it will have an average stock of money of $\frac{M}{2}$. In the second instance it will have an average money stock of $\frac{M}{k2}$, where k is the number of financial transactions (or sales of assets for money) which are made. The amount of money, which will on average be held, will depend on the k chosen.

The choice of k depends in turn on the cost of transactions and the opportunity cost of holding money. The transactions costs are the bank fees, brokerage commission or ATM charge that is incurred when an asset (e.g., a bond or savings account) is converted to a transaction fee and, also, the value of your time used in the transaction. The **opportunity cost** is the earnings foregone when money is held instead of the alternative financial asset (typically, a financial asset which earns interest). If k is larger, then there are more transactions and more fees paid. However, when k is larger the average cash balance is smaller and the opportunity costs—the foregone interest earnings—are smaller. The choice of k will depend on the size of the opportunity costs.

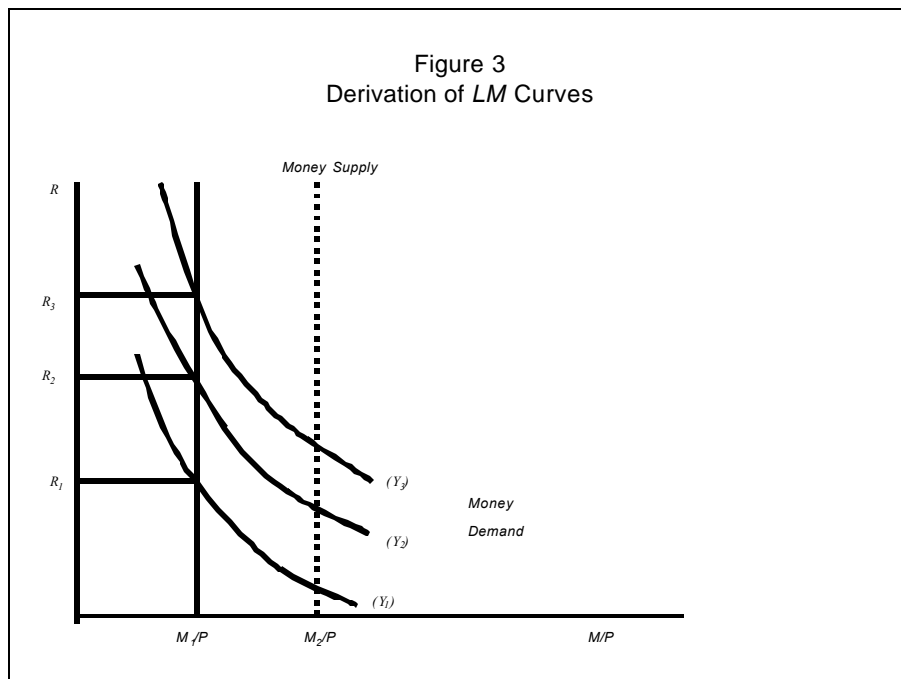
The direction of the argument should be clear. If the interest rate goes up, the opportunity cost of holding money increases and it will be advantageous to make more frequent financial transactions (choose a higher k). In other words, as interest rates rise, the average money balance held will decline. Thus, the demand for money is inversely related to the rate of interest.

Money Market Equilibrium

The supply of money or the stock in existence is determined by the interaction of the banking system and monetary policy. We will discuss this process later and for the present we will take the size of the real money supply to be exogenous. Equilibrium in the money market exists when the demand for money,

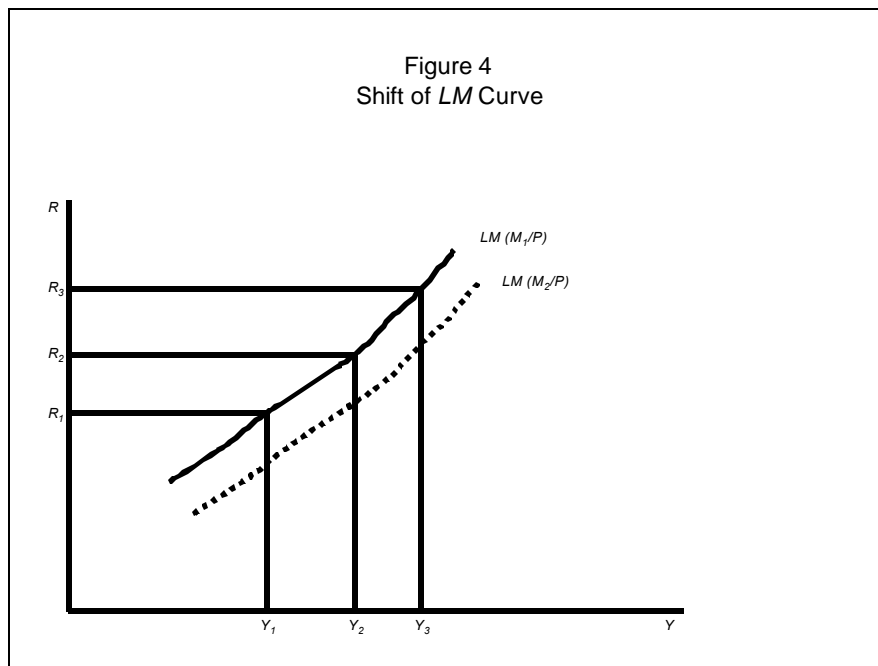
liquidity preference— $L(Y,R)$, is equal to the exogenous supply of real money balances, $\frac{M}{P}$.

Money market equilibrium is illustrated in Figure 3. The quantity of real money is measured along the horizontal axis and the interest rate is measured along the vertical axis. Since money demand is inversely related to interest rates, the money demand function is negatively sloped. There is a family of money demand curves shown because the relationship between money demand and R depends on the level of output as well. For any given level of real output, there is a negatively sloped money demand curve. The money demand curve shifts out as Y increases. In Figure 3, $Y_3 > Y_2 > Y_1$. Since the money supply is exogenous, the given money supply is indicated by a vertical line; it is unaffected by interest rates.



Thus for an exogenous real money supply of $\frac{M_1}{P}$, the possible money market equilibria are indicated. That is, there are different combinations of Y and R that lead to a level of real money demand equal to $\frac{M_1}{P}$ and therefore money market equilibrium.

These different points of money market equilibrium are summarized by the LM curve in Figure 4. It is called the LM curve because liquidity preference (money demand) is equal to the real money supply. We will shortly examine a model of equilibrium in the goods market (summarized by the IS curve) and equilibrium in the money market (summarized by the LM curve) that provides us with an explanation of how output and interest rates are determined. Figure 4 shows money market equilibrium at R_1 when output is Y_1 , R_2 for output Y_2 , etc. These are points along the solid LM curve in Figure 4. Similarly, for a higher money supply, the LM curve is shown by the dotted line.



Properties of the LM Curve

The LM curve represents a set of money market equilibrium. That is, supply equals demand:

$$\left(\frac{M}{P}\right)^D = \left(\frac{M}{P}\right)^S$$

We can drop the superscripts and write the LM curve for money market equilibrium as:

$$\left(\frac{M}{P}\right) = L(Y, R)$$

where the real money stock, M/P , is exogenously determined. It is clear from Figure 4 that the LM curve is positively sloped. We can see it as well by taking the total differential of money market equilibrium condition:

$$d\left(\frac{M}{P}\right) = L_Y dY + L_R dR$$

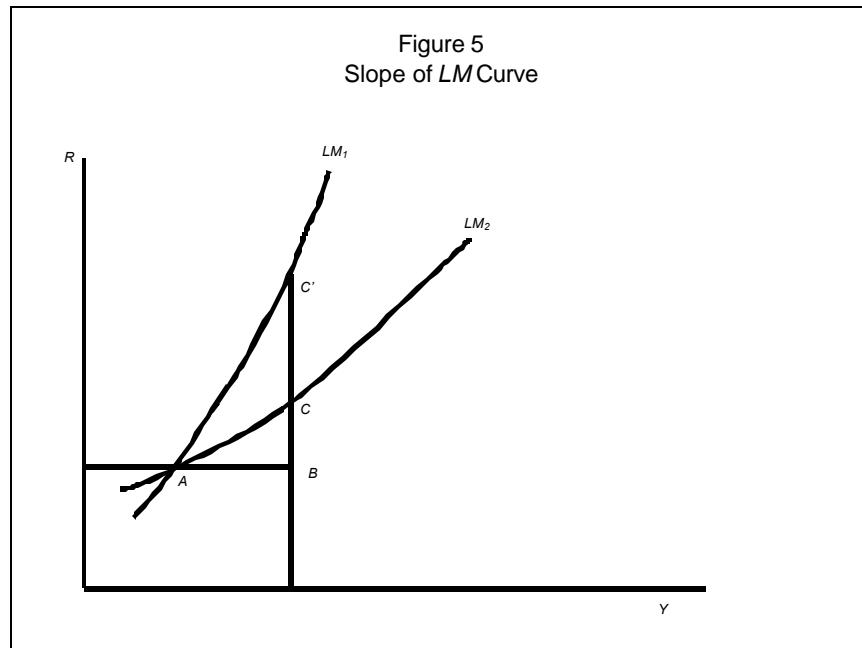
An LM curve is drawn for a given level of the real money stock so along an LM curve $d\left(\frac{M}{P}\right) = 0$ and the curve's slope is:

$$\frac{dR}{dY} = \frac{-L_Y}{L_R}$$

which is positive because L_Y , the effect of output on money demand is positive and L_R , the effect of interest rates on money demand is negative.

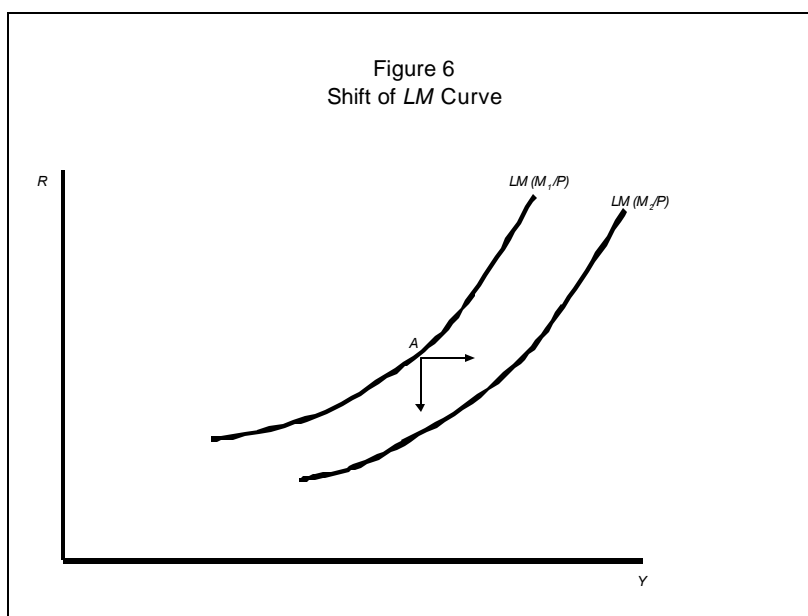
The size of the slope depends on the income and interest effects or elasticities of money demand. Of particular interest to economists is the interest elasticity of money demand. If the interest elasticity of money demand (the effect of a change in R on money demand) is very small, then L_R is small, and the slope is large and the LM curve is very steep. Similarly, a highly elastic or interest responsive money demand function will result in a relatively flat LM curve.

The slope and position of the LM curve can also be shown geometrically. To begin, assume that point A in Figure 5 is a combination of Y and R that results in money market equilibrium. To derive the LM curve, imagine that the level of output increases. Does point B also result in money market equilibrium? It cannot because the increase in Y increases money demand and the money supply is unchanged. Point B must be a situation of excess demand for money balances.



However, an increase in interest rates from the level at point *B* would lead economic agents to economize on their holding of the money asset and money demand would fall. If money demand is very responsive to changes in interest rates, a small increase in *R* would reduce money demand sufficiently to restore equilibrium. In that case point *C* would be an equilibrium and the *LM* curve is flat. Alternatively, if money demand does not respond very much to changes in interest rates, only a large increase in *R* would reestablish equilibrium. In this case the new equilibrium is at point *C'* and the *LM* curve is steep.

A similar geometric argument can be used to demonstrate the effect of change in the real money stock on the *LM* curve. Consider the situation at point *A* in Figure 6. With the initial real money supply M_1/P there is money market equilibrium at point *A*. If the real money supply increases (due to either an increase in *M* and/or a decline in *P*), will *A* still be an equilibrium? The answer is clearly no. With a higher real money stock, there will be excess supply of money at *A*. However, at some higher levels of output and/or lower levels of interest rates, money demand will be sufficiently increased to reestablish equilibrium. For the higher money supply (M_2/P) there is a new *LM* curve to the right of the original one.



Equilibrium Adjustments

Our discussion of the money market model has so far omitted one important issue—the adjustment mechanism. For the goods market we started the analysis with the adjustment mechanism—the quantity adjustment paradigm. It showed that if output and planned aggregate demand were not equal, then the production decisions of producers would bring the economy towards equilibrium. We have not introduced the adjustment mechanism for the money market. If the demand and supply for money are not in equilibrium, what adjustments occur and do they bring the market towards equilibrium?

Our approach to disequilibrium adjustments in the money market follows the Keynesian view that money is a financial asset, albeit the one used for transactions. Thus, a disequilibrium between the supply of and the demand for money will lead to changes in the holdings of other financial assets.

To see how such portfolio adjustments occur and affect the economy, consider that the demand for money at existing output and interest rate levels is greater than the real money balances held. That is, existing money holdings are lower than the desired level. Therefore, economic units will attempt to increase their

money holdings. In the Keynesian scheme this is done by selling other financial assets. As bonds are sold, their prices fall and market interest rates increase.² Thus, the attempt to replenish money holdings leads to an increase in interest rates and a movement toward the *LM* curve. Similarly, when there is an excess supply of money, bond purchases lead to lower rates.

Imagine that an increase in the level of economic activity leads to an increase in the demand for money for transactions and that monetary policy has not changed the available money stock. There is an excess money demand. Economic agents would like to maintain higher money balances. How can an individual increase his money holding? Individual agents will increase their money holding by selling other financial assets, or in other words, shifting part of their portfolios from bonds and deposits to money assets. However, as they sell bonds, the price of bonds goes down and the yield on existing bonds increases. In other words, interest rates go up. As interest rates increase, money demand falls and ultimately equilibrium is reestablished.

To summarize, the excess money demand sets in motion financial asset sales as individuals try to re-equilibrate their portfolios. These bond sales cause interest rates to rise which reduces money demand. Thus, interest rate adjustments are the mechanism that equilibrate the money market.

The approach presented here is that interest rate changes are the equilibrium adjustment mechanism in the money market. However, there are other adjustment processes that are possible. For example, in the case of an excess supply of money, we postulated that economic agents will rebalance their financial asset portfolios by buying bonds which leads to a change in interest rates. There are other ways that economic agents might respond to an excess supply of money. If I discover that my money holding exceeds my demand, I might respond by buying physical assets or by increasing my consumption expenditures. That is, the money market disequilibrium might have direct effects on the goods market. This possibility is emphasized by many monetarists and we will explore its implications later on. First, we will see how the *IS-LM* model with the Keynesian adjustment mechanisms works.

SUMMARY OF THE *IS-LM* MODEL

Before we proceed we can summarize the main features of the *IS* and *LM* curves:

***IS* Curve**

- Has a negative (downward) slope
- Is steep when PAD does not vary much with interest rates (investment is interest-inelastic)
- There is excess demand (supply) in the goods market to the left (right) of the *IS* curve.
- If there is excess demand or supply, horizontal movement toward the *IS* curve occurs as quantity adjustments eliminate the disequilibrium in the goods market.
- Government taxation and expenditure (fiscal) policy influences the position of the *IS* curve.
- Shifts outward (rightward) when there is an autonomous increase in PAD.

***LM* Curve**

- Has a positive (upward) slope.
- Is steep when money demand does not vary much with interest (rates demand is interest-inelastic).

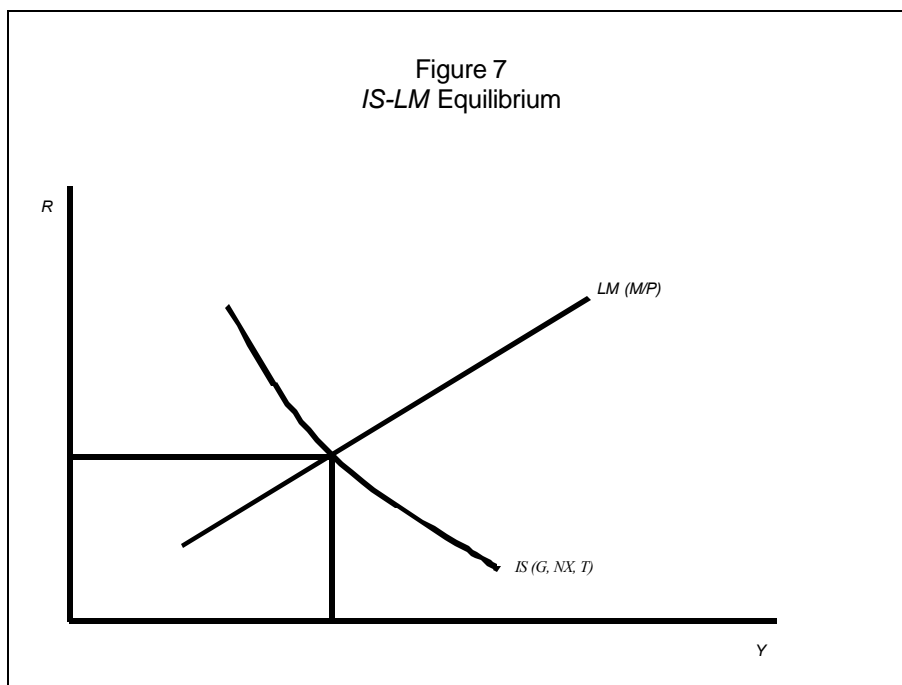
² Remember that bond prices fall when interest rates increase. Suppose that initially bonds are issued for \$100 and that each bond carries a \$10 annual interest coupon. The rate of interest on these bonds is 10 percent. In the above discussion market pressures push the price of bonds down, say, to \$90. The bond still pays a \$10 annual coupon. However, the purchaser of a bond now receives a \$10 annual return on a \$90 investment. The market interest rate on these bonds is $10/90 = 0.111$, or 11.1 percent.

- There is excess demand (supply) in money market to the right (left) of the LM curve.
- If there is excess demand or supply in the money market, vertical movement toward the curve occurs as interest rates change to eliminate the disequilibrium.
- Monetary policy operates through the real money stock (M/P) to position the LM curve.
- Shifts outward (rightward) when there is an autonomous increase in the real money stock (M/P).

ANALYSIS OF MONETARY AND FISCAL POLICY

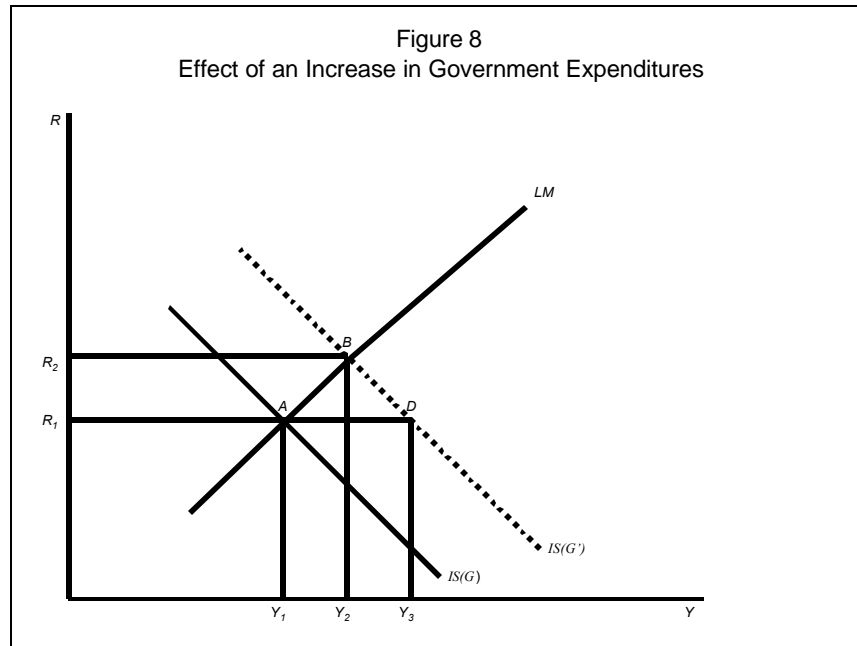
Probably the most important application of the $IS-LM$ model is for the analysis of the effects of monetary and fiscal policies. The purpose of this section is to examine the effects of policy changes with the model.

An $IS-LM$ equilibrium is shown in Figure 7. As noted in parentheses on the diagram the positions of the curves are fixed by the variables that are determined outside the model. The IS and LM curves are drawn for particular levels of the exogenous variables. The LM curve is a locus of money market equilibrium for a given level of the real money supply (M/P). Similarly, the IS curve is drawn for given levels of the fiscal policy variables, G and T , and NX . This is an important point because if any one of these variables changes, the IS or LM curve will shift and the equilibrium will change.



Fiscal Policy

We discussed earlier how a change in an exogenous component of planned aggregate demand will shift the IS curve. Consider now a fiscal policy expansion; the level of government expenditures increases from G to G' . The effects of this change are illustrated in Figure 8. The initial $IS-LM$ equilibrium is at point A . The increase in government expenditures shifts the IS curve and the new equilibrium is at point B .



The increase in output and interest rates that occur are the consequence of some important economic process and not just curve shifts. We need to examine these processes closely. The increase in government expenditure which shifts the IS curve causes income to increase because of a production response to increased demand (the multiplier process). Increased income (and transactions) increases the demand for money, and interest rates rise from R_1 and R_2 in order to maintain money market equilibrium.

Higher interest rates reduce investment and aggregate demand and constrain the multiplier-induced increases in output. Thus, the new equilibrium level of output is Y_2 . This is a Keynesian explanation of the channels by which fiscal policy affects the economy.

The response of the economy to the fiscal expansion is an example of the multiplier response. However, the simple model excluded any effects on the demand for money and feedback from the financial sector to the real sector. The simple Keynesian multiplier response would result in an output level of Y_3 in Figure 8. However, the interaction with the financial sector leads instead to an equilibrium at Y_2 because the multiplier expansion is held back by the increase in the interest rate.

The extent to which fiscal policy changes the equilibrium level of output depends on several factors:

- 1) The responsiveness of the demand for money to interest rates;
- 2) The responsiveness of investment demand to interest rates;
- 3) The magnitude of the expenditure multiplier.

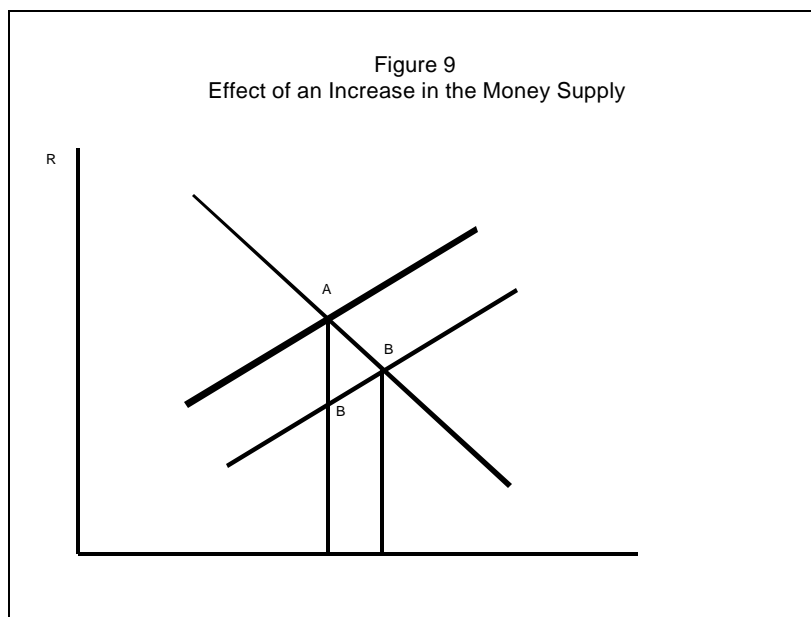
These factors affect the slopes of the IS and LM curves and determine whether a given stimulus will lead to a large or a small change in output. We will return to this issue later.

Another issue for later discussion is the length of time that equilibrium adjustments take to occur. That is, when fiscal policy changes, is the new equilibrium reached in 1 quarter, 1 year, or, 10 years? Finally, the analysis does not describe the precise path followed by the economy as it moves from point A to point B in Figure 8.

When fiscal policy changes and the multiplier process begins, it is likely that the demand for money will respond very quickly. However, the impact of a change in interest rates on investment may take a long time. Thus, the real goods market may respond to the stimulus from the increase in government expenditure and the rise in interest rates before investment is affected. A likely scenario may involve a movement from *A* to *D* and then to *B* in Figure 8. Other paths are also plausible, and distinguishing among them is very important for both policy analysis and forecasting. We must rely on empirical studies with econometric models to fill in the details.

Monetary Policy

We can now also describe the channels of the effects of monetary policy in the Keynesian world. An increase in the money supply will shift the *LM* curve to the right (see Figure 9), and the equilibrium point will shift from *A* to *B*. Once again we need an economic analysis to tell us why this movement occurs. An increase in the money supply creates an excess supply of money. With money supply *M'* and the economy at point *A*, economic units find that they are holding more money than they would like to. In order to adjust their money holdings, they purchase other financial assets. The price of bonds goes up, or equivalently, market interest rates decline. This in turn induces investment demand and an expansion of output from *Y₁* to *Y₂* through the multiplier process.



A likely sequence of these events is given by the path from *A* to *C* to *B*. This is based on the notion that the money markets equilibrate quickly when the money supply expands. Interest rates fall, and over time this leads to a gradual increase in investment expenditure and a multiplier expansion. The expansion of output increases money demand, which causes interest rates to rise.

Special Cases

First, consider the case where planned aggregate demand is unaffected by interest rates, and investment is interest inelastic. The effect on output of a fiscal expansion ($\frac{dY}{dG}$) is large. The logic of this is simple.

A fiscal expansion that starts the multiplier process going also leads to increased money demand and higher interest rates. However, if the higher interest rate does not affect PAD, the multiplier process is

not restrained. This is the case of a vertical IS curve. It is called the fiscalist case because it leads to a preference for fiscal policy over monetary policy.

It is also clear that the multiplier for a money supply change on output $\frac{dY}{d(M/P)}$ is zero when aggregate demand is unaffected by interest rates. The reason for this is that the channel that connects the monetary sector to the real sector is inoperative. An expansion of the money supply leads to lower interest rates that under normal circumstances effect PAD. That is the Keynesian channel for the influence of monetary policy. In this special case, the change in interest rates has no effect on the real sector so the monetary expansion is impotent.

If my empirical judgment is that PAD is relatively unaffected by interest rates, then I will view fiscal policy changes as very important and monetary policy of little consequence. Many of the Depression era Keynesian economists had this view. Their depression era experiences led them to conclude that PAD was interest inelastic. As a result, the Keynesians had a preference for fiscal policy over monetary policy. This preference was not due to any theoretical issue but instead was a consequence of the empirical judgment about the strength of the interest rate effect on aggregate demand.

The second case to consider is when money demand is interest inelastic. In this case, the monetary policy effect on output is very large and the fiscal policy effect on output is zero. Also, the reader should check why this corresponds to a vertical LM curve.

This case is called the monetarist case. Many economists who take a monetarist approach emphasize the transactions role of money and importance of the quantity of money available. They place little importance on the relationship between money demand and interest rates. In this case, the relationship is non-existent. As a consequence, monetary policy has an important effect on output and fiscal policy has none. Monetarist economists often emphasize monetary policy and eschew fiscal policy. They have other reasons for this view, which we will discuss later.

THE KEYNESIAN MODEL IN THE OPEN ECONOMY

The Keynesian model developed so far provides an analysis of equilibrium in the real sector (goods market) and the financial sector (money market). However, we have so far omitted any discussion of the foreign sector and the open economy implications of the model. This is an important issue to examine because the most significant structural change in the American economy in the last generation has been its internationalization. In this section we will examine the relationship between the domestic $IS-LM$ Keynesian equilibrium and the foreign sector.

Exchange Rates

A major influence of the foreign sector on aggregate demand is through the influence of exchange rates on planned aggregate demand. This relationship was already discussed in the previous chapter.

For the first 25 years after World War II, the major world currencies were on a fixed exchange rate regime. Exchange rates were set by international agreement and only occasional adjustments were made to the fixed rates. The rates were maintained via central bank intervention in the foreign exchange markets. The International Monetary Fund, which was established in the aftermath of the Second World War, provided an international lending facility which enabled the central banks to intervene and maintain the fixed exchange rates.

As the volume of international trade and financial flows grew rapidly, it became increasingly difficult to maintain fixed rates. The necessary interventions were often beyond the financing capabilities of the central banks. In addition, differences in growth rates and inflation rates among countries resulted in a

need for more frequent realignments of the fixed rates. As a consequence, the fixed exchange rate regime began to falter. By 1973, the major economies had moved to a floating exchange rate regime.

Exchange rates are now determined by market forces. Nevertheless, there are central bank efforts to intervene in the foreign exchange markets. The Federal Reserve and other central banks do enter the markets to effect the exchange rate. However, at this time, central bank interventions are probably not the major determinant of the exchange rate. The exchange rate is largely determined by the influences of supply and demand.

An appreciation (an increase in the exchange rate) will lead to an increase in real expenditures on imports and a decrease in real exports. Foreign goods will decline in relative price for Americans and the price of U.S. goods to foreigners will be relatively higher. Since net exports are exports less imports, we can write a function for net exports as:

$$NX = NX(Y,e)$$

The other major influence on planned net exports is U.S. income since import demand is highly income elastic. Import demand increases with income.

We showed earlier that the PAD curve will shift when the exchange rate changes. It is also the case that the *IS* curve shifts with a change in *e*. A dollar depreciation will increase net exports and shift the *IS* curve to the right.

Monetary Policy in an Open Economy

In this section we will assume that the economy starts with a domestic and international equilibrium, and we will examine the open economy implications of monetary and fiscal policy changes. That is, we will consider the exchange rate and international market implications of a change in domestic policy.

We know that an expansionary monetary policy will tend to increase output (*Y*) and reduce interest rates (*R*). What are the open economy implications of these changes?

As output and income increase, the demand for imports will go up. This effect is important because many imported consumer goods are highly income elastic. The demand for foreign exchange will lead to a depreciation of the dollar. Similarly, the fall in interest rates makes dollar denominated financial assets less attractive. If asset holders sell dollars and buy foreign financial assets, then the dollar depreciates.

In summary, an expansionary monetary policy puts downward pressure on the exchange rate. In a world of fixed exchange rates, this hampers the ability of the central bank to follow an expansionary monetary policy. Even in a flexible exchange rate regime, the central bank might be inhibited by concern with the exchange rate. It might avoid an expansionary monetary policy if it views a currency depreciation to be undesirable.

If a currency depreciation is undesirable, the central bank may hesitate to adopt an expansionary monetary policy even when domestic economic conditions, such as a recession and excess capacity, suggest that a looser policy is appropriate. Depreciation is undesirable because it increases the costs of important raw material imports and goods without any domestic substitutes. Thus, a currency depreciation leads to imported inflation even when there is no danger of excess demand inflation domestically.

We should not assume that in every instance a looser monetary policy causes a currency depreciation. It is notoriously difficult to generalize about exchange rate movements. They are often dominated by the tastes and expectations of asset holders around the world. A case in point occurred in 1990–1991. The Federal Reserve took distinct steps towards a looser monetary policy throughout the latter half of 1990. By early 1991 short-term interest rates in the U.S. were about 200 points lower than in the summer of

1990. The decline in interest rates should lead to a depreciation of the dollar. At the same time, the recession led to a weakening of import demand, but this effect was rather small and could not have put much upward pressure on exchange rates. Thus, it was rather surprising that in the first three months of 1991, the dollar increased in value by about 15% against most major currencies. In this episode exchange rate movements were determined by changing attitudes and expectations that were related to the military and political events in the Gulf. The economic fundamentals took the back seat.

Fiscal Policy in an Open Economy

A simple *IS-LM* analysis tells us that an expansionary fiscal policy leads to higher interest rates and a higher level of output. What are the open economy implications of these changes?

To begin, the expansion of output and income leads to an increase in imports. A negative trade balance will put downward pressure on the exchange rate.

At the same time, the increase in interest rates makes dollar denominated financial assets more attractive and leads to a dollar appreciation. The fiscal policy expansion sets in motion forces for both depreciation and appreciation of the currency. Which is stronger?

A generation ago, the answer would have been that the trade balance was the dominant determinant of exchange rate movements. At that time, international financial markets were not well integrated. Markets in different countries were segmented or separated from one another so differences in interest rates would not lead to large portfolio adjustments. The demand for foreign exchange was largely driven by the trade balance.

At the present time, the trade balance is small in comparison to the cross-border financial flows that occur daily. Although the U.S. trade deficit exceeded \$150 billion in 1987, this figure is small relative to the annual volume of foreign exchange trading that occurs in the world markets. In the short-run, the trade deficit is not large enough to exert much influence in the market. Of course, the trade deficit does affect the expectations and tastes of the foreign exchange traders. Short-run changes in the demand for foreign currency due to trade flows are simply not large enough to dominate other factors in the exchange market.

THE NEXT STEP: BACK TO LONG-RUN EQUILIBRIUM

The entire discussion of the full Keynesian equilibrium has been conducted within the quantity adjustment paradigm. However, even the most ardent Keynesian would acknowledge that price adjustments can occur, particularly when the *IS-LM* equilibrium is not at the overall long-run equilibrium level of output, Y^* . Specifically, if productive capacity is fully utilized and labor fully employed, then an increase in demand will inevitably lead to inflation.

The depression era Keynesians did not worry about such a situation since in their experience output was far below a capacity level. However, there have been large periods in the postwar decades when the economy was operating at or near full capacity. Furthermore, capacity output is really a broad range. Some industries and sectors begin to hit the constraints before others.

All in all, it is clear that in the modern macroeconomy, price adjustments to changes in demand are as important as quantity adjustments. Therefore, our next step in building a framework for macroeconomics is to incorporate price adjustments in our model and to develop an explanation of inflation. This step will help us link the long-run equilibrium concepts presented in Chapter II to the Keynesian short-run equilibrium discussed here.

CHAPTER IV INFLATION AND THE COMPLETE MACROECONOMIC MODEL

The last Chapter presented the Keynesian model of the real and monetary sectors of the economy. It showed how real sector quantity adjustments interact with the demand for money via the role of interest rates. The model determines the equilibrium values of real output and interest rates. The Keynesian approach is useful because it shows how the real and monetary sectors are affected by policy or any other shock to equilibrium. However, the main drawback of the model is that it retains the assumption that the only relevant real sector adjustments are quantities produced; prices were assumed unchanged for the entire discussion.

The assumption of fixed or sticky prices is in fact defensible when we restrict ourselves to very short periods of time, less than a year or so. However, any shock to the economy that is likely to have effects that extend beyond six months or a year is also going to have implications for price behavior. We will now extend the model so that we can discuss price behavior and inflation.

We begin this Chapter with an extension of the Keynesian model that allows us to discuss price flexibility and relate the Keynesian demand equilibrium (the topic of Chapter III) to the long-run neoclassical equilibrium (the topic of Chapter II). This analysis is called total supply and total demand analysis and it relates the level of output to the price level.

In the second half of this Chapter, we will take a more general view of price determination and develop an explanation of inflation—the rate of price change. Since most everyone thinks of the inflation rate as the relevant macroeconomic variable, this discussion will be used to discuss policy.

TOTAL DEMAND AND SUPPLY ANALYSIS

The Keynesian *IS-LM* model developed earlier is a model of only demand behavior. It tells us how the equilibrium between planned aggregate demand and output is achieved. It describes demand behavior but says absolutely nothing about supply behavior. Throughout the discussion of PAD in Chapter III, we always assumed (albeit implicitly) that the supply of output would and could adjust. Total supply and demand analysis breaks that implicit assumption and introduces the possibility of supply constraints or economic issues that can affect supply behavior.

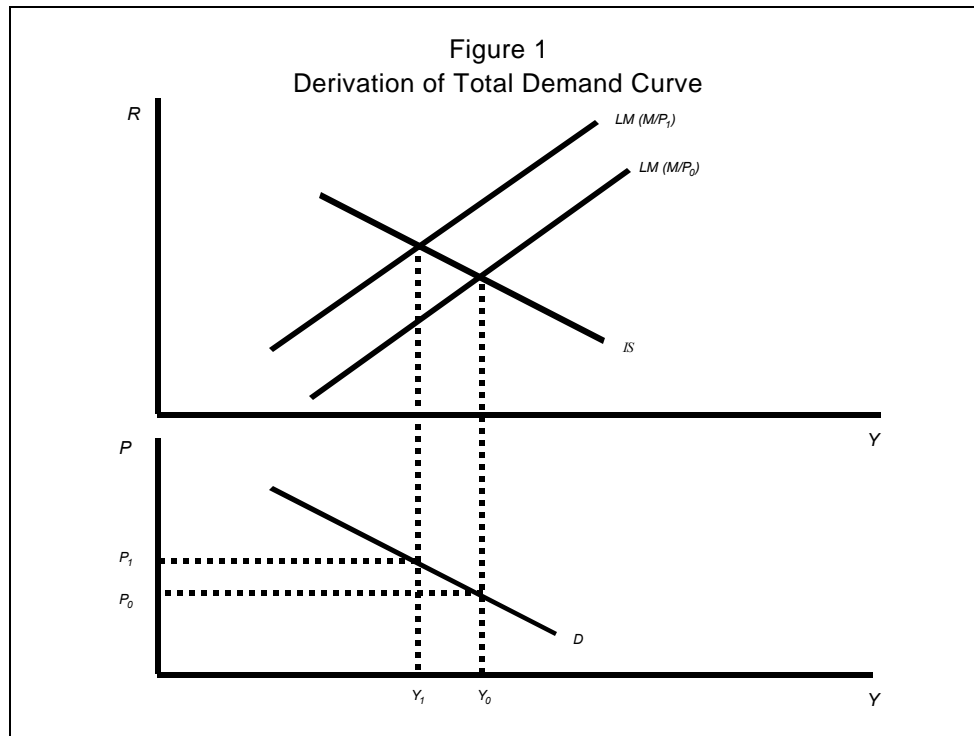
The total demand curve shows how the Keynesian equilibrium changes for different values of the price level. The total demand curve shows how changes in the price level affects the *IS-LM* or demand side equilibrium. We begin by deriving a total demand curve from the *IS-LM* analysis. We then turn to the determination of supply behavior and derive different approaches to the total supply curve.

Total Demand Curve

The price level is an exogenous variable in the Keynesian aggregate demand model. That is, it is not determined by the system, although it does appear in the model structure. The price level determines the real value, or purchasing power, of the nominal money supply, thus positioning the *LM* curve and determining the aggregate demand equilibrium. The position of the *LM* curve will change if prices change, and as it does, the output equilibrium changes as well. That is, in the *IS-LM* model there will be a different output equilibrium for each level of prices. The total demand curve summarizes this relationship between the price level and the output level determined by the intersection of the *IS* and *LM* curves.

The total demand curve is a locus of Keynesian aggregate demand equilibrium for different price levels. If the price level changes while everything else (including the nominal money supply, *M*) remains the

same, then the resulting change in the real money supply ($\frac{M}{P}$) causes the *IS-LM* equilibrium to change. To be more precise, consider the situation shown in the upper panel of Figure 1. For the price level P_0 , the *LM* curve is given by $LM(\frac{M}{P_0})$ and output is Y_0 . If the price level increases to P_1 , the real value of the money supply will be reduced. An increase in the price level with a given nominal money supply is equivalent to a contractionary monetary policy; it reduces the real money supply. The *LM* curve is given by $LM(\frac{M}{P_1})$ and the demand side output equilibrium is Y_1 . Similarly, for each value of the price level there will be a different aggregate demand equilibrium. Since a higher price level contracts the real money supply, the output equilibrium is lower when the price level increases.



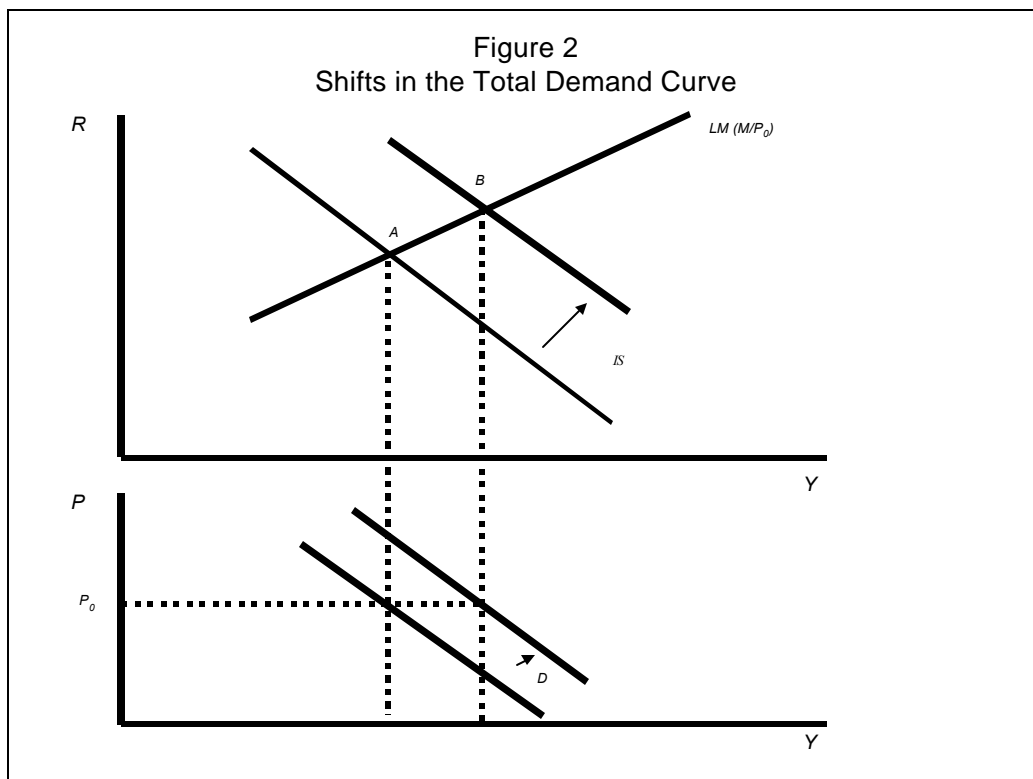
The lower panel of Figure 1 shows the equilibrium level of output that corresponds to each price level. The curve labeled *D* is a locus of *IS-LM*, or aggregate demand, equilibria. It is called the total demand curve because it summarizes equilibrium on the demand side of the economy.

The negative slope of the total demand curve follows from the derivation shown in Figure 1 and also agrees with one's intuition. When prices increase, nominal income also increases and the demand for money to be used for transactions goes up. If the nominal money supply is unchanged, there will be a shortage of transactions balances. This "tightness" in financial markets and/or the interest rate increases that will result restrain aggregate demand. The price increase is equivalent to a tighter money policy and similarly leads to a fall in the demand for output.

Another channel that relates the aggregate demand equilibrium to the price level is the foreign sector. An increase in the domestic price level tends to encourage the demand for imported goods and discourage exports. Thus an increase in prices reduces net exports and restrains aggregate demand. Consequently, a higher price level is associated with a lower output equilibrium.

The total demand curve in Figure 1 is drawn for given values of all the exogenous variables in the *IS-LM* model. A change in the level of government expenditure, taxes, or the nominal money supply will affect the aggregate demand equilibrium for every price level and thereby affect the position of the total demand curve. The total demand curve describes the amount of output that people are willing to buy at different price levels, but the levels of output and price that will actually emerge also depend on supply behavior.

The total demand curve shifts when an exogenous shock such as a policy change effects the *IS-LM* equilibrium. For example, consider an aggregate demand shock that shifts the *IS* curve to the right. For any price level that fixes the position of the *LM* curve, the demand equilibrium will be at a higher output level. Thus, the exogenous increase in demand shifts the total demand curve to the right. This particular case is shown in Figure 2. Point *A* is an *IS-LM* equilibrium for price P_0 . If there is an exogenous shift in aggregate demand, the *IS* curve shifts and the equilibrium for the same price level is at point *B*. Points *A* and *B* correspond to two points on separate total demand curves as shown in the bottom panel.



Total Supply Curve

A new concept in our discussions—the **total supply curve**—is both the more interesting and the more problematic aspect of this analysis. It describes the amount of output that producers are willing and able to supply to the goods market. We will have several different ways of looking at supply behavior.

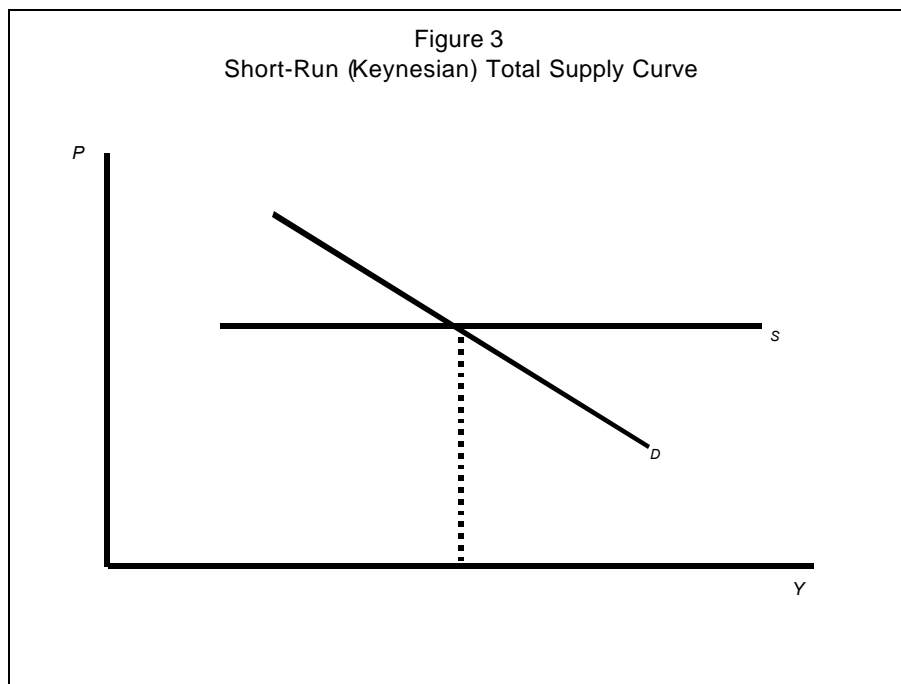
Keynesian Supply Curve

The total supply curve implicit in the Keynesian *IS-LM* model is based on the notion that there are no supply constraints and that prices are pre-determined in the short-run (one year or less). Thus, whatever output level is demanded will be produced and the total supply curve is a horizontal line. There is sufficient excess capacity so that an increase in demand leads to more production without increasing production costs and prices. For this reason the early Keynesian economists who were schooled by the

experiences of the depression used *IS-LM* analysis exclusively because they thought in terms of situations with a great deal of excess productive capacity. In this framework, we can view the price level as being set by existing contractual arrangements such as union contracts, sales agreements, and price lists. It is assumed that any level of output can be supplied at this given level of prices.

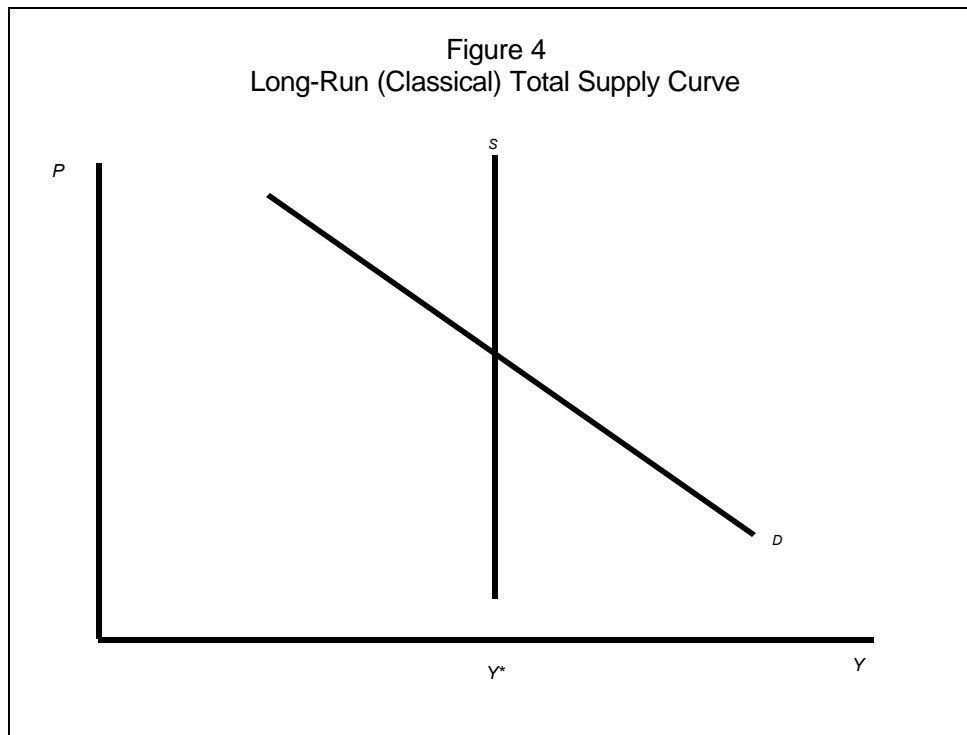
Figure 3 shows the horizontal total supply curve, which is labeled *S*. The output equilibrium is strictly determined by the *IS-LM* equilibrium, which positions the total demand curve, *D*.

The lack of realism in the Keynesian demand model quickly becomes clear when we consider Figure 3. The horizontal total supply curve (implicit in the simple Keynesian approach) implies that supply behavior is completely unaffected by the level of output or demand. A large shift outward in the *IS* or *LM* curves leads to a shift out in the *D* curve. The real and monetary sector interactions occur without any implications for prices, no matter how large the demand shift happens to be. Now, the horizontal supply curve may be appropriate when we are just considering very short-run periods (surely no more than 6 to 12 months) and when the output level is sufficiently far away from any capacity constraints. Over longer periods of time, we must consider whether producers will change their pricing decisions. In addition, if output increases, it is necessary to ask whether capacity constraints in particular industries are likely to lead to price adjustments. The higher the output level, the quicker and more likely such price adjustments will be.



Long-Run or Classical Supply Curve

At the opposite extreme to the Keynesian short-run horizontal supply curve lies the supply curve implicit in the long-run equilibrium or classical view of the macroeconomic world that was discussed in Chapter II. The classical view implies a vertical supply curve, as shown in Figure 4.



The classical view of macroeconomics is rooted in the idea that the macroeconomy is the aggregate of an infinite number of perfectly competitive markets. In this view each and every market for outputs and inputs reaches an equilibrium which determines both the relative price and the quantity for that market. The level of output supplied is simply the aggregate of all these outcomes for any overall price level. This is the case because each and every market equilibrium determines the relative price of the good in that market. As long as relative prices stay in equilibrium, the same level of total output will be supplied. A change in the aggregate price level does not disturb the relative price relationships between all pairs of goods. Thus, the total supply curve is vertical at this equilibrium output level no matter what the aggregate price level happens to be. This is the same equilibrium output level, Y^* , that we discussed in Chapter II.

The vertical classical total supply curve can be understood if we imagine that the aggregate price level doubles. If every price doubles in terms of the money unit of account in the economy, the aggregate price level doubles as well. Moreover, no relative price between pairs of goods changes, and thus the equilibrium level of output supplied remains unchanged. Therefore, the total supply curve in the classical model is vertical.

The classical vertical total supply curve and the Keynesian horizontal total supply curve represent two theoretical extremes, neither of which is a satisfactory representation of behavior in the real world. The traditional Keynesian approach leaves us without a theory of price determination. The classical approach introduces a theory of price determination, but at the cost of eliminating an explanation of fluctuations in real output. By assuming that competitive markets at all times generate equilibrium levels of output, the model cavalierly does away with fluctuations in output.

A more appropriate view of the total supply curve will be the middle road—a positively sloped total supply curve. In Chapter II, we used the quantity adjustment paradigm as an argument for very short-run price stickiness. There are some additional reasons why price adjustments occur slowly so that over the

medium-run the total supply curve has a positive slope. Such an approach is relevant for adjustments that occur for longer periods than the Keynesian short-run period (the 6 to 12 month period where quantity adjustments are dominant) and less than the long run (the period of several years after which the long-run equilibrium approach is dominant). We will begin with some additional reasons why prices are slow to adjust in the medium term.

Some More Reasons Why Prices are Sticky

Labor Contracts Many readers of this note have jobs and have had the experience of negotiating a price for their labor services—a wage rate or salary. However, these negotiations are unlikely to occur anew every morning at the start of the workday. In all likelihood, the negotiations take place about once a year. We observe that throughout the economy wages are usually set for relatively long periods. Formal wage contracts frequently cover a two or three year period. Informal wage setting agreements usually stipulate an annual review or renegotiation. Why is this the case? First, daily negotiations would place a costly burden in terms of time and effort on both employers and employees. Most everyone agrees that such frequent wage setting behavior would be a waste of resources. Second, employees often value some certainty about their wage and employment situation which longer-term agreements bring.

Implicit labor contracts is a term that refers to the type of agreements that are often made between employers and employees. Implicit contract theory stems from the observation that firms often do not vary employment levels when demand shifts. Instead, employees have implicit job guarantees at fixed wages. Such arrangements shift the risk of economic fluctuations to employers. Employees get some certainty about their job security and employers get a contractual commitment to pay a fixed wage. (Thus, the employer does not have to pay premium wages in boom times to retain workers.) Of course, these arrangements are not always permanent and layoffs do occur in recession periods. Nevertheless, the implicit contract arrangements seem to be common and seem to be satisfying to employers and employees alike.

The major implication of implicit contractual agreements is that they provided us another reason why wages do not vary very often. In fact, wages are somewhat unresponsive to changes in demand conditions.

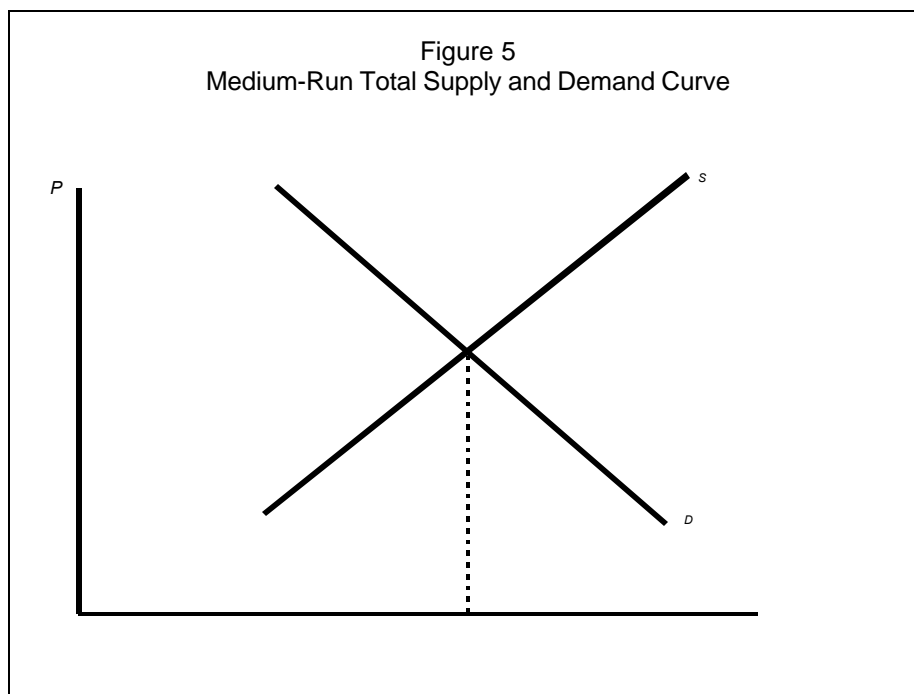
Career Labor Markets The career labor markets view is another reason why wage rates do not vary very often. The idea here is that employers adopt policies that promote the long-run attachment of workers to the firm. This is important to employers because finding able workers and providing training can be expensive. In addition, employees find it in their interest to agree to such arrangements. Some examples will show how career labor market arrangements work and why they lead to slow or sticky wage adjustments.

To begin, consider labor contracts that are often made with highly experienced and valued workers. It is common to pay experienced workers more than they could earn in another job in order to decrease the probability that a firm's experienced workers will leave. Thus, increased labor market tightness puts only minimal pressure on the wage rate because experienced workers, who are already earning a premium, do not leave for alternative positions. Such compensation arrangements are feasible for two reasons. First, part of the compensation that is available to experienced workers is often only available to workers with experience and will often disappear if the worker leaves. Examples of such forms of compensation are extra pay for seniority, pension rights, and additional vacation time that accrues with seniority. Second, experienced workers are compensated for their additional skills, which may be largely specific to a particular firm or job. They would lose the compensation for any job-specific skills when they change jobs. There is an additional reason for the allegiance of experienced workers to an existing job. That is, a job change imposes both real and psychological costs on the worker. These include the costs of job search and relocation and the psychological burden imposed by a change in work environment. People are typically reluctant to change jobs because of the uncertainties associated with a new work place, a new boss, and new coworkers.

Next, consider wage rates for entry-level jobs. Firms often maintain relatively high entry-level wages so there is always a normal queue of new job seekers. Thus, when demand conditions dictate an expansion of employment there is a queue of job seekers that can be tapped without increasing wage rates. With high entry-level wages, the human resources department will always have a pool of able workers available for hire.

In both instances, wage setting for experienced workers and wage rates for entry-level positions, specific institutional arrangements tend to emerge in many firms over time. Our quick analysis shows why such arrangements are attractive. The important macroeconomic implication in both instances is that wages do not change very often. Wages are sticky; they do not respond very quickly to changes in demand and supply conditions. Note that the conclusion that wages are sticky in the short and medium run does not imply that they are unchanging.

The earlier discussion (in Chapter III) of the quantity adjustment paradigm suggested that goods prices (at least for non-perishables) are fairly constant in the short-run. In fact, the Keynesian assumption of constant prices is applicable for short-run periods of up to 6–12 months. This discussion of wage determination indicates that wage rates are very slow to adjust to changing macroeconomic conditions for periods of time that extend beyond a year. As a consequence, for medium run time periods (say, about 6 months to several years), price adjustments occur but they occur slowly. Thus, the total supply curve should be drawn with a positive slope that indicates gradual upward pressure on prices when output exceeds the long-run equilibrium and vice versa. Such a situation is shown in Figure 5. In the next section, we discuss the effect of shock or disturbance on the macroeconomic supply and demand equilibrium.



A DEMAND SHOCK

Consider a disturbance or shock to equilibrium that originates on the demand side of the economy—a demand shock. Specifically, we will consider the case of an expansionary monetary policy that shifts the total demand curve to the right. What happens after the demand curve shifts—first, in the short-run, then in the medium-run and the long run?

For the short-run, we can consider the Keynesian analysis of Chapter II. The monetary policy expansion leads to a fall in interest rates which gets the multiplier process under way and output increases. Since we are no longer restricted to a fixed price, short-term adjustment situation, there are other effects to consider. Over time, the multiplier process that leads to an increase in output also leads to increases in prices. Why? As output increases some markets (for some goods or raw materials) begin to encounter bottlenecks or capacity restrictions. Quantity adjustments become less common and price adjustments become more common. Nevertheless, additional output is forthcoming, and after a period of several years, there are both output and price increases.

As prices rise, real money balances decline. Thus, the price increases offset to some extent the initial monetary policy expansion.

If bottlenecks, shortages, capacity constraints, etc., come into the picture rapidly when the multiplier expansion gets underway, the price increases caused by the expansionary policy may be quite large. The fall in real balances can then have a serious constraining influence, and the policy might have led to little growth in output. On the other hand, if there is plenty of excess capacity in the economy and producers can readily respond to increases in demand by increasing output, the new equilibrium can involve a healthy multiplier response with little upward pressure on prices.

The difference between an expansionary monetary policy that has a substantial effect on output and little effect on prices and one that has just the opposite effects depends on the situation. If the initial economic conditions were one where there was plenty of excess capacity in the economy, the former description may be more appropriate. If the economy were at its long-run equilibrium level of output to begin with, the policy expansion would lead to only a temporary increase in output. Price pressure would continue as long as output exceeded the equilibrium level (by definition there is excess demand pressure in such a situation). Continued upward pressures on prices would lead to a contraction of real money balances, which bring the level of output back to the long-run equilibrium. However, such long-run adjustments can take place slowly—over a period of as much as 3 to 5 years.

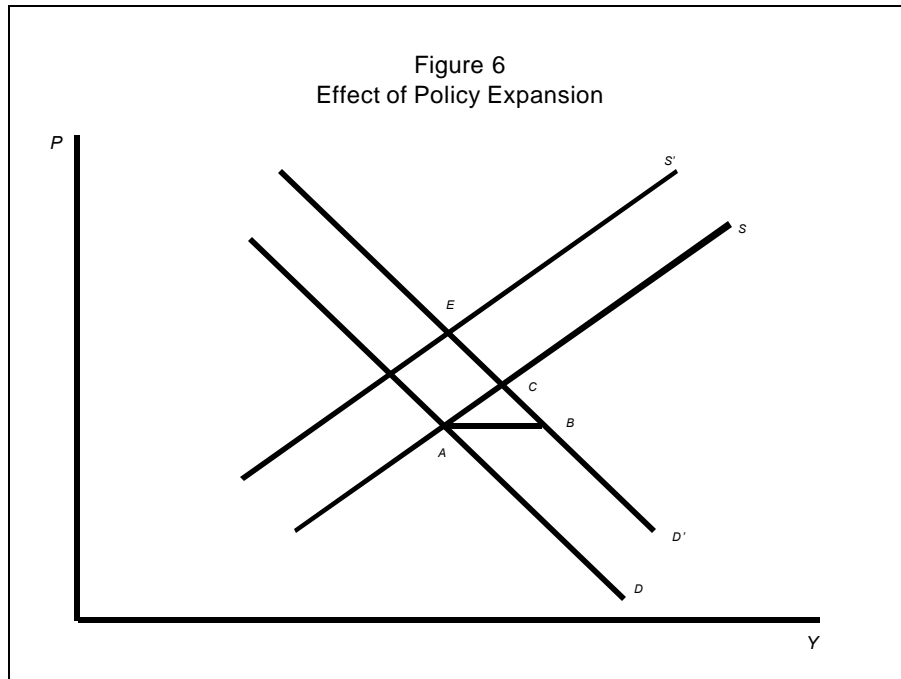
A Policy Expansion: Step by Step

Macroeconomic adjustments in the short-, medium- and long-runs can be quite complicated. In this section, we review the short-run Keynesian adjustments (from Chapter III), the medium-term adjustments introduced here and the long-term equilibrium (from Chapter II) with a step-by-step story.

Consider an expansionary monetary policy:

- The expansionary shock shifts the total demand curve in Figure 6 from D to D' . There is excess demand in the real sector, output begins to expand (the multiplier process) toward point B in Figure 6. This expansion begins within about six months and continues for two years or more.
- Within a year or so, price pressures begin to emerge because some markets begin to experience excess demand and price adjustments become more frequent. The price level begins to rise and the economy begins to move toward point C in Figure 6. The increase

in prices constrains the expansion because the real money supply falls as the price level increases.



Additional changes occur in the medium-run, which will have further implications on macroeconomic equilibrium:

- After a year or more, wage negotiations occur and wages begin to catch up with prices. As wages catch up, supply decisions begin to change. Output begins to fall as producers encounter higher labor costs. That is, over time we expect that new wage contracts will be negotiated at higher wage rates. With a higher price level, the equilibrium wage will also increase over time. As wages “catch-up” with prices, firms will no longer be willing to supply output in excess of the long-run equilibrium level. Such an effect on the output supply decisions of producers is reflected by a move back in the Supply curve to S' . The macroeconomic equilibrium is now point E in Figure 6.
- Price pressures will continue depending on the relationship between the level of output and the long-run equilibrium level of output, Y^* . If output is greater than the natural equilibrium level, Y^* , then some markets continue to experience excess demand pressures and price increases continue. As the price level continues to rise, $\left(\frac{M}{P}\right)$ falls and causes the level of output to move towards Y^* .
- This process goes on until the economy returns to a point of long-run equilibrium. Such long-run adjustments can take 3 years and can go on for substantially longer. Recall that the long-run equilibrium is the level of output where supply and demand in all individual markets balance so there are not pressures for change on either prices or production.

There are yet some additional adjustment issues to consider. Specifically, the whole process under discussion can lead to changes in the long-run equilibrium level of output, Y^* .

- The initial expansion in output from point *A* in Figure 6 resulted in several years of increased output and a possible change in the composition of output. Specifically, there would have been more investment expenditures than would have occurred otherwise. If additional investment occurs, then the capital stock will expand. This is important, because growth in the capital stock will shift the natural rate of output upward. Although the economy returns to a natural rate of output Y^* , several years of increased investment expenditure could have a significant impact on the level of real output associated with natural equilibrium.

Thus, the final element of the story is: The several years of output expansion leads to growth in the capital stock. Thus, there is a supply side effect which increases Y^* and can therefore lead to a long-run output effect of the expansionary shock. An issue of considerable controversy is whether these supply side effects on Y^* are very large. The supply side policy advisors in the early years of the Reagan administration argued that the shift in Y^* would be enormous. However, experience and most empirical studies suggest that supply side effects are much more modest. The idea reappeared in the 1996 Presidential campaign when Dole suggested tax cuts that were predicted to lead to enormous increases in the rate of growth of output. If the across-the-board tax cuts really could have enormous supply side effects on investment, capital stock, and growth, there would be additional government revenue and the tax cuts would not increase the deficit. The similarities between the Dole campaign and the rhetoric that came out of the Reagan White House in 1981 are striking and no more realistic this time around.

A major theoretical point of our discussion is the tendency to return to a long-run equilibrium—the normal, natural, or long-run equilibrium level of output termed Y^* . It is an output level associated with balance in the macroeconomy; in particular, there is an absence of inflationary or deflationary pressures at this output level. It is very important to add that this output need not be one where all resources are fully employed. It could very well be that modern economies begin to generate pervasive inflationary pressures when there are still considerable amounts of resources available. Thus, the natural equilibrium may not be a socially desirable goal for the economy.

The long-run equilibrium is “natural” because the economy tends to move toward it. The term does not convey a value judgment that this equilibrium is desirable or good. There may be more unemployment at Y^* than a democratic society would like to endure.

This possibility creates a serious quandary for decision-makers. If society sets as its policy goal an output level which exceeds Y^* , policy will have a built-in inflationary bias. Such a situation may have befallen the U.S. economy in the 1960s and 1970s. A frequently stated goal for the real sector at that time was an overall unemployment rate of about 4 or 5 percent. However, inflationary pressures exist in the U.S. economy unless the unemployment rate is higher than that. Most econometric estimates suggest that at that time the natural rate of unemployment (the unemployment rate associated with output at Y^*) was probably in the range of 6 or 7 percent. (We will discuss more recent estimates of the natural rate in the next section.) Thus, the policy goal for the unemployment rate for about two decades placed the output level about Y^* and helped generate an era of sustained inflation.

In order to explore the implications of inflation in more detail, we will develop a somewhat different approach to our theoretical discussion. That is, instead of total supply and demand analysis to determine output and the price level, we will develop an equation to explain the rate of inflation in the next section.

UNDERSTANDING INFLATION MOMENTUM

Since the inflation rate is probably the most closely followed macroeconomic phenomenon, it will be helpful to have a theoretical framework that concentrates on the determination of the inflation rate directly. Our inflation equation will also help us understand one of the most unusual characteristics of

inflation—its persistence. That is, we will discuss the momentum to inflation. For example, in 1982 when the unemployment rate approached 10%, it was clear that the level of output was far less than Y^* and had been so for well over a year. Nevertheless, the inflationary pressures that developed in the 1970s persisted with little sign of abatement. The momentum of inflation kept the inflation rate near 10% for another year before the output gap began to affect price setting behavior. The momentum of inflation has important consequences for policymaking. For example, in early 1990, it was clear, but not entirely certain, that the economy was approaching a downturn. Policy-makers were hesitant to respond with an easier policy stance because the inflation rate was accelerating at the same time. A price adjustment function will help us analyze the inflation process more closely.

Price Adjustment Function

The natural rate or long-run equilibrium level of output— Y^* —was defined (see Chapter II) as a point of overall market equilibrium. Consequently, at Y^* there would be no pressures for either quantity adjustment or price adjustments. Of course, there can be a short-run equilibrium between aggregate demand and output at a level of output different than Y^* . However, over the long run, the output level will tend back to Y^* . For example, if output exceeds Y^* , there will be upward pressures on prices.

In general, inflationary pressures emerge and increase with the size of the gap between Y and Y^* . In this section, we will examine inflation by developing a dynamic specification of price adjustments. We will replace the static total supply curve with a specification of a dynamic price adjustment function.

To begin, the price adjustment function relates the inflation rate to the gap between actual output and the long-run equilibrium level of output:

$$p = g(Y - Y^*) \quad (1)$$

where p is the inflation rate; the -1 subscript (e.g., P_{-1}) indicates that it is the value in the prior time period. An important implication of this specification of equation (1) is that when $Y = Y^*$, there are not inflationary pressures.

The price adjustment equation (1) reflects the influence of both supply and demand phenomenon on the inflation rate. As the level of output rises about Y^* , there will be more and more sectors of the economy where quantity adjustments give rise to price adjustments. As Y grows there will be fewer markets where quantity adjustments are feasible. Thus, the inflation rate increases with the gap between Y and Y^* .

Phillips Curve

The term Phillips curve refers to the empirical relationship between wage or price inflation and the unemployment rate. Since Phillips' early econometric studies in the 1950s, the relationship has been extended and developed into an important analytic tool for understanding the inflation process. It is common to present the Phillips curve with the unemployment rate instead of the gap between Y and Y^* . Just remember that as U goes up, $Y - Y^*$ goes down.

An empirical Phillips curve relationship is shown in Figure 7. It shows the relationship between the unemployment rate for all workers and the rate of price inflation (in the implicit price deflator for personal consumption expenditures) in the U.S. from 1950 to 1969.

In the 1960s policy-makers were confident that the relatively flat Phillips curve (such as the one shown in Figure 7) could be used to describe their options. That is, an expansionary policy would reduce unemployment but would also lead to a higher inflation rate. Policy-makers believed that the Phillips curve provided an estimate of the trade-off between unemployment and inflation. Typical of such discussion is the following from the *Economic Report of the President*, 1963 (p.84):

At present, considerable latitude exists in the American economy to increase output by bringing unemployed labor and unused capital back to work: this is a principal reason why a tax reduction is needed. While the record of the postwar years indicates that wages tend to rise more rapidly in years when unemployment is low, given the present high unemployment rate, demand for labor can expand substantially without resulting in much additional pressure on labor markets.

If the Phillips curve is flat, an expansionary policy can reduce unemployment without increasing inflationary pressures very much. Indeed, the unemployment rate averaged 5.5 percent in 1962–1964 and the inflation rate for personal consumption expenditures averaged 1.1 percent per year. A cornerstone of the Kennedy-Johnson economic program was a large reduction in personal taxes enacted in 1964. In 1966–1968 the average values were 3.7 percent for the unemployment rate and 2.8 percent for the inflation rate. The experience of the 1960s gives rather strong support for the idea of a Phillips trade-off:

- An expansionary policy could reduce the unemployment rate at the cost of only a small increase in the inflation rate.

The gradually sloped Phillips curve relationship shown in Figure 7 for the early postwar years disappears when we add more recent data. Figure 8 shows a scatter of data for unemployment and inflation rates for the period 1950–1987. These data indicate that the world is more complicated than the simple Phillips curve would suggest. Throughout the 1960s and 1970s, as the breakdown of the Phillips relationship became apparent, economists struggled with attempts to explain and understand the phenomenon.

In the late 1960s, contributions by Edmund Phelps and Milton Friedman led economists to focus on the role of the expected rate of inflation. In the next section we introduce the expectations augmented Phillips curve which will enable us to understand the relationship between unemployment and inflation shown in Figure 8 and also explain the persistence of inflation.

Expectations of Inflation

Economic agents who set prices are affected by two major influences. The first is the overall supply and demand conditions in the market and the second is the rate of inflation that these agents expect to prevail in the economy. The overall supply and demand conditions are measured in the Phillips curve analysis by the unemployment rate or the $Y - Y^*$ gap, either of which indicate the strength of economic activity and the extent of slack in the economy. However, the influence of market supply and demand on the inflation rate that emerges will depend as well on the expected inflation rate. A given degree of slack will result in a higher or lower overall inflation rate depending on how much inflation the price-setting agents expect to occur.

The inflation-unemployment relationship for the 1950s and 1960s lies along the single smooth Phillips curve shown in Figure 7 because in this period, the inflation rate was relatively low and agents generally believed that it would continue to be small. Throughout this period the expected rate of inflation was both very small and relatively unchanging. The economic expansion that occurred in the 1960s pushed the economy into an overheated situation in a few years. The inflation rate accelerated from year to year and agents began to change their expectations of inflation. As the expected rate of inflation changed, the relationship between inflation and unemployment changed as well; the Phillips curve shifted.

Figure 7
U.S. Phillips Curve, 1950-69

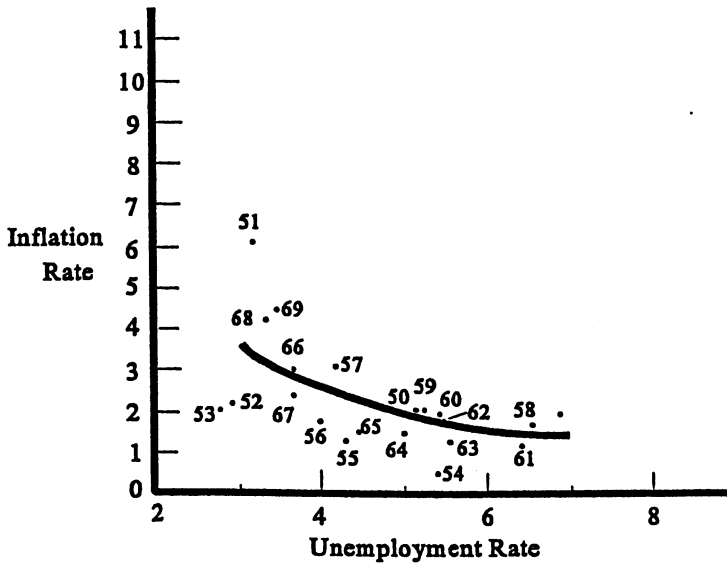
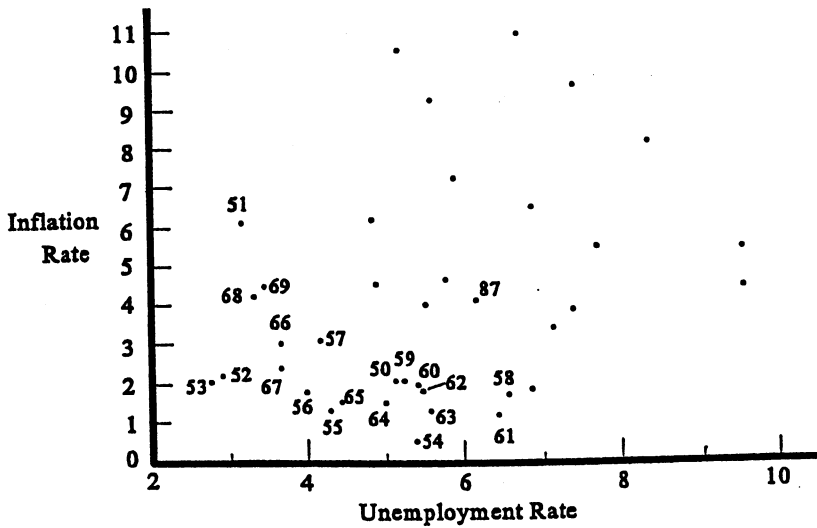


Figure 8
U.S. Tradeoff, 1950-1987



The so-called trade-off between inflation and unemployment that had been observed in the 1960s turned out to be not so favorable. Economists soon realized that the Phillips curve model had to be expanded to include the role of expectations of inflation. The expectations-augmented Phillips curve indicates that the attractive short-run trade-off is only a temporary phenomenon. The optimistic views of the trade-off expressed in the *Economic Report of the President*, 1963 quickly disappeared.

This new view was forcefully presented by Milton Friedman in his 1967 Presidential Address to the American Economic Association,¹ as follows:

“Implicitly, Phillips wrote his article for a world in which everyone anticipated that nominal prices would be stable and in which that anticipation remained unshaken and immutable whatever happened to actual prices and wages.”

Friedman went on to note that if inflation is anticipated, wages will rise at the same rate to maintain real wages and labor market equilibrium. Friedman concluded his view of the Phillips curve trade-off by stating:

“There is always a temporary trade-off between inflation and unemployment; there is no permanent trade-off. The temporary trade-off comes not from inflation per se, but from unanticipated inflation, which generally means from rising inflation A rising rate of inflation may reduce unemployment, a higher rate will not.”

Friedman’s development of the Phillips curve analysis has the important implication that there is no permanent way of trading off unemployment for inflation. In the long run the unemployment rate will tend towards the natural rate or the unemployment rate (see Chapter I) consistent with long-run equilibrium, Y^* .

The intuition underlying the natural rate Phillips curve was presented by Friedman in terms of the labor market. Curiously, this story makes use of the Keynesian assumption that wages tend to be sticky; that is, nominal wages move slowly relative to prices. Consider a situation where aggregate demand expands; as output increases, prices also begin to move upward. However, wages do not increase as quickly; there is some nominal wage stickiness in the short and medium-run. Therefore, real wages fall and employment increases as employers move along their downward sloping demand for labor curve. As employment increases, output increases. Furthermore, additional labor supply is forthcoming because workers do not perceive the real wage decline. There is a general misperception of the real wage decline for some period of time. However, over time as workers perceive the changes in real wages that have occurred, they adjust their expectations of inflation, labor supply adjusts, and the labor market returns to its equilibrium at the natural rate of unemployment. Thus, a trade-off between inflation and unemployment (or output)—the Phillips curve trade-off—exists as long as workers do not perceive what is happening. After a while, workers realize that inflation erodes real wages and they adjust their expectations and negotiating behavior accordingly. In the long run, there is no trade-off.

The relationship between the short-run and long-run trade-off can be illustrated by developing an augmented price adjustment equation that includes the role of inflationary expectations.

Augmented Price Adjustment Equation

The effect of an increase in expectations of inflation on the actual rate of inflation can be seen by envisioning a particular price or wage negotiation. The parties in a particular negotiating session will be influenced by supply and demand conditions in the market and also by their expectations of aggregate inflation. The price or wage agreement that emerges from the negotiations will be higher if both parties expect more inflation to take place in the overall economy. Thus, if more inflation is expected, the

¹ Milton Friedman, “The Role of Monetary Policy,” *American Economic Review*, March 1968, p.8.

bargainers are likely to settle on a higher price even with supply and demand conditions unchanged. The inflation rate that emerges from price setting throughout the economy will depend on both supply and demand conditions and the decision-makers' expectations of inflation. This suggests the following specification of the **expectations-augmented price adjustment equation**:

$$p = g(Y - Y^*) + b p^e \quad (5)$$

where p^e is the expected rate of inflation and the coefficient b measures the impact of expectations on the inflation rate.

Equation (5) represents a family of Phillips curves; there is a trade-off relationship between the inflation rate and the output gap (or unemployment rate) for each level of expected inflation. The expectations-augmented Phillips curve is illustrated in Figure 9. A flat trade-off curve is shown for each level of the expected inflation rate, p^e . Over time, the expected rate of inflation changes and the longer-run trade-off curve is given by the steeper line shown by ABC in Figure 9—a long-run trade-off.

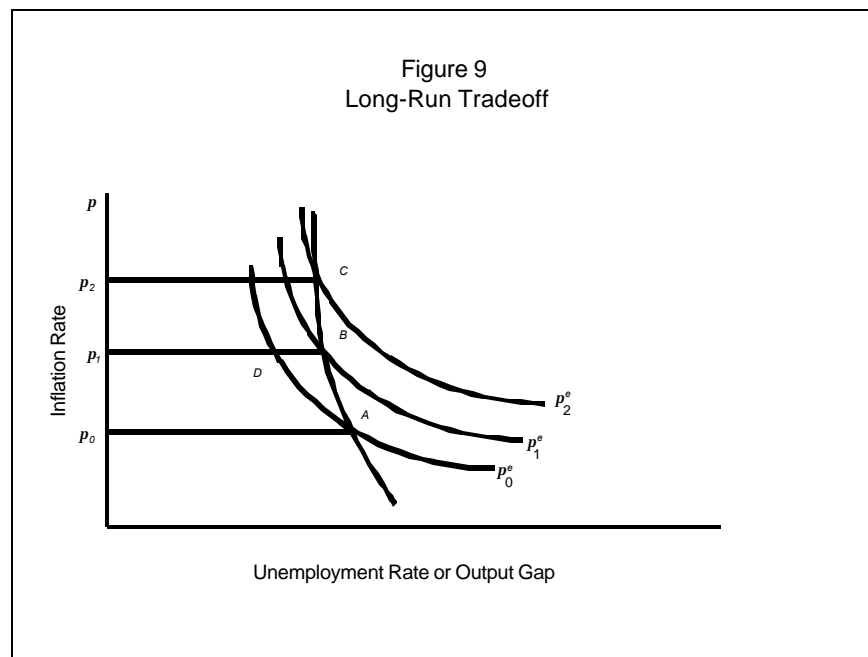
To develop the long-run trade-off, recall that changes in the actual inflation rate are likely to be associated with changes in the expected inflation rate as well. Ultimately, expected inflation should adjust to the actual inflation rate. With this in mind, we know that a characteristic of long-run equilibrium is that actual and expected inflation rates are equal. The condition that defines the long run is:

$$p = p^e \quad (6)$$

Substituting the long-run equilibrium condition (6) into (5) yields:

$$p = \frac{g}{1-b}(Y - Y^*)$$

The slope coefficient on the output gap is $g / 1 - b$; since $b < 1$, the long-run trade-off is steeper than the short-run curves.



The economic importance of the distinction between the long-run and short-run Phillips curves can be seen by reconsidering the expansionist policy recommendations of the early 1960s. The Keynesian economists of that era viewed the trade-off between unemployment and inflation as relatively flat and thought that a policy expansion would move the economy along to higher output with little effect on inflation. That is, as the economy expands, bottlenecks and inflationary pressures emerge and the inflation rate increases as the unemployment rate falls. However, the early Keynesians overlooked the fact that as inflation increases, expectations of inflation will increase as well.

The long-run trade-off in Figure 9 is drawn with the assumption that $b < 1$. This implies that expectations do not have a full impact on inflation. The true long-run is reached when all adjustments to change have taken place. If all the long-run adjustments have occurred, we expect that the bargaining process that determines wage and price inflation should fully reflect the expected inflation rate. In this case, a change in p^e has a one-for-one impact on the inflation rate, p , in the long run. In this view, $b < 1$.

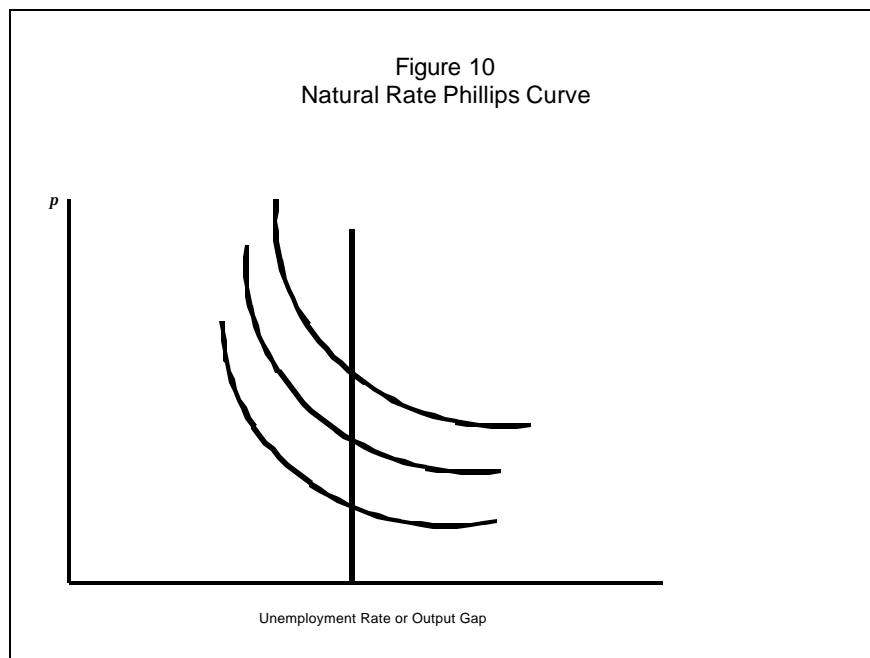
With a full impact of expectations, the expectations-augmented price adjustment equation should be written as:

$$p = g(Y - Y^*) + p^e$$

If we now solve for the long-run trade-off by substituting the long-run condition ($p = p^e$), we get:

$$Y = Y^*$$

In the long run, after all inflation adjustments have taken place, there is no trade-off between output and inflation; output simply returns to the natural level, Y^* . The relationship between inflation and the output gap or unemployment rate is shown in Figure 10. The long-run trade-off is given by the vertical line. Thus, the long-run equilibrium in this analysis is the same one developed in Chapter II.



Friedman's view is often called the accelerationist approach because of one of its major implications. That is, to keep the unemployment rate below the natural rate requires an increasing expansionary stimulus and a continuously rising inflation rate. If the initial policy change that moved the economy from A to D in Figure 9 was an increase in the rate of growth of the money supply, the growth rate must be continuously increased in order to prevent a return to the natural rate of unemployment. An accelerating rate of inflation would keep wages lagging behind prices, keep the real wage below its overall equilibrium, and maintain an unemployment rate below the natural rate.

The unemployment level associated with Y^* is called the natural rate of unemployment; it is not directly observable and the value of the measured unemployment rate that we call the natural rate might change over time. The *Economic Report of the President, 1983* states that

“... while it is not easy to pinpoint the inflation threshold unemployment rate precisely, it probably lies between 6 and 7 percent.”

In the 1990 *OECD Economic Surveys: United States*, the Organization of Economic Cooperation and Development reports that:

“Estimates of the unemployment rate at which inflation would stabilize—the so-called natural rate—vary from about 5 to 7 percent. Wage behavior has been restrained enough in the last decade to suggest that the natural rate is in the middle of the 5 to 6 percent range—an improvement in the 1970s probably due in part to an older and more educated work force, greater confidence that inflation would not be allowed to accelerate and improved productivity growth. But it is still higher than that which was obtained in the 1960s, when wages tended to accelerate only when the unemployment fell towards 4 percent.”

The actual employment rate was at least 7% in every year from 1980 to 1986 and the inflation rate declined steadily over this period. The unemployment rate declined rapidly from 7.0% in 1986 to 5.5% in 1988 and 5.3% in 1989 and the inflation increased during this period.

More recently, there has been considerable controversy about the natural rate of unemployment. From mid-1994 through 1996, the unemployment rate in the U.S. has been consistently less than 5.5 percent. The Congressional Budget Office's estimate of the natural rate of unemployment is 5.8 percent. Thus, there has been serious concern about the start of inflationary pressures. However, the inflation rate in the CPI has remained around 3 percent. In early 1997, policy-makers adopted a more restrictive policy that would reduce the rate of economic growth (real GDP grew at an annual rate of around 4.5% in the first half of the year) and keep the unemployment rate from dipping any further below the estimated natural rate. Such a policy would be a preventive measure, designed to stave off the inflationary pressures that are deemed to be inevitable sooner or later with the unemployment rate below the natural rate.

The consensus interpretation of the economic situation, however, has been challenged by many observers. These economists note that the unemployment rate has been below the estimated natural rate for some time without there being any evidence of inflationary pressures. Rather than concluding that increased inflation is just around the corner, they conclude that the estimates of the natural rate of unemployment must be mistaken. This approach argues that structural change in the economy and in labor force behavior imply that there can be lower unemployment and more growth without the danger of inflationary pressures. Indeed, to make policy restrictive when the unemployment rate is still above 5% would unfairly rob the U.S. economy of its growth potential.

Readers of this Chapter will have the experience of 1997 and later to help them conclude which interpretation was correct.

Summary

We have not developed three distinct views of the Phillips curve—inflation adjustment model. The first is the simple trade-off model that was used by policy-makers in the 1960s. The second was the expectations-augmented Phillips curve. With this model, a trade-off between inflation and unemployment exists in the long run, although it is not as favorable as the short-run trade-off because inflation begins to generate inflationary momentum and expectations of future inflation. The third model is derived from the assumption that expectations adjustments are complete. The full impact of inflationary expectations on inflation brings us full circle back to the long-run equilibrium model of Chapter II.

RATIONAL EXPECTATIONS

Our discussions of the formation of expectations have indicated that expectations adjust slowly when inflation changes. This has often turned out to be an accurate description of reality, but it is not necessarily true. Starting in the early 1970s an alternative hypothesis about the formation of expectations had a very profound effect on economic thinking. We will begin with an explanation of the rational expectations hypothesis and then examine its implications for our macroeconomic model. Although the hypothesis is at times unrealistic, its implications are important.

The rational expectations hypothesis states that expectations are knowledgeable and informed predictions of the actual outcome. That is, expectations are formed by individuals with an understanding of the workings of the economy and with available information on all relevant phenomena. Expectations of inflation are thus based on all available information that relates to price determination and with an understanding of how prices are in fact determined. If expectations of inflation are rational, expectational errors will not be systematic. Any known and systematic determinant of inflation will be taken into consideration when expectations are formed. Therefore, expectations of inflation and the actual inflation rate will differ only when a nonsystematic phenomenon or random shock affects the inflation rate.

With rational expectations, the expected inflation rate can be expressed as the actual inflation rate plus a random error term:

$$p^e = p + I$$

where I is a random error term, or “noise,” which is on average, zero. If it differs from zero in any systematic fashion, expectations, p^e , are not rational. With rational expectations, any systematic deviation will be observed and incorporated into the process by which expectations are formed. The rational expectations hypothesis represents the extreme position that expectations are always formed with full knowledge of how the world operates. Although factual considerations do not fully support this assumption, we will examine its implications for the Phillips curve model and policy-making.

The implication of the rational expectations hypothesis can be seen if we re-write the inflation adjustment equation as:

$$Y - Y^* = g(p - p^e)$$

and add the above definition of the difference between actual and expected inflation.

The Phillips curve/price adjustment model with rational expectations, has some startling implications. With expectations formed according to the hypothesis of rationality, there is no trade-off between unemployment (or the $Y - Y^*$ gap) and inflation at all. Unemployment is equal to the natural rate plus a deviation that is purely random. By the very definition of rationality, there is no opportunity for the systematic emergence of unanticipated inflation. Only surprises that are due to random events or an unanticipated policy change can lead to an inflation rate that differs from the expected rate. Thus the unemployment rate differs from the natural rate only when there is a shock or surprise. Furthermore, such

deviations must be short-lived since a surprise cannot last beyond the current period. In the next period it becomes part of the available information set used by economic agents to determine p^e rationally. Of course, a shock may appear to persist because its magnitude and extent were also a surprise. The OPEC oil-price shock could be viewed in this way. The change in energy prices was thought at first to be a temporary change associated with the oil embargo, but it turned out to be a permanent change in the relative price. However, some shocks seem to persist longer than the rational expectations hypothesis would suggest.

The rational expectations theory has been presented here as a particular hypothesis for the formation of expectations in a natural rate Phillips curve model. It is, however, a major building block in a whole class of macroeconomic models that can be termed the new classical approach.

New Classical Macroeconomics

The natural rate Phillips curve model implies that the unemployment rate differs from the natural rate when inflation is unanticipated. With rational expectations, unanticipated inflation is always a random or unpredictable phenomenon. Therefore, all deviations, including short-run deviations, of the unemployment rate from the natural rate are random events.

The policy implication of rational expectations can be seen by considering again an equilibrium disturbed by an expansionary policy. If the policy is known to economic agents, they will use their knowledge of how such a policy affects the economy. In our model with a natural rate, an expansionary policy leads to inflation and does not change the natural rate equilibrium. If expectations are formed rationally, expectations of inflation will adjust immediately and the new equilibrium will be established immediately. The immediate adjustment of p^e to all the implications of the policy expansion has eliminated the distinction between the long-run and short-run response. Hence the policy ineffectiveness proposition of the rational expectations school:

In an equilibrium (e.g., natural rate) model, where inflationary expectations are formed rationally, a fully anticipated policy will have no effect on the level of real economic activity.

Put succinctly, we have the startling and troublesome implication that “policy doesn’t matter.”

If the policy is not fully anticipated, real-sector effects do occur. A policy which surprises the public can lead to an error in inflation prediction and a deviation of the unemployment rate from the natural rate. However, such effects must be short-lived because economic agents will respond and adjust their expectations as soon as the policy becomes apparent.

It is important to note that the policy ineffectiveness proposition is a consequence of two elements of the model structure. First, the natural rate Phillips curve indicates that there is a unique and stable natural unemployment rate. Second, expectations are formed rationally, with complete knowledge of the underlying equilibrium structure and all relevant information. If the public is well informed about the economic structure, economic conditions, and policy goals, it will be able to forecast the inflation rate. That makes it impossible to set a policy that makes the unemployment rate differ systematically from the natural rate. As long as expectations are an informed forecast which utilize the actual structure of the economy, the unemployment rate will differ from the natural rate only if there is a policy surprise. That is, an effective policy must be unanticipated. Only by systematically fooling the public can policy-makers force the unemployment rate to deviate from the natural equilibrium.

There may be some reasons, though, why the public may be fooled into making errors in forecasting inflation. First, economic agents may have an imperfect understanding of the workings of the economy; second, their information may be limited. Policy-makers may have better and more up-to-date

information about the structure of the economy. In addition, the policy-makers may be able to fool the public by pursuing an expansionary policy without saying so. Finally, institutional constraints such as long-term labor contracts may lead to expectational errors.

Deviations of the unemployment rate persist for much longer periods of time than the rational expectations hypothesis suggests. Nevertheless, the implications of the rational expectations Phillips curve are important. The policy ineffectiveness proposition has had a profound effect on our understanding of what macroeconomic policy can accomplish.

There is an additional lesson to be learned from the new classical approach. It is often called the policy evaluation or uncertainty proposition. This proposition has very important real-world implications. The proposition is as follows:

In a model where expectations of inflation are based on all available information and formed with an understanding of the structure of the economy (i.e., expectations are rational), the responses of the economy to economic policy initiatives are variable and uncertain.

This proposition is not as unrealistically strong as the policy ineffectiveness proposition. Nevertheless, it implies that it may be impossible to use discretionary policy to guide the economy because the implementation of policy alters the responses.

SOURCES OF INFLATION

In this section we will review the phenomena that generate inflation and maintain its momentum. The distinction among the different sources of inflation is somewhat artificial because they can all be present and are often related to one another. However, different inflationary episodes can be often be ascribed to a particular dominant causal factor. The sources examined are:

- monetary growth
- excess demand
- relative price shocks
- wage-price spiral or cost push
- expectations

Monetary Growth

The first thing to note is that a sustained inflation cannot be maintained forever without monetary expansion. If the money supply is growing less rapidly than the price level, then the stock of real money balances (M/P) is declining. Unless monetary policy accommodates the ongoing inflation, monetary policy is contractionary and will ultimately pull the economy into a slowdown, which will over the long run remove the inflationary pressures.

More generally, the usual context for examining the relationship between inflation and money growth is the quantity theory of money that was developed in Chapter II. More appropriately termed the quantity identity, it is a functional relationship between the money stock and nominal income:

$$M V = P Y$$

where M is the stock of money, P the price level, Y real output and V , called velocity, is defined by the relationship. We can view velocity as the rate at which the money stock turns over in order to generate some level of nominal income, PY .

To discuss inflation, examine the quantity equation in terms of growth rates:

$$\% \Delta M + \% \Delta V = \% \Delta P + \% \Delta Y$$

Our earlier examination of money demand indicated that velocity—the speed of turn over of money balances—will vary with the level of interest rates. Over long periods of time, velocity varies around an increasing trend. The trend is increasing because technological improvements in the financial system have increased the rate of turnover. Thus, in the long run we can view the $\% \Delta V$ as constant.

Furthermore, in the long run the economy reverts to Y^* and the long-run rate of growth of output is the rate of growth in Y^* . Thus, the $\% \Delta Y$ is determined by technology and growth in the labor force and capital stock. Thus, in the long run, inflation $\% \Delta P$, is determined by the rate of monetary growth. This is the reason for the adage often suggested by monetarists that ‘inflation is a monetary phenomenon’.

In the short-run, there can be wide variations in monetary growth which have little apparent impact on the inflation rate. However, over long periods of time there is a very strong correlation between inflation rates and monetary growth rates.

Excess Demand

The major thrust of our theoretical model (the total supply and demand framework) is that disequilibrium between output and demand leads to quantity adjustment and price adjustments. When there is available productive capacity, we expect that quantity adjustments would be dominant. However, if there is excess demand and limited availability of capacity and other productive resources, then price adjustments will dominate.

Relative Price Shocks

Relative price shocks have on occasion had a major impact on the inflation rate. A relative price shock is a sudden change in the price of particular goods that play an important role in overall production. The common examples are weather conditions that can affect the prices of agricultural products and change in the prices of raw materials like the oil price shocks of the 1970s. We already analyzed such supply shocks that can reduce the level of output and increase prices.

Supply shocks can have a lasting effect on inflation. Imagine that the OPEC cartel is once again able to impose a large increase in the price of oil. The price of energy increases and this effects prices generally. In addition, expenditures on energy increase and less income is available for other expenditures. In all likelihood, monetary policy will accommodate the price increase from the initial shock and increase the money supply. If monetary policy does not accommodate the price shock, the economy will contract into a recession. Since policy-makers are anxious to avoid recession, the shock is likely to be accommodated. Thus, the initial shock raises prices and leads to monetary accommodation, which leads to the price increases spreading through the economy.

Wage Price Spiral

The momentum of inflation is hard to break because price increases will be quickly passed around the economy from one price sector to another. In particular, wage setting behavior (in both unionized and non-unionized sectors) tends to reflect changes in the cost of living or the aggregate inflation rate. Furthermore, many firms determine the prices of finished goods (at least in the short-run) by adding a markup onto costs.

The combined influence of markup pricing and the cost of living on wages means that inflation has a great deal of momentum. Any inflationary shock that enters the spiral of price and wage setting can effect the process for years. Even in the presence of large changes in demand, the momentum of the inflation process goes on for some period of time.

Inflation Expectations

Another element of the price setting spiral is the role of inflationary expectations. Inflation expectations have a strong influence on the actual rate of inflation. In addition, inflation expectations can be very slow to change.

A policy-induced contraction can have only a minor effect on the inflation even in the presence of considerable excess capacity because of the resilience of expectations. Inflationary expectations may be unaffected because of the lack of credibility of the anti-inflationary policy. If economic agents do not think that the policy-makers will stick to anti-inflation policy, then expected inflation might not change when the economy enters a slowdown. As a consequence, the momentum of inflation driven by expectations can be a very important determinant of the inflation rate.

Inflation expectations are important for examining inflation for another reason as well. That is, the interest rate relevant for decision making is the ex ante real rate.

Although the expected real rate of interest is difficult to measure, economic theory tells us that it is the relevant determinant of behavior. For example, the decision to make a capital investment will depend on the expected real return from the project and the expected real cost of financing. These measures may or may not change when there are changes in the observed nominal interest rate. If nominal interest rates rise from 5% to 10% and at the same time expected inflation rates increase by the same amount, then the expected real return is unchanged. If all economic agents shared this expectation, then we would not expect the change in interest rates to have any implications for behavior. However, expectations of inflation are not likely to be shared by all economic agents.

Thus, all of our discussions of the effects of interest rates on expenditure decisions must be qualified. A decrease in nominal interest rates will increase investment expenditures if it is perceived to be a decrease in the real rate of interest. We now understand why the response of expenditure to changes in interest rates can be so unpredictable and variable. The response depends on how economic agents interpret a given change in nominal interest rates.

An easing of monetary policy will have an immediate effect on short-term interest rates, but it may not affect the long-term interest rates that are relevant for investment decisions. For example, the move towards an easier monetary policy between mid-1990 and early 1991 reduced short-term interest rates by almost 2 percentage points. At the same time long-term interest rates declined by only one-half of a percentage point. The reason for this is simply that financial market participants did not expect the easing to be permanent and over the long-run expectations of inflation still persisted. Thus, long-term rates moved very little and the impact on investment expenditures is likely to be fairly small.

CHAPTER V BANKS, MONEY AND MACRO POLICY

In this chapter, we turn our attention to financial institutions and their relationship to macroeconomics. The reason to devote an entire chapter to money and banking is that the monetary sector and monetary policy are currently the focus of most macroeconomic policy discussions. There are two reasons why this is so:

First, there is general agreement that in the long run, monetary policy and the growth of the money stock determine the inflation rate. Thus, as policy making in recent years has emphasized the importance of maintaining a stable inflation rate that is as close to zero as possible, policy making has focused on monetary policy.

Second, in the short run, monetary policy through its influence on interest rates and credit availability is the principal tool of macroeconomic stabilization policy. Persistent government deficits and the slow macroeconomic responses to fiscal policy changes have resulted in increased emphasis on monetary policy in the short-run. Thus, when recession or international payments crises loom, the principal policy responses come from monetary policy and central banks are the most important influences on monetary policy.

In Chapter II, we presented the formal definition of the stock of money that is used in the United States by the central bank—the Federal Reserve System. At the present it suffices to recall that *M1* is currency and the transactions deposits at banks and other intermediary institutions. The broader definitions *M2* and *M3* add various deposit balances and other financial assets that can be converted into transactions balances with little risk or cost. These non-transactions components of the broader money supply definitions are so close to being transactions assets that they are often called near-moneys. Substitution between transactions assets and other components of the broader aggregates is now both very common and very easy. (You probably do it often, and in just a few seconds, at your local ATM when you transfer funds from a savings or money market account to a checking account or to cash.) As a result, the broader aggregates are probably the best current definition of money. When we think of the money supply (that is, the stock of money outstanding), we will usually have an *M2* aggregate in mind. This was not always the case; prior to the deregulation of bank activities and the simultaneous development of the technology for electronic funds transfers (all of which started in the 1960s and speeded up in the 1980s), the most commonly used definition of money was a narrow one, *M1*.

BANKS AS CREATORS OF MONEY

The most important part of the money supply is the deposit liabilities of financial institutions. Such deposits are created by the banking institutions themselves. When a bank purchases an asset it pays for it by crediting the account of the seller and in so doing it creates a deposit. Similarly, when a bank makes a loan it is also buying an asset—the lender's IOU—and it simply credits (increases) the lender's deposit account by the amount of the loan. That is, when a bank makes a loan, it creates a new asset (the loan contract or IOU) and a new liability for the bank (the increase in the bank's deposit balances).

Of course, a bank's willingness and ability to buy assets (or make loans) and increase the money supply will be guided by its profit maximizing behavior and tastes for risk. Thus, good business behavior will place restraints on the amount of money creation that the bank will do. For example, it needs to maintain adequate reserves in the form cash in the vault or deposits at other banks or at the central bank.¹ Reserves are held for two reasons. First, a bank will want to hold reserves that are sufficient to meet the requests of

¹ Reserves are the liquid or cash assets of the bank. Reserves (usually in the form of vault cash or deposits at the central bank) may be required by the regulatory authorities. The use of the term reserves here should not be confused with the loan loss reserve, a part of the bank's capital which is set aside to cover bad loans.

depositors to withdraw their funds. Thus, the bank needs reserves so that any depositor's request for cash can be readily met. Second, deposit-issuing institutions are often required by regulatory authorities to hold reserves that are at least as great as a certain fraction of their deposits.

The Federal Reserve imposes reserve requirements for two reasons. The first is to promote the safety and stability of the banking system, i.e. to be sure that every bank has adequate reserves. The second is to control the aggregate money supply. Required reserves can be held in the form of cash in the bank's vaults or as the bank's own deposits at the central bank (the Federal Reserve).

The practice of holding a fraction of deposits in cash reserves, called fractional reserve banking, is a key part of the money creation roles of banks. We will start with a brief examination of the money creation process that shows how a banking organization works. The section that follows will develop a more formal model of the banking system and money creation process that will show how monetary policy is conducted.

Fractional Reserve Banking

Suppose that banks keep one-fifth of their deposits as reserves either as a legal requirement or by choice in order to be prepared for possible withdrawals. Also, for the simplest possible framework, we will assume that the banking system always maintains exactly that ratio of reserves to deposits. In this section we will explore some of the implications for the banking business.

To begin, imagine that someone comes along and deposits \$10,000 in currency in Bank A. Let's see what might happen. Bank A experiences an increase in its cash (asset) and its deposits (liabilities) of \$10,000. The cash is placed in the vault and there is an increase in the reserves (cash in the vault) and deposits of Bank A. We depict these balance sheet changes in the first panel of the **T-accounts** in Table 1, which shows changes in bank's assets (on the left) and liabilities (on the right).

TABLE 1

Changes in Bank A's Balance Sheet

<u>Assets</u>	<u>Liabilities</u>
*BANK RECEIVES DEPOSIT	
Reserves +10,000	Deposits +10,000
Loans and Investments No change	Net Worth No change
Total +10,000	Total +10,000
*BANK MAKES LOAN	
Reserves No change	Deposits +8,000
Loans and Investments +8,000	Net Worth No change
Total +8,000	Total +8,000
*MS. SMITH SPENDS	
Reserves -8,000	Deposits -8,000
Loans and Investments No change	Net Worth No change
Total -8,000	Total -8,000
**BANK A SUMMARY	
Reserves +2,000	Deposits +10,000
Loans and Investments +8,000	Net Worth No change
Total +10,000	Total +10,000

Bank A realizes that it does not need to use all of the \$10,000 in additional cash as reserves. Since it only needs to keep one-fifth of the additional deposit as reserves, it finds that it only needs to keep \$2000 as additional reserves. It then has excess reserves of \$8,000, which it can use to buy other assets—make loans and investments.

A bank customer, Ms. Smith, now comes to the bank to take out a loan to finance an increase she would like to make in her business inventories. The bank judges this to be a good loan and it lends Ms. Smith the funds (\$8,000) and creates a demand deposit for her. That is, it simply writes up an additional balance in Ms. Smith's checking account. She can then write a check to Ms. Jones, her supplier, to pay for the new inventories.

Ms. Jones then deposits the check in Bank B. Bank B then presents the check to Bank A for payment. The check is cleared through the Federal Reserve System where the deposits (reserves) of Bank A are decreased by \$8,000 and the deposits of Bank B are increased by \$8,000.

We can trace the effect of this series of transactions on the balance sheet of Bank A. As a result of the increase in deposits of \$10,000, Bank A is able to expand its loans and investments by an additional \$8,000. In addition, the money supply has been expanded by \$8,000 since Ms. Jones now has a demand deposit (in Bank B) of that amount which did not exist before. In summary, Bank A's activities have increased both the money supply and credit by \$8,000.

Next, we follow the process to Bank B where Ms. Jones has deposited the check from Ms. Smith. Bank B has cleared the check through the regional Federal Reserve Bank and finds that it has both deposits and reserves that have increased by \$8,000, as shown in Table 2.

Bank B does not need to keep the full \$8,000 as reserves since its reserve requirements are also only one-fifth of deposits. Bank B's increase in required reserves is only $\$8,000/5$ or \$1,600. Bank B now has excess reserves of $\$8,000 - 1,600 = \$6,400$ that it can use to make loans and investments.

Suppose that the bank decides to buy \$6,400 of bonds from a depositor, Mr. Green. The bank credits Mr. Green's account by \$6,400 in return for the bonds, and Mr. Green uses the proceeds of the sale to buy some furniture from Ms. Stone's furniture store. Ms. Stone takes the check from Mr. Green and deposits it in Bank C.

We can see several things from the T-accounts for Bank B. The inflow of deposits of \$8,000 into Bank B required this bank to increase its reserves by only \$1,600. The bank found that it had excess reserves so that it was able to expand loans and investments by \$6,400. When it did so, it increased the deposits of Mr. Green by that amount, increasing the quantity of money held by the public by an equal amount. Thus Bank B has contributed to an expansion of money by \$6,400—this amount eventually becomes additional deposits in Bank C. The money supply has increased by \$8,000 due to Bank A's activities and by \$6,400 from Bank B's activities.

TABLE 2

Changes in Bank B's Balance Sheet

<u>Assets</u>	<u>Liabilities</u>
*BANK RECEIVES DEPOSIT	
Reserves +8,000	Deposits +8,000
<u>Loans and Investments No change</u>	<u>Net Worth No change</u>
Total +8,000	Total +8,000
*BANK BUYS BOND	
Reserves No change	Deposits +6,400
<u>Loans and Investments +6,400</u>	<u>Net Worth No change</u>
Total +6,400	Total +6,400
*MS. GREEN SPENDS	
Reserves -6,400	Deposits -6,400
<u>Loans and Investments No change</u>	<u>Net Worth No change</u>
Total -6,400	Total -6,400
**BANK B SUMMARY	
Reserves +1,600	Deposits +8,000
<u>Loans and Investments +6,400</u>	<u>Net Worth No change</u>
Total +8,000	Total +8,000

What is the total effect on the banking system of the initial cash deposit of \$10,000 at Bank A? Clearly the effect of this initial increase in deposits has spread beyond Bank A. We could continue the process by tracing through the transactions that result from Bank C's deposit inflow of funds and see that Bank C will also be able to expand bank credit and expand the money supply. The process goes on indefinitely. An increase in reserves (from the initial deposit of cash) has a multiplier effect on the money supply. In this simple example—where every bank always holds exactly one-fifth of its deposits in the form of reserves—it is easy to figure out the total effect of the initial inflow of reserves on amount of money and credit created by the banking system.

As a result of the initial increase in deposits of \$10,000, Bank A finds that it has excess reserves of \$8,000. It was this quantity of excess reserves that started the process of credit creation and money supply expansion in motion. Bank B creates money and credit equal to $\$6,400 = (4/5)*8,000$. Bank C is able to create money and credit equal to $\$5120 = (4/5)*6400 = (4/5)*(4/5)*8,000$. The process continues indefinitely, with each successive bank in the process expanding money and credit by $4/5$ the amount of the previous bank. The net effect of the process is that the money supply will be expanded by an amount that is $1/(1 - (4/5)) = 5$ times the initial increase in excess reserves.

The Money Multiplier

If an amount of excess reserves is made available to the banking system, the banking system as a whole can increase the money supply by a multiple of the amount of excess reserves. This multiple depends on the size of the reserve ratio, which is the ratio of reserves held (because of legal requirements or as a consequence of prudent banking business) to deposits.

Suppose that banks hold reserves equal to a fraction, k , of their deposits. For now, we will continue to assume that banks have only one kind of deposits and that banks never hold excess reserves or have

reserve deficiencies. If we let R be the volume of reserves of the banking system and let D be the volume of demand deposits held by the public, then:

$$R = kD.$$

If we divide both sides of this equality by k , we find that:

$$\frac{R}{k} = D$$

Furthermore, the change in deposits for a change in reserves is:

$$\frac{dD}{dR} = \frac{1}{k}$$

If the reserve ratio is $k = .2$ and if the volume of reserves is 2000 then the volume of deposits must be 10,000. If deposits exceeded this amount, banks would not be holding sufficient reserves. If deposits fell short of this amount, banks would be holding excess reserves.

The money supply consists of currency (C), checkable (demand) deposits (D) and time (savings) deposits (T):

$$M = C + D + T$$

That is, the money supply is an $M2$ —broad money—definition, but we will call it M for simplicity. We will assume that the public's holding of currency and time deposits are both determined exogenously. The banks are required to hold reserves that are equal to a fraction of their demand deposit liabilities:

$$RR = k D$$

where k is the reserve requirement ratio. Reserves of the banks are held in the form of deposits at the central bank (for simplicity, we will assume that the banks do not hold vault cash but deposit all their currency with the central bank). Total reserves (TR) are the sum of required reserves (RR) and excess reserves (RE):

$$TR = RR + RE$$

The assets of the central bank (the Federal Reserve or the Fed in the U.S.) consists of a portfolio of securities, which we call Q and loans from the central bank to banks. The borrowing by banks from the central bank is denoted by B . Central bank assets consist of the deposits held by banks and other financial institutions at the central bank. These deposits are total reserves TR (since for simplicity, the banks do not hold any cash reserves). These deposit balances are the reserve assets of the banks and deposit liabilities of the Fed. In addition, the currency in circulation is a liability of the Federal Reserve. Take a look at the dollar bills in your pocket, they are notes—Federal Reserve notes—issued by the Fed.

The central bank's balance sheet is:

Q	TR
B	C

For simplicity, we do not include some other assets and liabilities that are actually on the balance sheet of the Fed. For example, the Treasury and foreign governments will also have deposit balances at the Fed. The most important omitted asset is the foreign exchange holdings of the Fed.

The securities on the Fed's balance sheet are U.S. government notes, bills, and bonds. By buying and selling securities, the Fed is able to exercise considerable control over the quantity of bank reserves. As we will see shortly, this is the most important way that the Fed exerts control over the money supply and the economy. Loans to banks are made by the Fed to banks that find themselves short of reserves. We will also see that the interest rate charged on these loans, the discount rate, is another means by which the Fed carries out monetary policy.

The money supply function is derived from the central bank's balance sheet identity:

$$Q + B = TR + C$$

Substituting the definitions for TR , RR yields:

$$\begin{aligned} Q &= TR - B + C \\ &= RR + RE - B + C \\ &= kD + RE - B + C \end{aligned}$$

Finally, use the definition of the money supply to substitute for D and define free reserves as $RF = RE - B$:

$$Q = k(M - C - T) + RF + C$$

Solving for M , gives:

$$M = (1/k) \{Q - (1 - k) C - RF\} + T$$

One additional element needs to be added to the relationship—the banking system's demand for free reserves. Some banks may hold excess reserves while some may be borrowing reserves from the central bank. The overall position of the banking system is RF , which can be positive or negative. It will depend on the behavior of the banks, which we will consider next.

To examine bank demand for reserves, it will be useful to have the aggregate balance sheet for the banking system as a reference:

Loans	D
Securities	T
TR	B

Assets of the banking system include loans and securities and the reserve deposits at the central bank, TR . The liabilities of the banking system include their deposit liabilities—the demand and time deposits of the public (D and T)—and bank borrowing from the central bank— B .

Some banks may wish to hold excess reserves while some banks may have deficiencies that are met by borrowing reserves from other banks or from the Fed. Each bank will manage its balance sheet differently and the overall situation for the banking system is reflected by the amount of free reserves, RF . We will examine the costs and benefits of reserve positions in order to derive the demand for free reserves.

In the U.S., reserve deposits at the Fed do not earn any interest. Thus, if interest rates are high, there is a substantial opportunity cost to free reserves. If interest rates are low, excess reserves or liquidity is less costly and some banks may prefer the cushion provided by excess reserves. In summary, the demand for excess reserves— RE —will be decreasing with market interest rates, r . If interest rates are high, banks will be less willing to hold non-interest-earning reserves and will prefer to purchase earning assets (e.g., loans or securities).

The demand for borrowing— B —will depend on the cost of borrowing from the central bank. In the U.S. the interest rate on Fed loans to the banks is known as the discount rate — r^d . Borrowing will be a declining function of r^d . In addition, as market interest rates— r —increase, profit maximizing banks will have an incentive to borrow reserves (increase B) and buy more assets (loans or securities). Thus, borrowing is an increasing function of market interest rates and a decreasing function of the discount rate.

The demand for free reserves can be written as:

$$RF = RF(r, r^d)$$

RF increases with r^d and decreases with r . Substitution yields our money supply function:

$$M = (1/k) \{ Q - (1 - k) C - RF(r, r^d) \} + T$$

The money supply depends on behavior of the public, and the banks, as well as central bank policy decisions. The public determines its hold of Currency and Time deposits— C and T . The banks determine their demand for liquidity, the level of free reserves, RF . Finally, the central bank sets the level of the reserve requirement ratio, k , the discount rate, r^d , and the size of its portfolio of securities, Q . These correspond to the three tools of monetary policy—reserve requirement ratios, discount rate and open market operations.

The money supply is an increasing function of market interest rates since $RF(r, r^d)$ decreases as r increases. With higher interest rates, the banks will find it profitable to make more loans and hold reduce free reserves. Also, the money supply changes when the policy tools (Q , k , and r^d) change or the public changes its demand for currency. M^s increases when the central bank adds to its bond portfolio (an increase in Q , which supplies additional reserves), reduces reserve requirements (k), or reduces the discount rate (r^d). An increase in demand for currency (C) absorbs reserves and decreases M^s . We will first discuss the structure of the Federal Reserve system and then return to a full discussion of each of the monetary policy tools.

STRUCTURE OF THE FEDERAL RESERVE

Early in the history of the United States there was a strong central bank. The First bank of the United States was established in 1791 by Alexander Hamilton with substantial powers over the banking system. However, in the populist era of Andrew Jackson, opposition to the concentration of economic power and the extension of federal government authority over the states emerged. The charter of the central bank was not renewed, and the country entered a period of free banking in which entry into the banking business and the expansion of credit were largely unregulated. The National Currency Act of 1863 imposed some reserve requirements and solved the problem of not having a uniform national currency. However, there was no central bank to provide a loan to the banks when they had temporary liquidity problems or to influence the overall availability of reserves and the growth of the money supply. The need for a strong central bank became apparent during the panic of 1907, and the Federal Reserve Act was passed in 1913.

The act created the Federal Reserve system that exists today. Although the structure of the central banking system has been largely the same for over 80 years, the functions and activities of the system are now very different from what was originally envisioned. Today we emphasize the macroeconomic monetary policy role of the Federal Reserve but this function emerged only in the last few decades. Originally, the primary function of the central bank was to smooth out fluctuations in the banking system by providing liquidity when needed.

The Federal Reserve System consists of twelve regional banks and the seven-member Board of Governors which is located in Washington, D.C. Nineteenth-century wariness of central government power led to great emphasis on a decentralized structure. The Federal Reserve Bank of New York has considerably more influence than the other regional banks because some of the important policy operations of the system (open market operations, the buying and selling of securities, and the foreign exchange market interventions) are conducted by the New York bank, and most of the nation's largest banks are located in the New York district. However, the center of power in the entire system has tended over the years to shift away from the regional banks and the New York Fed in particular and toward the Board of Governors, and the board chairman in particular.

Each of the twelve regional Federal Reserve banks is owned by the member banks and controlled by a Board of Directors. However, the profits of the Federal Reserve banks, which are substantial, are turned over to the Treasury.

The Federal reserve Banks and the Board of Governors play a diverse set of roles, not all of which relate to macroeconomic policy. Therefore, the functions of the banks will be briefly outlined here.

The functions of the Federal Reserve banks can be separated into "chore" functions and policy functions. The chore functions are to:

- Examine the member banks
- Review merger applications
- Provide check-clearing services and electronic funds transfers
- Act as an agent for the distribution of new currency and for the sale of Treasury securities.

The policy functions of the Federal Reserve banks are to:

- Set the discount rate, which is the rate which financial institutions pay for borrowing from the Fed. In practice, the Board of Governors determines the rate that the regional banks are allowed to set; it is almost always uniform across the country.
- Administer the discount window. Banks are discouraged from making use of borrowings from the Federal Reserve on a continuing basis, and each regional bank determines policy on how much borrowing to allow.
- Participate on the Federal Open Market Committee (FOMC). The regional bank presidents sit on the FOMC and have voting privileges on a revolving basis, with the exception of the president of the New York Fed, who is a permanent voting member. The important role of the FOMC will be discussed below.

In addition, the Federal Reserve Bank of New York has several functions that are not shared with the other banks:

- The Open Market Desk where transactions in government securities are conducted is located in the New York Fed.
- The New York Fed holds the deposits and securities of foreign central banks, and any U.S. central bank intervention in the foreign exchange markets is directed there.
- The New York Fed holds and manages the country's gold stock.

The Board of Governors consists of seven members who hold 14-year terms and are appointed by the President. The chairman is the center of power for the whole Federal Reserve System. He or she is a board member appointed to a 4-year term as chairman by the President. Interestingly, the term does not coincide with the President's. Membership on the board is neither particularly remunerative nor very exciting, and so the governors rarely stay for full terms. The purpose of the long term was to make the governors who are appointed by the President subject to congressional approval, independent of political influence or concerns.

The formal functions of the Board include the following:

- Approve bank mergers.
- Set the regulations that determine what activities commercial banks are allowed to engage in.
- Set reserve requirements and approve changes in the discount rate.
- Direct open market operations through the FOMC.

The Chairman of the Board directs the Board staff and is the most powerful individual in the system. Recent chairmen have become national public figures while in office (Alan Greenspan, Paul Volcker, Arthur Burns). The chairman is also the dominant figure on the Federal Open Market Committee.

A major feature of the Federal Reserve system is its independence from the rest of the government. The regional banks are not government agencies and the Board of Governors is largely independent of the President and the Congress. Monetary and banking policy are set by the Fed and not directly by elected officials.

However, there is some modest oversight of the Fed's policy making. The Full Employment and Balanced Growth Act of 1978 requires that the Board chairman testify before congress twice a year on the monetary policy objectives of the Federal Reserve. Fed officials do not meet with the President or his staff when considering monetary policy decisions.

Although enormously important to society, the bank regulatory responsibilities of the Fed are less relevant to our focus on macroeconomic policy and will not be discussed in detail. The macroeconomic policy role of the Fed is centered on the activities of the FOMC.

The Federal Open Market Committee (FOMC) consists of the Board's governors and the presidents of the regional banks. Five of the twelve Federal Reserve Bank presidents are voting members of the FOMC at any one time. The committee meets in Washington about 8 times a year and reviews economic conditions and the behavior of the financial system. The purpose of the meeting is to formulate a directive on monetary policy. This directive provides instructions to the managers of the Open Market Desk (at the New York Fed) about how to conduct open market operations. If changes in economic and financial conditions warrant, the committee holds a telephone consultation between its regular meetings and changes the directive.

The way in which the FOMC directs monetary policy has changed over the years. At the present time, the FOMC establishes a target for a particularly important short-term interest rate—the Federal Funds

rate—which provides guidance for the purchase and sale of securities by the open market desk at the Federal Reserve Bank of New York. The Federal Funds rate is the rate at which reserve deposits (funds at the Fed or Fed Funds) are lent by one financial institution to another. A bank with more reserve deposits than it would like to hold can lend them (sell Fed Funds) to another bank. Thus, the Fed Funds rate is a market determined interest rate; it is not set by the Fed. However, as we will see below, the buying and selling of securities by the Fed will have a very direct influence on the market Funds rate so the Fed can keep the Fed Funds rate at or near the target set by the FOMC. It does so by buying and selling securities from the Fed’s portfolio (Q in the model above). In the next section we will show how changes in Q affect interest rates, the money supply and the economy.

The other Fed policy instruments are the reserve requirement ratio and the discount rate. The reserve requirements are set (within a legislative framework) by the Board of Governors and the discount rate is determined by the Board as well. In the next section, we return to our model framework and explain how each of the three monetary policy tools work.

TOOLS OF MONETARY POLICY

A simple and informative way of exploring the money supply relationship is to examine how changes in each of the policy variables affect the money supply.

Open Market Operations

The Federal Reserve can increase its holdings of government securities, Q , by making open market purchases. The money supply function clearly implies that M increases when Q goes up. However, we need to understand why this is the case—how does an open market purchase of securities by the Fed lead to an increase in the stock of money. Since the 1960s open market purchases and sales have become the dominant policy tool used to influence the money supply.

The best way to see how open market operations work is to trace the effects of a purchase of an existing government bond by the central bank on the Fed, the banks and the public. Imagine that THE FED buys a government bond from the public (ME). I get a check from the Fed for the price of the bond and the check is written against the Fed itself. Here’s what happens to the balance sheets:

THE FED buys MY T-Bill:

<u>ME</u>		<u>THE FED</u>	
Fed check +		T-bill	Fed check
T-bill -			

Of course, I take the Fed check and deposit it in my account at MY BANK. So, there is a change in my balance sheet and my bank’s:

I deposit the check in MY BANK:

<u>ME</u>		<u>MY BANK</u>	
Bank deposit +		Fed check	Bank deposit
Fed check -			

My part in the transaction—the open market operation—is now complete. I sold my T-bill and I now hold money—a bank deposit—instead. MY BANK presents the check to the Fed for collection. The Fed credits MY BANK’s account at the Fed with the amount of the check and the open market purchase is concluded:

MY BANK takes check to THE FED:

<u>MY BANK</u>		<u>THE FED</u>	
Fed check -		Fed check	Reserve deposit
Reserve deposit +			

The summary below highlights the important implications of the Fed’s open market purchase on the balance sheet of the Fed and of the banks. An open market purchase adds to the Fed’s portfolio of government securities. The Fed pays for this by creating reserve deposits. For the banking system, the deposit balances of the public have increased and the bank’s reserves held at the Fed have also increased.

There are important implications of this increase in reserve availability. First, there are more reserves available and the price at which reserves are borrowed and lent among financial institutions will change. The open market purchase increases reserves and the Fed Funds rate will fall. The Fed will use open market activities to influence the Fed Funds rate. Second, the increased availability of reserves will lead to changes in bank behavior. The banks will buy more assets and make more loans.

The open market purchase increases reserves and leads to more bank lending and thus an expansion in the money supply.

SUMMARY: The Effect of an Open Market Purchase on Securities

<u>Federal Reserve Bank</u>		<u>My Bank</u>	
Assets	Liabilities	Assets	Liabilities
T-Bill	Reserve deposit	Reserve deposit	Public’s deposit

The Fed creates bank reserves when it purchases Treasury bills or other securities. The effect of this creation of bank reserves is to increase the money supply initially by the amount of the purchase. In addition, the amount of reserves in existence increases by the same amount. The banking system will find that it has excess reserves since its deposit liabilities and reserves have increased by the same amount.

These excess reserves can then be used for loans and investments, and the ultimate impact on the money supply will exceed the initial increase due to the multiple expansion of deposits that was discussed above.

Now suppose that the Fed sold \$1m in government securities in the open market. If Mr. Y purchases these securities from the Fed and pays with a check drawn on an account at HIS BANK, the Fed's assets will decline by \$1 million as well as its liabilities. HIS BANK finds that its deposits have declined by \$1 million and its reserves will also decline by this amount when the Fed clears the check by debiting HIS BANK's account at the Fed. After the open market sale of securities by the Fed, HIS BANK finds itself short of reserves since deposits and reserves have declined by the same amount. As HIS BANK then reduces its loans and investments in order to increase its reserves, the process of multiple expansion of deposits will begin to operate in reverse, and the ultimate contraction of the money supply that results from the open market sale will exceed the initial size of the sale.

When the Fed **purchases** securities, it **adds** reserves to the banking system and the Fed Funds rate **declines**.

When the Fed **sells** securities, it **drains** reserves from the banking system and the Fed Funds rate **increases**.

The Fed uses open market operations very actively on a daily basis. They are the primary tool of monetary policy which are used to increase and decrease reserves in the banking system in order to keep the Federal Funds rate very close to the target set by the FOMC. The target FED funds rate is determined by the FOMC as the rate that is associated with a desired level of economic activity and money growth. In addition to the open market operations required to meet policy goals, there is a very substantial volume of "defensive" purchases and sales of securities. The Fed takes defensive operations in order to offset the large and frequent random variations in the volume of reserves that occur as part of the normal operations of the banking system.

Discount Rate

The money supply function shows that an increase in the discount rate leads to a lower money supply. A higher discount rate, a higher cost of borrowing from the Fed, will discourage the banks from incurring reserve deficiencies since the cost of doing so increases. Thus, when the cost of borrowing from the central bank increases, the banks will tend to maintain higher reserve positions and be more careful to avoid costly borrowing of reserves from the Fed. Thus, an increase in the discount rate will tend to reduce the money supply.

Discount borrowing was once the major tool of monetary policy used by the Fed. With the development of the Federal Funds—interbank market for reserves—market in the post-war period, discount borrowing became less important. A bank that finds itself short of reserves can borrow the reserves from other banks or sell assets, such as the banks own holding of securities, in order to increase its reserves position.

Nevertheless, the discount window (as the borrowing facility is called) is still important. First, many small banks do not have ready access to the Federal Funds market where the minimum size of transactions is \$5 million and most transactions exceed \$25 million. Second, if a bank is in economic difficulty, it may not be able to borrow from other banks in the Fed Funds market. Lenders will not want the risk. The Fed uses the discount window to provide support for banks in trouble. As a result, the Fed will manage access to the discount window.

The Fed usually keeps the discount rate below (about $\frac{1}{4}$ of a percentage point) the Fed Funds rate. This helps small banks that use the discount window. It could also be an inducement for banks to use the discount window except that the Fed limits access to the window. While the Fed does not have a formal rationing system for borrowed reserves, it discourages discount borrowing by monitoring individual banks' behavior and discussing with them any perceived excessive use of the discount window for

borrowing reserves. The Federal Reserve actively manages the privilege of borrowing at the discount rate. Banks and other depository institutions are discouraged from using discount borrowings on a regular basis. Borrowings are allowed to enable banks to meet temporary reserve deficiencies, but not as a continued source of reserve funds. Exceptions to this are seasonal borrowing by small banks which experience wide seasonal fluctuations in deposits or loans and the so-called extended credit facility, which the Fed uses to provide liquidity to banks in serious financial trouble. In any event, banks are not anxious to use the discount facility, because other banks may interpret such behavior as a signal that the bank is in financial difficulty.

Changes in the discount rate can lead to changes in the level of borrowings (B , borrowed reserves) and therefore to changes in the level of bank reserves. However, this tool of Fed policy is generally agreed to have little direct effect on the money supply. The reasons for this are that borrowed reserves are generally not a large source of bank reserves, and that the Fed can easily offset any change in borrowing by member banks through its open market operations. However, changes in the discount rate are thought to be important as signal of Fed policy.

Reserve Requirements

If reserve requirements are increased, a bank that held not excess reserves before the reserve requirement increase will find that its reserve holdings are deficient and will have to adjust its portfolio in order to satisfy the reserve requirement. An increase in reserve requirement reduces the money supply because it forces the banks to reduce their loans outstanding or to sell other assets (e.g., securities) in order to hold more reserves. In both instances, deposit balances will fall (as the public pays back loans or pays for securities) and the money supply declines.

Changes in the reserve requirement ratios—the amount of reserves that banks must hold for every dollar of deposits—are rarely used as a tool of monetary policy. Although the Fed has the ability to set reserve requirements within broad ranges dictated by Congress, changes in legal reserve requirements are made only infrequently. Reserve requirement ratio changes would be a very clumsy and inefficient tool for influencing the money supply. The Fed does not use reserve requirements for short-run control of the monetary system.

Reserve requirements were extended and made largely uniform by a major banking reform bill in 1980, the Depository Institutions Deregulatory and Monetary Control Act (DIDMCA). Reserve requirements were extended to all depository institutions, and the structure of reserve requirements for various deposit types was vastly simplified. The changes were made to improve the Fed's control over the money stock, which includes the liabilities of depository institutions other than commercial banks. However, since reserve balances at the Fed are not interest earning, the banks view them as an unfair burden that other financial institutions do not have. As a consequence, the Fed has, over time, reduced the level of reserve requirements.

Most recently, the reserve requirements on time and savings deposits (which had been 3%) were eliminated in December 1990 and the reserve requirement ratio on transaction accounts (including demand deposits) was reduced from 12% to 10% in April 1992. These steps were taken at this particular time to both ease the burden on the banks of reserve accounts that are non-interest earning and to influence monetary policy.

In 1990–91, the Federal Reserve took several distinct steps to loosen monetary policy as an anti-recessionary policy. Interest rates declined but the volume of new bank loans did not increase. The banks simply did not increase their loans to business and the money supply did not expand. The banks were not willing to expand their loan portfolios because of their concern with their own weak capital positions and the large amount of non-performing loans already in their portfolios.

This episode was called a credit crunch because businesses without access to other sources of borrowing were squeezed out of the markets. The credit crunch increased the severity of the recession. The Fed took the unusual step of reducing reserve requirements in order to offset the crunch. The changes in reserve requirements were being used at this time as a monetary policy action. The reduction in reserve requirement ratios directly increases bank profitability (the banks can hold interest earning assets instead of reserve deposits at the Fed) and was intended to induce the banks to make more loans to business and to help end the recession.

The changes in reserve requirements are also a reflection of changes in the way the Fed determines monetary policy. In the 1970s the Fed was strongly influenced by the monetarist point of view which gave great emphasis to the rate of growth of the money stock. As noted above, the 1980 banking legislation, DIDMCA, extended reserve requirements to all banking institutions and thus gave the Fed close control of *M1*. Soon thereafter, innovations in banking technology made *M1* a less important monetary aggregate than *M2*. However, most of *M2* is not subject to reserve requirements. In fact, since 1990 only transactions deposits are subject to reserve requirements. Thus, the Fed now places less emphasis on reserve requirements, the quantity of reserves available and direct control over the monetary aggregates. Although the growth of the money aggregates—particularly *M2*—is an important medium-term and long-term indicator of how monetary policy is working, the Fed now pays little attention to money growth in the short-run. The Fed now uses the target for the Fed Funds rate as its short-run policy target and relies on open market operations to influence interest rates. Of course, interest rate changes do have an impact on the growth of the money aggregates.

The policy tool or instrument most actively used by the Fed is open market operations. The operating target that guides policy in the short-run is no longer the rate of growth of a money aggregate like *M1* or *M2*. Instead, the operating target is the target for the Fed funds rate chosen by the FOMC. Of course, the FOMC still monitors money growth and it will influence their decision-making but money growth rates are too unpredictable to be used as a policy guide. The next issue to address is how policy actions effect the economy. The next section discusses the relationship between monetary policy and macroeconomic activity.

EFFECTS OF MONETARY POLICY

Changes in the availability of reserves affect the macroeconomy in two ways. First, the availability of reserves affects the banks' willingness and ability to create money and generally expand the amount of borrowing and credit. Second, open market sales (purchases) will push down (up) the price of government bonds, which raises (lowers) interest rates. Changes in monetary policy affect interest rates for additional reasons as well. For example, if reserves are in short supply, the banks will raise the price of borrowing; that is, loan rates go up. Thus, a tight monetary policy is associated with higher interest rates, which inhibit borrowing and the expansion of economic activity.

What are the magnitudes and timing of these effects of monetary policy on the economy? There is a rather clear consensus among economists concerning the answer to this question. In the short-run, monetary policy has a strong affect on real output; the effects on the inflation rate appear with a lag. In the long run, the economy tends to its natural output equilibrium and money growth determines the inflation rate.

Milton Friedman's description of the relationship of monetary policy to the economy is anecdotal but informative:

Over any long period of time ... Higher monetary growth means high inflation, and high inflation produces high interest rates ...

Over shorter periods ... the situation is more complex. The initial impact of slower monetary growth is to raise interest rates; of faster monetary growth to lower interest

rates. However, as the markets have learned the longer-term relationships, the duration of the short-term effects has become shorter and shorter ...

The rapid response of interest rates reflects market anticipations of the future effects of monetary growth. The economy itself does not respond as rapidly. The typical pattern is that higher monetary growth is followed some three to nine months later by higher growth in total spending. The higher spending, in turn, first takes the form of higher growth in output and employment, and only later still of higher inflation The total delay between monetary change and inflation is on the order of 18 to 24 months.

[*Newsweek*, December 29, 1980]

More specific answers are obtained from simulations of large-scale econometric models of the economy. The MPS model is maintained by the Federal Reserve Board staff and is used for forecasting and policy analysis. The model consists of a large number of estimated behavioral equations that are supposed to represent the macroeconomy. Policy effects are estimated from the comparison of a model simulation with historical (actual) policy to another simulation based on a simulation disturbed by a policy shock. The policy disturbance is introduced at one point and maintained at its new level. In addition, the simulations are carried out dynamically over a long period so that the effects of the policy change over time can be estimated.

Table 3 shows the effects of a permanent 1% increase in the level of M1 based on simulations of the 1981–85 period. The results show that a monetary expansion quickly affects interest rates. Both short-term and long-term interest rates fall dramatically in the quarter after the increase in the money supply. A robust real sector expansion begins in about a year and inflation effects do not appear until after 2 or 3 years. Higher interest rates begin to depress output after 4 years. The long-run characteristic of the model (not shown) is that a 1% increase in M1 increases the price level by 1% and leads to no permanent change in output.

TABLE 3

Monetary Policy Simulations with the MPS Model

	Quarters after disturbance:				
	<u>1</u>	<u>4</u>	<u>8</u>	<u>12</u>	<u>16</u>
Real GNP (in %)	0.3	1.2	1.1	0.5	-0.7
GNP deflator (in %)	0.0	0.1	0.6	1.4	2.1
Short-term interest rate, Commercial paper (in percentage points)	-2.3	-0.2	0.5	0.9	0.6
Long-term interest rate, Corporate bonds (in percentage points)	-0.8	-0.2	0.3	0.5	0.4

SOURCE: Flint Brayton and Eileen Mauskopf, "Structure and Uses of the MPS Quarterly Econometric Model of the United States," *Federal Reserve Bulletin*, February 1987.

In addition to effects on output, inflation and interest rates, monetary policy affects the exchange rate. The exchange rate effects of monetary policy are the consequence of changes in interest rates, which make dollar denominated assets more or less attractive to hold and changes in inflation and expected inflation which alter the relative value of the dollar. The effect of monetary policy on the economy through the exchange rate affects has become stronger as international trade becomes more important.

A looser monetary policy will lead to currency depreciation, which will affect the economy in two ways. First, the depreciation of the dollar will lead to higher dollar prices of imported goods. Since the ability, particularly in the short-run, to substitute domestically produced goods for imports is limited, this leads to an increase in prices. Thus, a looser monetary policy leads to some imported inflation. For example, the dollar depreciated in 1990 because the Federal Reserve eased monetary policy while the central banks of Germany, Japan and the U.K. maintained tight policy and high interest rates. The imported inflation that resulted kept the inflation rate from falling even as the domestic economy weakened. Second, a dollar depreciation makes U.S. exports more competitive in international markets. Thus, the balance of trade improves when monetary policy is loosened.

FISCAL POLICY

The Keynesian approach to macroeconomics gave a great deal of emphasis to fiscal policy. Thus, the academic responses to the tragedy of the Great Depression were often calls for fiscal policy activism. In addition, the Keynesian approach gained respectability in the first few post-World War II decades and policy-makers advocated the discretionary use of fiscal policy to stabilize the economy. By the 1970s and 1980s fiscal policy was no longer viewed as a viable tool for short-run macroeconomic policy management. The supply side approach to macroeconomics emphasizes the importance of a fiscal policy that supports long-term growth of the economy.

Contemporary discussion of fiscal policy are concerned with the long-run impact of government taxation and expenditure policies on economic growth. More specifically, the focus of current interest is the persistent government deficit. Most economists agree that large and persistent deficits should have a detrimental effect on the economy. There is less agreement about the extent to which the large deficits that persisted throughout the 1980s have had an effect on the economic performance of the U.S. economy.

There are a large number of fiscal policy tools available. Tax policy tools include both the personal and corporate income taxes. A change in the level of taxes collected will have a direct effect on disposable incomes and aggregate demand. Often tax changes are designed to change tax rates in a way that promotes one type of expenditure. For example, the proposed reduction in capital gains taxation would increase the returns to equity investments and hopefully increase the amount of such investments. Finally, tax changes are sometimes made on a temporary basis.

The expenditure policies of the government have a clear impact on aggregate demand. For many years there have been proposals to give the President greater discretion over Federal government expenditure through a line item veto. These proposals were never very popular in congress which guards its legislative powers closely. Thus, it was somewhat surprising that Congress passed a limited line item veto in March 1996. If the legislation survives judicial review (it may or may not), the line item veto will shift some power to the President and could have a long-run influence on government expenditures.

Since the budget deficit and increasing government expenditures have become a widely acknowledged problem in the U.S., there have been numerous attempts to reform the budgeting process and bring discipline to fiscal decision making.

The Balanced Budget and Emergency Deficit Control Act (Graham-Rudman-Hollings Act) of 1985 established new procedures for deficit control. It set specific targets for reduction of the deficit and set out means for assuring the outcome. The Graham Rudman target, that the deficit be eliminated by 1991, was never met and the required sequestration of funds was always avoided. However, the act was not a total failure; the legislative structure and budget goals have had some effect on the Congress. There is a gradually emerging willingness to bring expenditure growth under control.

In the Fall of 1990, administration and Congressional leaders negotiated a new set of budget procedures and deficit reduction plans. The Budget Enforcement Act of 1990 eliminates fixed deficit targets. The act specifies that the Congress set spending limits and includes procedures that are designed to assure that the targets are met.

Throughout the 1980s, the focus of the budget discussions has been how to bring the Federal budget into balance over a long period of time. The Budget Enforcement Act of 1990 set fiscal policy for the next five years and precluded any changes except emergency spending. The congressional Budget Office reported that:

As a result, by 1995, the total federal deficit is likely to fall below \$100 billion for the first time in 15 years and below 1 percent of gross national product for the first time in 20 years.

Of course, these targets were revised soon after they were put in place because unforeseen growth of non-discretionary expenditures (particularly Medicare and Medicaid) and the lingering effects on tax revenues of the 1990–91 recession necessitated changes. The 1995 Federal deficit was a bit more than \$160 billion, still a little more than 2 percent of GDP. Throughout 1995, the Congress and the President fought a long political battle concerning the budget and deficit reduction plans that culminated with a short shut down of the Federal government. With vigorous economic growth in 1996, it is likely that the budget deficit as a fraction of GDP was smaller than it had been in any year since 1974.

Fiscal Policy Activism: 1960s and 1970s

It was only in the post-World War II era that the Federal government assumed the responsibility for management of the economy. The Employment Act of 1946 sets out an objective for economic policy—maintaining full employment. The act also established the Joint Economic Committee of the Congress and the Council of Economic Advisers, which together made macroeconomic policy an explicit government function for the first time. The goals of policy management were reiterated in the Full Employment and Balanced Growth Act of 1978 (the Humphrey-Hawkins Act).

The first administration to explicitly advocate fiscal policy activism was the Kennedy administration. The Kennedy inauguration took place at the tail end of a mild 10-month recession which had followed a relatively short (two-year) expansion. Thus, there was substantial sentiment in favor of a fiscal policy expansion. However, the Congress was more conservative in its approach and broad tax cuts were not enacted until 1964. The Kennedy-Johnson tax cuts proved the efficacy of fiscal policy. The average annual unemployment rate for the years 1958 through 1963 was 5.8%. It fell to 5.0% in 1964 and 4.4% in 1965.

The confidence in the efficacy of fiscal policy disappeared very quickly when future attempts to use fiscal policy for macroeconomic stabilization were less successful.

The Vietnam war led to the expansion of aggregate demand in the late 1960s and inflationary pressures began to build as the economy pressed against capacity constraints. The average unemployment rate reached 3.4% in 1969 and the CPI inflation rate was up to 6.2%. That might not seem like a great deal of inflation but the average annual CPI inflation rate for the first half of the 1960s was only 1.2% per year. A restrictive fiscal policy seemed the appropriate move to take. With much reluctance and delay the Congress finally enacted a one-year income tax surcharge in 1968. The increase in taxes reduced disposable income and in theory was supposed to reduce consumer expenditures and reduce demand pressures in the economy. However, by most indication, consumers were barely affected. It seems that the consumer sector understood that a temporary tax increase has very little impact on their permanent or lifetime well being and therefore there was no necessity to adjust expenditure patterns.

In 1975, in the midst of the long and deep recession that followed the first OPEC oil price shock, the congress passed an income tax rebate. The oil shock in October 1973 pushed the inflation rate up and inflation was the primary concern of policy-makers throughout 1974. Even as the recession emerged in mid-and late 1974, the President was calling for a tax increase to hold down demand and inflation. An expansionary fiscal policy was not proposed until January 1975. It took some time for the legislation to get through Congress and it was not until May 1975 that the rebates were paid. The personal savings rate increased for several months and consumer expenditure patterns did not begin to adjust to the increase in disposable income until year-end. By that time the expansion in economic activity had already begun.

These two episodes illustrate two major problems with the use of fiscal policy for macroeconomic stabilization. First, it is difficult to design the appropriate size and type of policy change because the responses of economic agents to a policy change is not always predictable. Consumer and business responses to policy changes, particularly tax changes, will vary from episode to episode. The experience with the tax surcharge in 1968 showed that the response of the macroeconomy to a fiscal policy change is hard to predict.

Second, the delays inherent in putting a policy change in place often make it impossible to make a fiscal policy response at the appropriate time. The experience of the 1975 tax rebate showed that it might be impossible to time an appropriate fiscal policy response to a cyclical downturn.

Policy Lags

Generally speaking, the lags in the effect of any policy change on the economy can be characterized into an inside lag and an outside lag. The inside lag is the time it takes to formulate and execute a policy change. It can be separated into three components:

Recognition lag – the time between an economic disturbance and the recognition by policy-makers that a policy response is required. The recognition lag occurs because a) forecasting is not reliable enough to use as a basis for a major change in policy; b) data on macroeconomic developments are not immediately available and do not immediately present an unambiguous picture. Estimates of the recognition lag suggest that it is usually about 3 to 6 months.

Decision lag – the time it takes for the policy-makers to make a decision about policy. For monetary policy the decision lag is relatively short (about one month) because there are regular meetings of the Federal Open Market Committee. For fiscal policy, however, the decision lag can be lengthy because congressional action is necessary.

Implementation lag – the time it takes for a policy change to be put in place. For monetary policy the implementation lag is very short; open market operations in response to an FOMC meeting decision will start within hours. For fiscal policy the lag can again be several months. A task as simple as promulgating and putting into action new withholding rates can take at least several weeks.

All in all the inside lags are likely to take at least six months. Once policy has been changed there is an outside lag or the time it takes for the policy change to effect the macroeconomy. A fiscal policy change leads to a multiplier effect, which takes several quarters to get going, and the full effect is not felt for about a year.

In the case of an economy entering a slowdown, it would typically take 3 months after the peak in activity to recognize that a recession has begun. The decision lag is likely to be at least another 2 or 3 months. The implementation lag for fiscal policy is another 2 or 3 months. Finally, several months pass before the multiplier process has a significant effect on aggregate demand. Thus, the fiscal policy lag is about a year. But in the post-war period, recession has on average lasted only about 11 months. Thus, fiscal policy is hardly a useful tool for short-run stabilization policy.

Measuring Fiscal Policy

It is not correct to say that an increasing deficit is an indication that fiscal policy is expansionary. The deficit could be larger because of changes in government expenditure and revenue that occur because of economic conditions. Thus, the deficit widens in recession even without any change in the stance of fiscal policy. This occurs because expenditures (particularly, transfer payments) increase automatically in a recession and revenues decline as income falls.

A better measure of fiscal policy and the deficit is one that removes the effects of cyclical change on the deficit:

The cyclically adjusted budget is an estimate of what the Federal budget would be after removing the automatic responses of receipts and expenditures to economic fluctuations.

The calculation starts with a measure of trend Gross Domestic Product in constant dollars. Then government expenditures and revenues that would occur at the trend level of GDP are calculated. Changes in the cyclically adjusted deficit then represent true change in the fiscal policy stance of the government.

EFFECTS OF FISCAL POLICY

The magnitude and timing of the effect of a change in fiscal policy on the economy can best be estimated from a large-scale econometric model. Simulations from different models will differ but there is some general agreement about the effects of policy. We will present simulation results from one particular model, the MPS model used by the Federal Reserve Board staff, the same model that was used for the monetary policy simulation earlier.

Table 4 shows three simulations with versions of the MPS model which illustrate how the effects of fiscal policy interact with monetary policy. In each case the control simulation is an historical dynamic simulation for the period 1981–85. The disturbed simulation introduces a permanent increase in real government expenditures equal to 1 percent of real GNP.

Table 4Fiscal Policy Simulations with the MPS Model

	Quarters after disturbance:				
	<u>1</u>	<u>2</u>	<u>4</u>	<u>8</u>	<u>12</u>
1. Exogenous supply and fixed interest rate (Simple Keynesian)					
Real GNP growth	1.3	1.7	2.3	3.1	3.7
2. Exogenous supply and fixed money supply (IS-LM)					
Real GNP growth	1.2	1.4	1.4	1.2	1.1
Federal Funds rate	1.0	1.2	1.1	0.8	0.9
3. Full Model and fixed money supply (Total Supply and Demand)					
Real GNP growth	1.3	1.4	1.3	1.0	0.2
Federal Funds rate	1.0	1.4	1.5	1.7	2.2
GNP deflator growth	0.0	0.1	0.3	0.9	1.8

SOURCE: Flint Brayton and Eileen Mauskopf, "Structure and Uses of the MPS quarterly Econometric Model of the United States," *Federal Reserve Bulletin*, February 1987.

The first case corresponds to a simple Keynesian real sector multiplier model. Interest rate effects of the changes induced by the fiscal expansion are assumed not to exist. Interest rates are held constant and the money supply increases to satisfy any increase in money demand. In addition, supply is viewed as exogenous; there are not capacity constraints.

It takes a few quarters from the start of the policy disturbance for the multiplier process to get started. The multiplier effect is only 1.3 in the first quarter but it is 2.3 at the end of one year, and 3.7 after three years. Thus, the full multiplier effect is quite large, over three, and it takes about two years for the multiplier effects to occur. The multiplier is large because the model is simple—there are no changes in interest rates or prices.

The second case shows that the multiplier effects in a simple Keynesian model are "crowded out" by changes in the financial markets. This case corresponds to an *IS-LM* model framework. A fiscal expansion leads to increased economic activity that increases the demand for money. In the absence of any accommodating expansion of monetary policy, interest rates will rise and hold back the multiplier process. This is called "transactions crowding out." Comparison of cases one and two for the MPS model indicates that transactions crowding out is sizable. In the second case the multiplier effect on real GNP peaks at 1.4 after a year and then falls. The effect of the fiscal expansion on the Federal Funds rate is also shown.

The third case represents a simulation of the full MPS model. There is no monetary accommodation so the money supply is fixed but all the interest rate, price and supply effects that are started by the fiscal

expansion are allowed to come into play. This simulation corresponds to the short-run and long-run analysis of the total supply and demand model.

The fiscal expansion has a short-run effect on real GNP but it is not large and peaks within a year. The effects on interest rates and inflation become more important and by the end of three years the real GNP effect is barely more than zero. The inflation effects do not appear for several quarters and then grow larger. In the long run, inflation and interest rates are substantially higher and real output is virtually unaffected.