



Bonga University
College of Business and Economics
Department of Economics
Macroeconomics-II Reading Material

By
Birku Andualem (MSc.)

March, 2023

Table of Contents

Chapter One.....	1
Behavioral Foundations: Theories of Consumption	1
1.1 The Keynes Consumption function	1
1.2 The Early Empirical Successes.....	2
1.3 secular Stagnation, Simon Kuznets and the consumption puzzle	3
1.4 Irving Fisher’s Model.....	4
1.5 Modigliani’s Life cycle Hypothesis model	10
1.6 Friedman’s Permanent Income Hypothesis.....	12
1.7 Hall’s Random-Walk Hypothesis	13
Chapter Two: Theories of Investment	15
2.2 Types of Investment.....	15
1.2.1 Business Fixed Investment.....	15
2.3 Neoclassical Model	16
2.4 The Keynesian Explanation	26
2.5 The Stock Market and Tobin’s q	28
Chapter Three: Money Demand and Money Supply	30
3.1 What is money?.....	30
3.1.1 The Functions of Money	30
3.1. Money supply and monetary expansion mechanism	31
3.2 Money demand –portfolio and transaction theories	36
Chapter Four: Theories of Labor Market.....	41
4.1 Definition of unemployment	41
4.3 Causes and Types of Unemployment.....	44
4.4 Classical and Neoclassical Microeconomic model of the labour market.....	51
4.5 New – Keynesian view of unemployment	53
Chapter Five: Macroeconomic Policy Debates.....	57
5.2 Ignorance, Expectations, and the Lucas Critique	60
5.3 Should Policy Be Conducted by Rule or by Discretion?	62
Chapter Six: Macroeconomic Growth Model.....	68
6.1 Introduction: Modeling Economic Growth	68
6.2 Theories of economic Growth	69

6.2.1 The Harrod–Domar Model.....	69
6.2.2 Neoclassical Growth Model.....	73
6.3 The Endogenous Growth Theory	88

Chapter One: Behavioral Foundations: Theories of Consumption

✂ This chapter surveys the most prominent work on consumption.

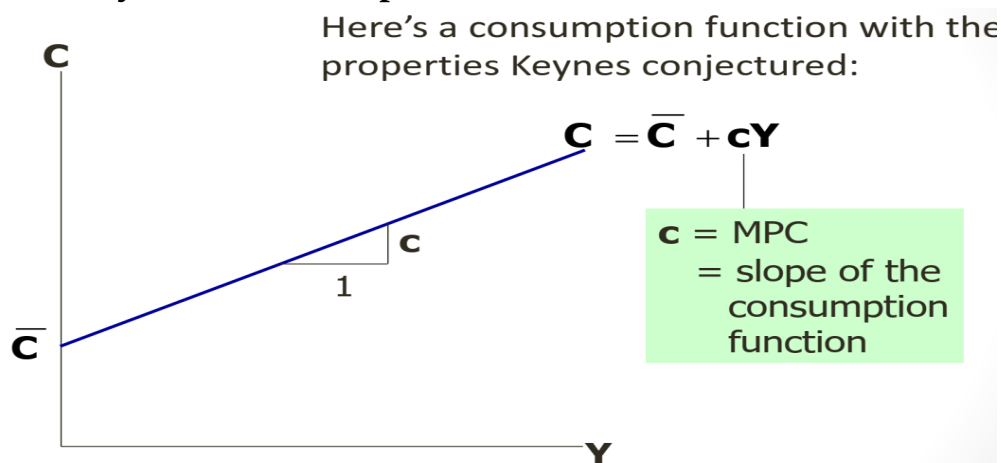
- John Maynard Keynes: consumption and current income
- Irving Fisher: Inter-temporal Choice
- Franco Modigliani: the Life-Cycle Hypothesis
- Milton Friedman: the Permanent Income Hypothesis
- Robert Hall: the Random-Walk Hypothesis
- David Laibson: the pull of instant gratification

1.1 The Keynes Consumption function

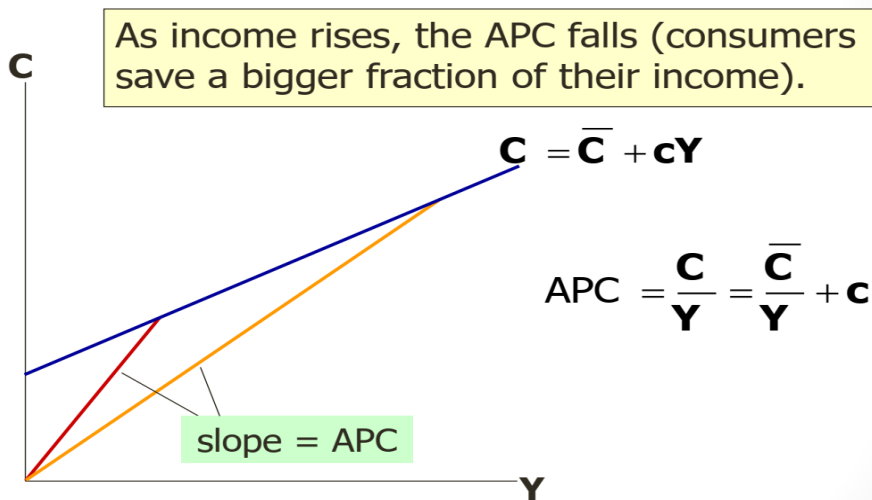
Keynes's Conjectures

1. $0 < MPC < 1$
2. **Average propensity to consume (APC)** falls as income rises.
($APC = C/Y$)
3. Income is the main determinant of consumption.

The Keynesian Consumption Function



The Keynesian Consumption Function



1.2 The Early Empirical Successes

Results from Early Studies

- Households with higher incomes:

- Consume more

$\Rightarrow MPC > 0$

- Save more

$\Rightarrow MPC < 1$

- Save a larger fraction of their income

$\Rightarrow APC \downarrow$ as $Y \uparrow$

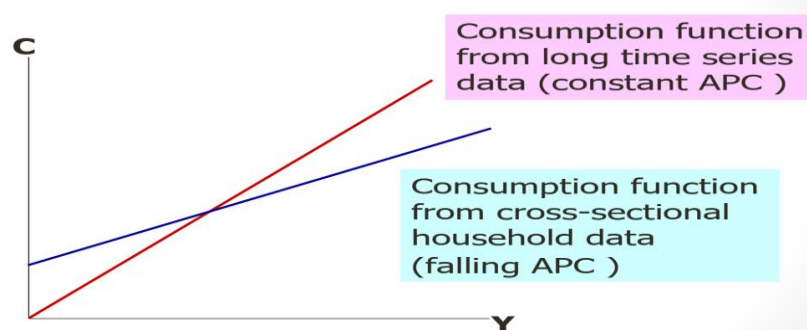
Very strong correlation between income and consumption

$\Rightarrow$ income seemed to be the main determinant of consumption

Problems for the Keynesian Consumption Function

- Based on the Keynesian consumption function, economists predicted that C would grow more slowly than Y over time. This prediction did not come true:
 - As incomes grew, the APC did not fall, and C grew just as fast.
 - Simon Kuznets showed that C/Y was very stable in long time series data.

The Consumption Puzzle



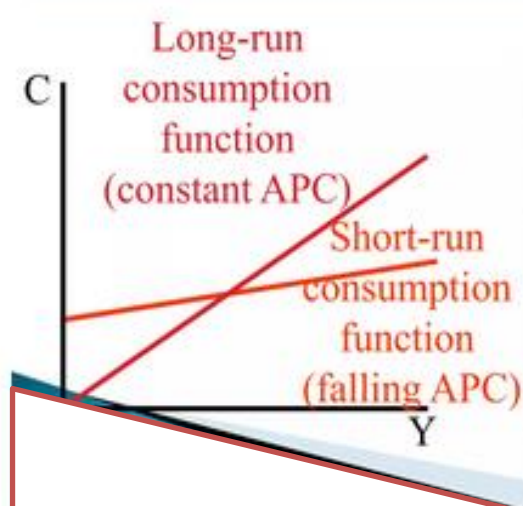
1.3 secular Stagnation, Simon Kuznets and the consumption puzzle

During World War II, on the basis of Keynes' consumption function, economists predicted that the economy would experience what they called *secular stagnation*-- a long depression of infinite duration-- unless fiscal policy was used to stimulate aggregate demand. It turned out that the end of the war did not throw the U.S. into another depression, but it did suggest that Keynes' conjecture that the average propensity to consume would fall as income rose appeared not to hold.

Simon Kuznets constructed new aggregate data on consumption and investment dating back to 1869 and whose work would later earn a Nobel Prize. He discovered that the ratio of consumption to income was stable over time, despite large increases in income; again, Keynes' conjecture was called into question.

Consumption Puzzle

The failure of the secular-stagnation hypothesis and the findings of Kuznets both indicated that the average propensity to consume is fairly constant over time. This presented a puzzle: why did Keynes' conjectures hold up well in the studies of household data and in the studies of short time-series, but fail when long time series were examined?



Studies of household data and short time-series found a relationship between consumption and income similar to the one Keynes conjectured--this is called the **short-run consumption function**. But, studies using long time-series found that the APC did not vary systematically with income--this relationship is called the **long-run consumption function**.

1.4 Irving Fisher's Model

- ✓ The basis for much subsequent work on consumption.
- ✓ Assumes consumer is forward-looking and chooses consumption for the present and future to maximize lifetime satisfaction.
- ✓ Consumer's choices are subject to an intertemporal budget constraint, a measure of the total resources available for present and future consumption

The basic two-period model

- Period 1: the present
- Period 2: the future
- Notation

Y_1 is income in period 1

Y_2 is income in period 2

C_1 is consumption in period 1

C_2 is consumption in period 2

$S = Y_1 - C_1$ is saving in period 1 ($S < 0$ if the consumer borrows in period 1)

Derivation of the inter-temporal budget Constraint

- Period 2 budget constraint:

$$\begin{aligned} C_2 &= Y_2 + (1+r)S \\ &= Y_2 + (1+r)(Y_1 - C_1) \end{aligned}$$

- Rearrange to put C terms on one side and Y terms on the other:

$$(1+r)C_1 + C_2 = Y_2 + (1+r)Y_1$$

- Finally, divide through by $(1+r)$:

The intertemporal budget constraint

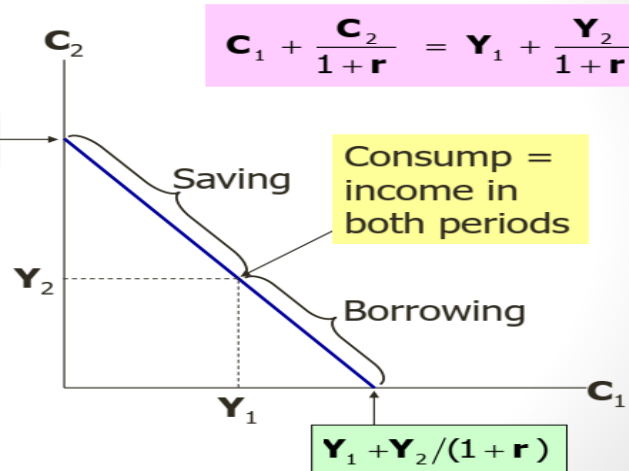
$$C_1 + \frac{C_2}{1+r} = Y_1 + \frac{Y_2}{1+r}$$

present value of
lifetime consumption

present value of
lifetime income

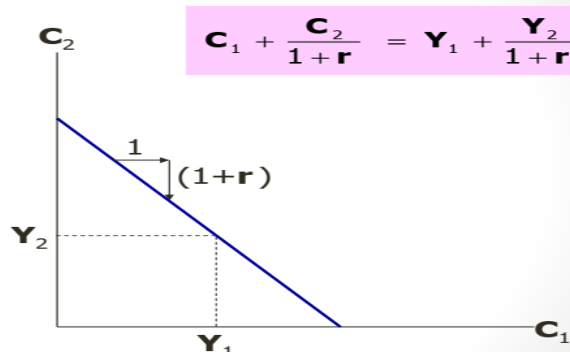
The intertemporal budget constraint

The budget constraint shows all combinations of C_1 and C_2 that just exhaust the consumer's resources.



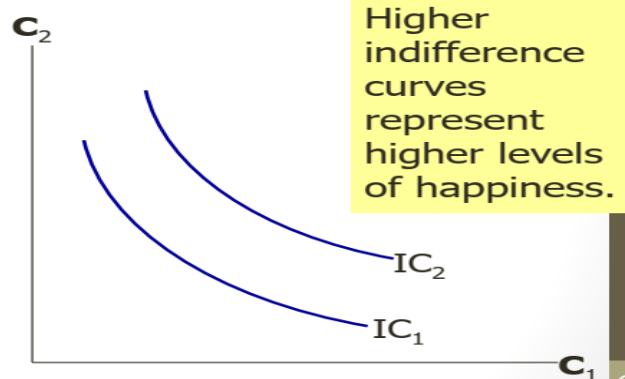
The intertemporal budget constraint

The slope of the budget line equals $-(1+r)$



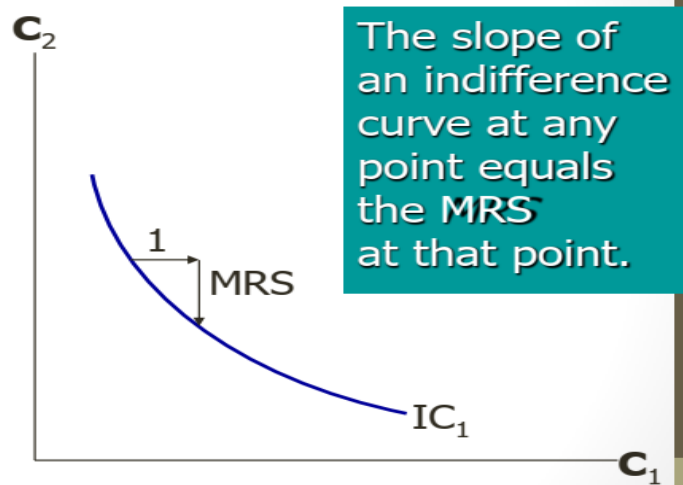
Consumer preferences

An indifference curve shows all combinations of C_1 and C_2 that make the consumer equally happy.

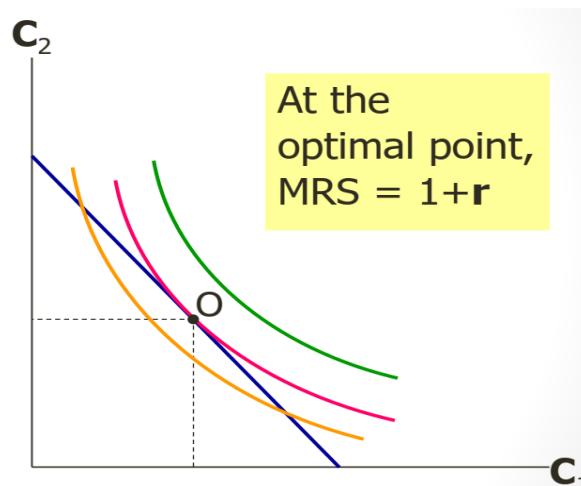


Consumer preferences

Marginal rate of substitution (MRS): the amount of C_2 consumer would be willing to substitute for one unit of C_1 .



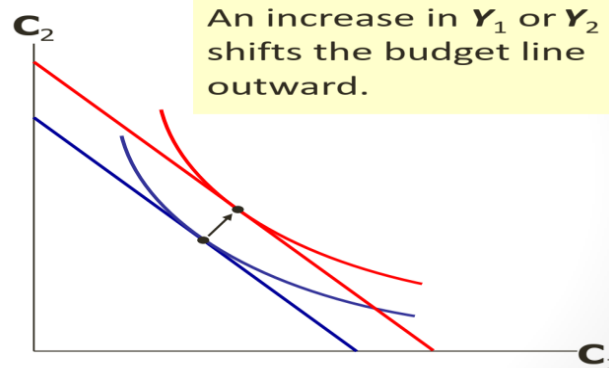
The optimal (C_1, C_2) is where the budget line just touches the highest indifference curve.



How C responds to changes in Y

Results:
Provided they are both normal goods, C_1 and C_2 both increase,

...regardless of whether the income increase occurs in period 1 or period 2.



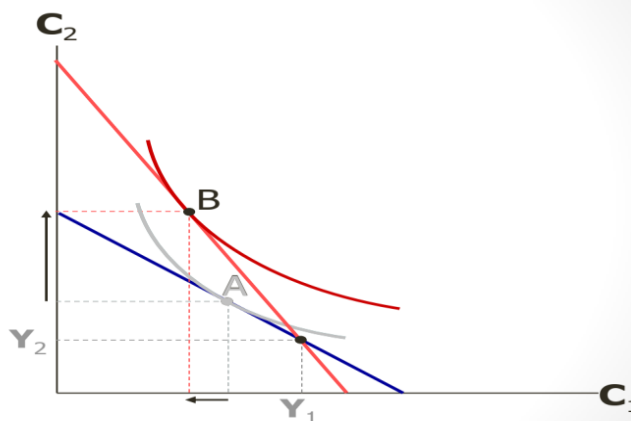
Keynes vs. Fisher

- Keynes:
current consumption depends only on current income
- Fisher:
current consumption depends only on the present value of lifetime income;
the timing of income is irrelevant
because the consumer can borrow or lend between periods.

How C responds to changes in r

An increase in r pivots the budget line around the point (Y_1, Y_2) .

As depicted here, C_1 falls and C_2 rises. However, it could turn out differently...



How C responds to changes in r

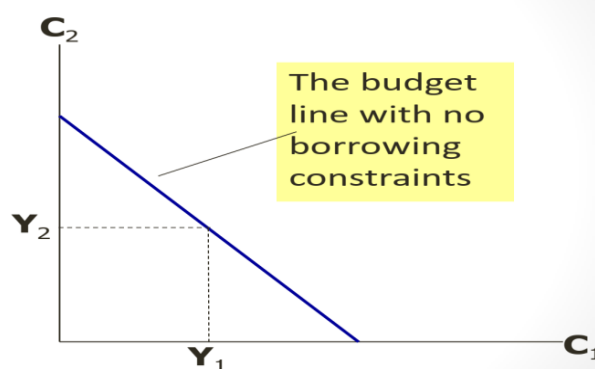
- income effect
If consumer is a saver, the rise in r makes him better off, which tends to increase consumption in both periods.
- substitution effect
The rise in r increases the opportunity cost of current consumption, which tends to reduce C_1 and increase C_2 .
- Both effects $\Rightarrow \uparrow C_2$.
Whether C_1 rises or falls depends on the relative size of the income & substitution effects.

Constraints on borrowing

- In Fisher's theory, the timing of income is irrelevant because the consumer can borrow and lend across periods.
- Example: If consumer learns that her future income will increase, she can spread the extra consumption over both periods by borrowing in the current period.

- However, if consumer faces **borrowing constraints** (aka “liquidity constraints”), then she may not be able to increase current consumption and her consumption may behave as in the Keynesian theory even though she is rational & forward-looking

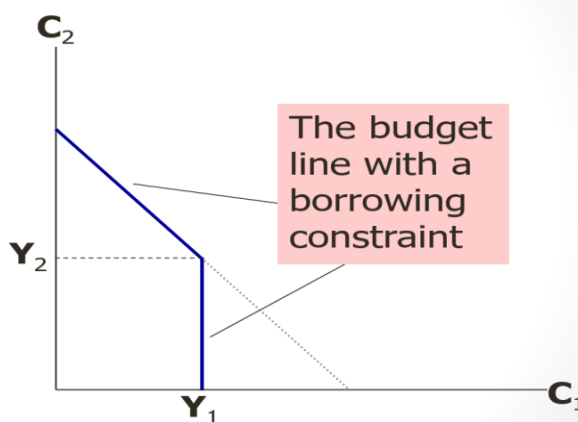
Constraints on borrowing



Constraints on borrowing

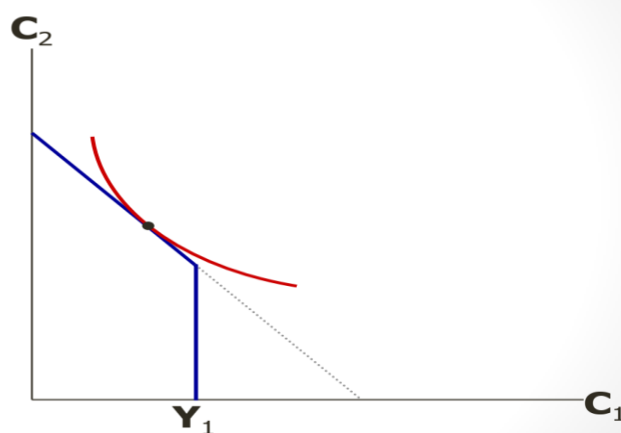
The borrowing constraint takes the form:

$$C_1 \leq Y_1$$



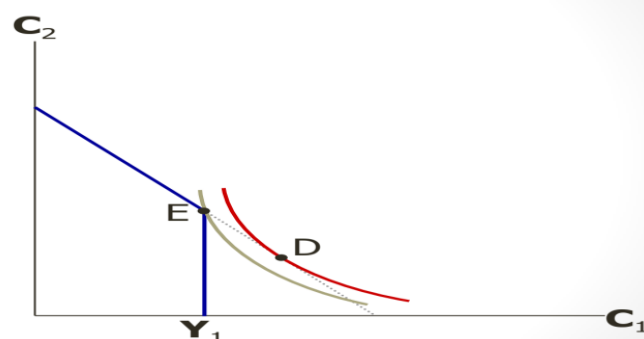
Consumer optimization when the borrowing constraint is **not binding**

The borrowing constraint is not binding if the consumer's optimal C_1 is less than Y_1 .



Consumer optimization when the borrowing constraint **is binding**

The optimal choice is at point D. But since the consumer cannot borrow, the best he can do is point E.



1.5 Modigliani's Life cycle Hypothesis model

The Life-Cycle Hypothesis

- due to Franco Modigliani (1950s)
- Fisher's model says that consumption depends on lifetime income, and people try to achieve smooth consumption.
- The LCH says that income varies systematically over the phases of the consumer's "life cycle," and saving allows the consumer to achieve smooth consumption.

The Life-Cycle Hypothesis

- The basic model:
 W = initial wealth
 Y = annual income until retirement (assumed constant)
 R = number of years until retirement
 T = lifetime in years
- Assumptions:
 - zero real interest rate (for simplicity)
 - consumption-smoothing is optimal

The Life-Cycle Hypothesis

- Lifetime resources = $W + RY$
- To achieve smooth consumption, consumer divides her resources equally over time:

$$C = (W + RY)/T, \text{ or}$$

$$C = \alpha W + \beta Y$$

where

$\alpha = (1/T)$ is the marginal propensity to consume out of wealth

$\beta = (R/T)$ is the marginal propensity to consume out of income

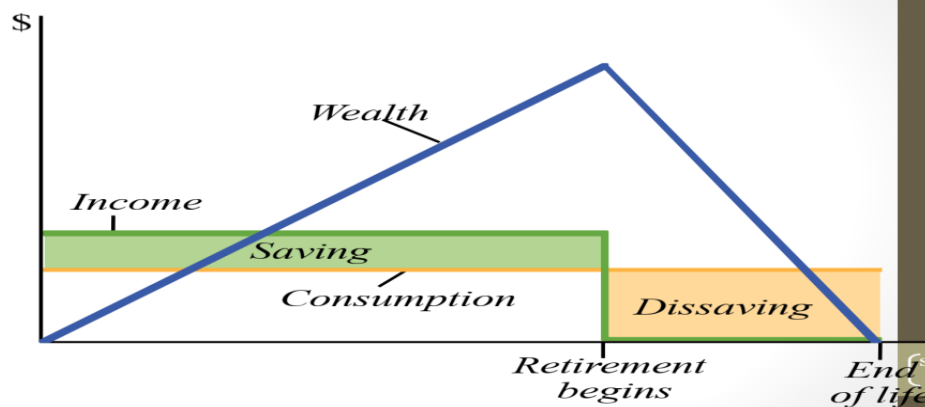
Implications of the Life-Cycle Hypothesis

The Life-Cycle Hypothesis can solve the consumption puzzle:

- The APC implied by the life-cycle consumption function is
 $C/Y = \alpha(W/Y) + \beta$
- Across households, wealth does not vary as much as income, so high income households should have a lower APC than low income households.
- Over time, aggregate wealth and income grow together, causing APC to remain stable.

Implications of the Life-Cycle Hypothesis

The LCH implies that saving varies systematically over a person's lifetime.



1.6 Friedman's Permanent Income Hypothesis

- Due to Milton Friedman (1957)
- The PIH views current income Y as the sum of two components: permanent income Y^P (average income, which people expect to persist into the future) transitory income Y^T (temporary deviations from average income).

The PIH can solve the consumption puzzle:

- The PIH implies $APC = C/Y = aY^P/Y$
- To the extent that high income households have higher transitory income than low income households, the APC will be lower in high income households.
- Over the long run, income variation is due mainly if not solely to variation in permanent income, which implies a stable APC.

PIH vs. LCH

- ✓ In both, people try to achieve smooth consumption in the face of changing current income.
- ✓ In the LCH, current income changes systematically as people move through their life cycle.
- ✓ In the PIH, current income is subject to random, transitory fluctuations.
- ✓ Both hypotheses can explain the consumption puzzle.

1.7 Hall's Random-Walk Hypothesis

- Due to Robert Hall (1978)
- based on Fisher's model & PIH, in which forward looking consumers base consumption on expected future income.
- Hall adds the assumption of **rational expectations**, that people use all available information to forecast future variables like income.
- If PIH is correct and consumers have rational expectations, then consumption should follow a **random walk**: changes in consumption should be unpredictable.
- A change in income or wealth that was anticipated has already been factored into expected permanent income, so it will not change consumption.
- Only unanticipated changes in income or wealth that alter expected permanent income will change consumption.

Implication of the R-W Hypothesis

- ✓ If consumers obey the PIH and have rational expectations, then policy changes will affect consumption *only if they are unanticipated*.

The Psychology of Instant Gratification

- ✓ Theories from Fisher to Hall assume that consumers are rational and act to maximize lifetime utility.
- ✓ Recent studies by David Laibson and others consider the psychology of consumers.
- ✓ Consumers consider themselves to be imperfect decision makers.
- ✓ Laibson: The “pull of instant gratification” explains why people don't save as much as a perfectly rational lifetime utility maximizer would save.

Two Questions and Time Inconsistency

1. Would you prefer

- (A) a candy today, or
- (B) Two candies tomorrow?

2. Would you prefer

- (A) a candy in 100 days, or
- (B) Two candies in 101 days?

✎ In studies, most people answered A to question 1, and B to question 2.

- ✓ A person confronted with question 2 may choose B. 100 days later, when he is confronted with question 1, the pull of instant gratification may induce him to change his mind.

Summing up

✎ Keynes suggested that consumption depends primarily on current income.

✎ Recent work suggests instead that consumption depends on

- current income
- expected future income
- Wealth
- interest rates

📖 Economists disagree over the relative importance of these factors and of borrowing constraints and psychological factors.

Chapter Two: Theories of Investment

While spending on consumption goods provides utility to households today, spending on investment goods is aimed at providing a higher standard of living at a later date. Investment is the component of GDP that links the present and the future.

Investment spending plays a key role not only in long-run growth but also in the short-run business cycle because it is the most volatile component of GDP. When expenditure on goods and services falls during a recession, much of the decline is usually due to a drop in investment.

2.2 Types of Investment

There are three types of investment spending.

1. **Business fixed investment** includes the equipment and structures that businesses buy to use in production.
2. **Residential investment** includes the new housing that people buy to live in and that landlords buy to rent out.
3. **Inventory investment** includes those goods that businesses put aside in storage, including materials and supplies, work in process, and finished goods.

Investment usually falls during recessions.

Now we build models of each type of investment to explain the fluctuations. The models will shed light on the following questions:

- Why investment is negatively related to the interest rate?
- What causes the investment function to shift?
- Why does investment rise during booms and fall during recessions?

1.2.1 Business Fixed Investment

The largest piece of investment spending, accounting for about three-quarters of the total, is business fixed investment. The term “business” means that these investment goods are bought by firms for use in future production. The term “fixed” means that this spending is for capital that will stay put for a while, as opposed to inventory investment, which will be used or sold within a

short time. Business fixed investment includes everything from office furniture to factories, computers to company cars.

2.3 Neoclassical Model

The standard model of business fixed investment is called the **neoclassical model of investment**. The neoclassical model examines the benefits and costs to firms of owning capital goods. The model shows how the level of investment (the addition to the stock of capital) is related to the marginal product of capital, the interest rate, and the tax rules affecting firms.

To develop the model, imagine that there are two kinds of firms in the economy. *Production firms* produce goods and services using capital that they rent. *Rental firms* make all the investments in the economy; they buy capital and rent it out to the production firms. Most firms in the real world perform both functions. For simplicity and clarity of analysis, let separate these two activities by imagining that they take place in different firms.

The Rental Price of Capital

Let's first consider the typical production firm. The firm rents capital at a rental rate **R** and sells its output at a price **P**; the real cost of a unit of capital to the production firm is **R/P**. The real benefit of a unit of capital is the marginal product of capital MPK, the extra output produced with one more unit of capital. The marginal product of capital declines as the amount of capital rises: the more capital the firm has, the less an additional unit of capital will add to its output. It concluded that, to maximize profit, the firm rents capital until the marginal product of capital falls to equal the real rental price.

For the reasons discussed above, the marginal product of capital determines the demand curve. The demand curve slopes downward because the marginal product of capital is low when the level of capital is high. At any point in time, the amount of capital in the economy is fixed, so the supply curve is vertical. The real rental price of capital adjusts to equilibrate supply and demand.

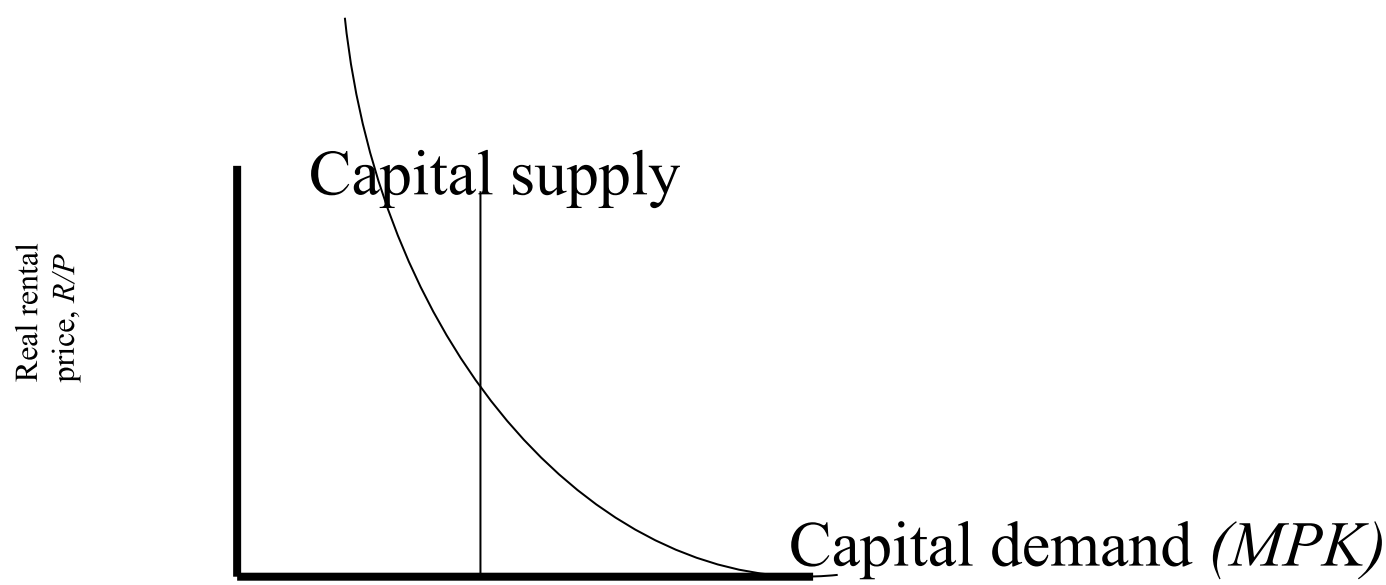


Figure 1 The Rental Price of Capital. The real rental price of capital adjusts to equilibrate the demand for capital (determined by the marginal product of capital) and the fixed supply.

What are the variable that influence the equilibrium rental price?

To see these let's consider a particular production function. Many economists consider the Cobb–Douglas production function a good approximation of how the actual economy turns capital and labor into goods and services. The Cobb–Douglas production function is $Y = AK^\alpha L^{1-\alpha}$ where Y is output, K is capital, L is labor, A is a parameter measuring the level of technology, and α is a parameter between zero and one that measures capital's share of output. The marginal product of capital for the Cobb–Douglas production function is

$$MPK = \alpha A(L/K)^{1-\alpha}$$

Because the real rental price R/P equals the marginal product of capital in equilibrium, we can write

$$R/P = \alpha A(L/K)^{1-\alpha}$$

The expression identifies the variables that determine the real rental price. It shows the following:

- The lower the stock of capital, the higher the real rental price of capital.
- The greater the amount of labor employed, the higher the real rental price of capital.
- The better the technology, the higher the real rental price of capital.

Events that reduce the capital stock (an earthquake), or raise employment (an expansion in aggregate demand), or improve the technology (a scientific discovery) raise the equilibrium real rental price of capital.

The Cost of Capital

Next consider the rental firms. These firms like *car-rental* companies, merely buy capital goods and rent them out. Because our goal is to explain the investments made by the rental firms, we begin by considering the benefit and cost of owning capital. The benefit of owning capital is the revenue earned by renting it to the production firms. The rental firm receives the real rental price of capital R/P for each unit of capital it owns and rents out.

The cost of owning capital is more complex. For each period of time that it rents out a unit of capital, the rental firm bears three costs:

1. When a rental firm borrows to buy a unit of capital, it must pay interest on the loan. If P_k is the purchase price of a unit of capital and i is the nominal interest rate, then iP_k is the interest cost.
NB: This interest cost would be the same even if the rental firm did not have to borrow: if the rental firm buys a unit of capital using cash on hand, it loses out on the interest it could have earned by depositing this cash in the bank. In either case, the interest cost equals iP_k .
2. While the rental firm is renting out the capital, the price of capital can change. If the price of capital falls, the firm loses, because the firm's asset has fallen in value. If the price of capital rises, the firm gains, because the firm's asset has risen in value. The cost of this loss or gain is $-\Delta P_k$. (The minus sign is here because we are measuring costs, not benefits.)
3. While the capital is rented out, it suffers wear and tear, called **depreciation**. If δ is the rate of depreciation, the fraction of capital's value lost per period because of wear and tear, then the dollar cost of depreciation is δP_k .

Therefore, the total cost of renting out a unit of capital for one period is

$$\begin{aligned}\text{Cost of Capital} &= iP_k - \Delta P_k + \delta P_k \\ &= P_k(i - \Delta P_k/P_k + \delta)\end{aligned}$$

The cost of capital depends on the *price of capital*, the *interest rate*, the rate at which capital prices are changing, and the depreciation rate.

For example, consider the cost of capital to a car-rental company. The company buys cars for \$10,000 each and rents them out to other businesses. The company faces an interest rate i of 10 percent per year, so the interest cost iP_k is \$1,000 per year for each car the company owns. Car prices are rising at 6 percent per year, so, excluding wear and tear, the firm gets a capital gain ΔP_k of \$600 per year. Cars depreciate at 20 percent per year, so the loss due to wear and tear δP_k is \$2,000 per year. Therefore, the company's cost of capital is

Cost of Capital = \$1,000 - \$600 + \$2,000 = \$2,400. The cost to the car-rental company of keeping a car in its capital stock is \$2,400 per year.

To make the expression for the cost of capital simpler and easier to interpret, we assume that the price of capital goods rises with the prices of other goods. In this case, $\Delta P_k/P_k$ equals the overall rate of inflation π . Because $i - \pi$ equals the real interest rate r , we can write the cost of capital as **Cost of Capital = $P_k(r + \delta)$** .

This equation states that the cost of capital depends on

- The price of capital,
- The real interest rate, and
- The depreciation rate

Finally, we want to express the cost of capital relative to other goods in the economy. The **real cost of capital**, the cost of buying and renting out a unit of capital measured in units of the economy's output, is

$$\text{Real Cost of Capital} = (P_k/P)(r + \delta)$$

This equation states that the real cost of capital depends on the relative price of a capital good P_k/P , the real interest rate r , and the depreciation rate δ .

The Determinants of Investment

Now consider a rental firm's decision about whether to increase or decrease its capital stock. For each unit of capital, the firm earns real revenue R/P and bears the real cost $(P_k/P)(r + \delta)$. The real profit per unit of capital is

Profit Rate = Revenue – Cost

$$= R/P - (P_k/P)(r + \delta).$$

Because the real rental price in equilibrium equals the marginal product of capital, we can write the profit rate as

$$\text{Profit Rate} = MPK - (P_k/P)(r + \delta).$$

The rental firm makes a profit if the marginal product of capital is greater than the cost of capital. It incurs a loss if the marginal product is less than the cost of capital. The firm's decision, whether to add to its capital or to let it depreciate, depends on profitability of owning and renting out capital. The change in the capital stock, called **net investment**, depends on the difference between the marginal product of capital and the cost of capital.

If the marginal product of capital exceeds the cost of capital, firms find it profitable to add to their capital stock. If the marginal product of capital falls short of the cost of capital, they let their capital stock shrink.

To make a conclusion separating firms is not important. For a firm that both uses and owns capital, the benefit of an extra unit of capital is the marginal product of capital, and the cost is the cost of capital. Like a firm that owns and rents out capital, this firm adds to its capital stock if the marginal product exceeds the cost of capital. Thus, we can write

$$\Delta K = I_n[MPK - (P_k/P)(r + \delta)].$$

Where $I_n()$ is the function showing how much net investment responds to the incentive to invest.

We can now derive the investment function. Total spending on business fixed investment is the sum of net investment and the replacement of depreciated capital. The investment function is

$$I = I_n[MPK - (P_k/P)(r + \delta)] + \delta K$$

 Business fixed investment depends on the marginal product of capital, the cost of capital, and the amount of depreciation.

This model shows why investment depends on the interest rate. A decrease in the real interest rate lowers the cost of capital. Therefore, it raises the amount of profit from owning capital and increases the incentive to accumulate more capital. Similarly, an increase in the real interest rate raises the cost of capital and leads firms to reduce their investment. For this reason, the

The model also shows what causes the investment schedule to shift. Any event that raises the marginal product of capital increases the profitability of investment and causes the investment schedule to shift outward. For example, a technological innovation that increases the production function parameter A raises the marginal product of capital and, for any given interest rate, increases the amount of capital goods that rental firms wish to buy.

Finally, consider what happens as this adjustment of the capital stock continues over time. If the marginal product begins above the cost of capital, the capital stock will rise and the marginal product will fall. If the marginal product of capital begins below the cost of capital, the capital stock will fall and the marginal product will rise. Eventually, as the capital stock adjusts, the marginal product of capital approaches the cost of capital. When the capital stock reaches a *steady-state level*, we can write

$$MPK = (P_k/P)(r + \delta)$$

Thus, in the long run, the marginal product of capital equals the real cost of capital. The speed of adjustment toward the steady state depends on how quickly firms adjust their capital stock, which in turn depends on how costly it is to build, deliver, and install new capital.

Taxes and Investment

Tax laws influence firms' incentives to accumulate capital in many ways. Sometimes policymakers change the tax code to shift the investment function and influence aggregate demand. Here we consider two of the most important provisions of corporate taxation: *the corporate income tax and the investment tax credit*.

The **corporate income tax** is a tax on corporate profits. The effect of a corporate income tax on investment depends on how the law defines "profit" for the purpose of taxation. Suppose, first, that the law defined profit as we did previously, the rental price of capital minus the cost of capital. In this case, even though firms would be sharing a fraction of their profits with the government, it would still be rational for them to invest if the rental price of capital exceeded the cost of capital and to disinvest if the rental price fell short of the cost of capital. A tax on profit, measured in this way, would not alter investment incentives.

Yet, because of the tax law's definition of profit, the corporate income tax does affect investment decisions. There are many differences between the law's definition of profit and ours. For example, our definition of profit deducts the *current* value of depreciation as a cost. That is, it bases depreciation on how much it would cost today to replace worn-out capital. By contrast, under the corporate tax laws, firms deduct depreciation using *historical* cost. That is, the depreciation deduction is based on the price of the capital when it was originally purchased. In periods of inflation, replacement cost is greater than historical cost, so the corporate tax tends to understate the cost of depreciation and overstate profit. As a result, the tax law sees a profit and levies a tax even when economic profit is zero, which makes owning capital less attractive. For this and other reasons, many economists believe that the corporate income tax discourages investment.

Policymakers often change the rules governing the corporate income tax in an attempt to encourage investment or at least mitigate the disincentive the tax provides. One example is the **investment tax credit**, a tax provision that reduces a firm's taxes by a certain amount for each dollar spent on capital goods. Because a firm recoups part of its expenditure on new capital in lower taxes, the credit reduces the effective purchase price of a unit of capital P_k . Thus, the investment tax credit reduces the cost of capital and raises investment.

Residential Investment

Residential investment includes the purchase of new housing both by people who plan to live in it themselves and by landlords who plan to rent it to others. To keep things simple, however, it is useful to imagine that all housing is owner-occupied.

The Stock Equilibrium and the Flow Supply

There are two parts to the model.

- First, the market for the existing stock of houses determines the equilibrium housing price.
- Second, the housing price determines the flow of residential investment.

The panel A of figure 2 shows how the relative price of housing P_H/P is determined by the supply and demand for the existing *stock of houses*. At any point in time, the supply of houses is fixed. We represent this stock with a *vertical supply curve*. The demand curve for houses slopes downward, because high prices cause people to live in smaller houses, to share residences, or sometimes even to become homeless. *The price of housing adjusts to equilibrate supply and demand.*

- *The higher the relative price of housing, the greater the incentive to build houses and the more houses are built.*

The flow of new houses, residential investment, therefore depends on the equilibrium price set in the market for existing houses.

This model of residential investment is similar to the q theory of business fixed investment. According to the q theory, business fixed investment depends on the market price of installed capital relative to its replacement cost; this relative price, in turn, depends on the expected profits from owning installed capital. According to this model of the housing market, residential investment depends on the relative price of housing. The relative price of housing, in turn, depends on the demand for housing, which depends on the imputed rent that individuals expect to receive from their housing. Hence, the relative price of housing plays much the same role for residential investment as Tobin's q does for Business Fixed Investment.

Changes in Housing Demand

When the demand for housing shifts, the equilibrium price of housing changes, and this change in turn affects residential investment.

Reasons shift in demand curve for housing

- An economic boom raises national income and therefore the demand for housing.
- A large increase in the population, perhaps because of immigration, also raises the demand for housing.

One important determinant of housing demand is the *real interest rate*. Many people take out loans, **mortgages**, to buy their homes; the interest rate is the cost of the loan. Even the few people who do not have to borrow to purchase a home will respond to the interest rate, because the interest rate is the opportunity cost of holding their wealth in housing rather than putting it in a bank. A reduction in the interest rate therefore raises housing demand, housing prices, and residential investment.

Another important determinant of housing demand is credit availability. When it is easy to get a loan, more households buy their own homes, and they buy larger ones than they otherwise might, thus increasing the demand for housing. When credit conditions become tight, fewer people buy their own homes or trade up to larger ones, and the demand for housing falls.

Inventory Investment

Inventory investment: the goods that businesses put aside in storage. It is at the same time negligible and of great significance. It is one of the smallest components of spending, averaging about 1 percent of GDP. Yet its remarkable volatility makes it central to the study of economic fluctuations. In recessions, firms stop replenishing their inventory as goods are sold, and inventory investment becomes negative. In a typical recession, more than half the fall in spending comes from a decline in inventory investment.

Reasons for Holding Inventories

Inventories serve many purposes. Let's discuss in broad terms some of the motives firms have for holding inventories.

The first reason for holding inventories is *to smooth the level of production over time*. Consider a firm that experiences temporary booms and busts in sales. Rather than adjusting production to

match the fluctuations in sales, the firm may find it cheaper to produce goods at a steady rate. When sales are low, the firm produces more than it sells and puts the extra goods into inventory. When sales are high, the firm produces less than it sells and takes goods out of inventory. This motive for holding inventories is called **production smoothing**.

A second reason for holding inventories is that they may allow a firm to operate more efficiently. Retail stores, for example, can sell merchant disempower effectively if they have goods on hand to show to customers. Manufacturing firms keep inventories of spare parts to reduce the time that the assembly line is shut down when a machine breaks. In some ways, we can view **inventories as a factor of production**: the larger the stock of inventories a firm holds, the more output it can produce.

A third reason for holding inventories is *to avoid running out of goods* when sales are unexpectedly high. Firms often have to make production decisions before knowing the level of customer demand. For example, a publisher must decide how many copies of a new book to print before knowing whether the book will be popular. If demand exceeds production and there are no inventories, the good will be out of stock for a period, and the firm will lose sales and profit. Inventories can prevent this from happening. This motive for holding inventories is called **stock-out avoidance**.

A fourth explanation of inventories is dictated by the production process. Many goods require a number of production steps and, therefore, take time to produce. When a product is only partly completed, its *components are counted as part of a firm's inventory*. These inventories are called **work in process**.

How the Real Interest Rate and Credit Conditions Affect Inventory Investment

Like other investment, inventory investment depends on the real interest rate. When a firm holds a good in inventory and sells it tomorrow rather than selling it today, it gives up the interest it could have earned between today and tomorrow. Thus, the real interest rate measures the opportunity cost of holding inventories.

When the real interest rate rises, holding inventories becomes more costly, so rational firms try to reduce their stock. Therefore, an increase in the real interest rate depresses inventory

investment. For example, in the 1980s many firms adopted “just-in-time” production plans, which were designed to reduce the amount of inventory by producing goods just before sale. The high real interest rates that prevailed during most of this decade are one possible explanation for this change in business strategy.

Inventory investment also depends on credit conditions. Because many firms rely on bank loans to finance their purchases of inventories, they cut back when these loans are hard to come by.

Accelerator model of inventory investment

The *accelerator model* assumes that firms hold a stock of inventories that is proportional to the firm’s level of output. Thus, if N is the economy’s stock of inventories and Y is output, then

$$N = \beta Y$$

Where, β is a parameter reflecting how much inventory firms wish to hold as a proportion of output. Inventory investment I is the change in the stock of inventories ΔN . Therefore,

$$I = \Delta N = \beta \Delta Y$$

The accelerator model predicts that inventory investment is proportional to the change in output.

- When output rises, firms want to hold a larger stock of inventory, so inventory investment is high.
- When output falls, firms want to hold a smaller stock of inventory, so they allow their inventory to run down, and inventory investment is negative.

The model says that inventory investment depends on whether the economy is speeding up or slowing down.

2.4 The Keynesian Explanation

One continuing source of debate among economists is whether stock market fluctuations are rational.

According to the **efficient markets hypothesis**, the market price of a company's stock is the fully rational valuation of the company's value, given current information about the company's business prospects.

According to this theory, the stock market is *efficient*: it reflects all available information about the value of the asset. Stock prices change when information changes. When good news about the company's prospects becomes public, the value and the stock price both rise. When the company's prospects deteriorate, the value and price both falls. But at any moment in time, the market price is the rational best guess of the company's value based on available information.

One implication of the efficient markets hypothesis is that stock prices should follow a *random walk*. This means that the changes in stock prices should be unpredictable from available information. According to this theory, the only thing that can move stock prices is news that changes the *market's perception* of the company's value. But such news must be unpredictable, otherwise, it wouldn't really be news. For the same reason, changes in stock prices should be unpredictable as well.

What is the evidence for the efficient markets hypothesis? Its proponents point out that it is hard to beat the market by buying allegedly undervalued stocks and selling allegedly overvalued stocks. Statistical tests show that stock prices are random walks, or at least approximately so.

But some economists are less convinced that the stock market is so rational. These economists point out that many movements in stock prices are hard to attribute to news. They suggest that when buying and selling, stock investors are less focused on companies' fundamental values and more focused on what they expect other investors will later pay.

John Maynard Keynes proposed a famous analogy to explain stock market speculation. In his day, some newspapers held "beauty contests" in which the paper printed the pictures of 100 women and readers were invited to submit a list of the five most beautiful. A prize went to the reader whose choices most closely matched those of the consensus of the other entrants. A naive entrant would simply have picked the five most beautiful women in his eyes. But a slightly more sophisticated strategy would have been to guess the five women whom other people considered the most beautiful. Other people, however, were likely thinking along the same lines. So an even

more sophisticated strategy would have been to try to guess who other people thought other people thought were the most beautiful women. And so on. In the end of the process, judging true beauty would be less important to winning the contest than guessing other people's opinions of other people's opinions.

Similarly, Keynes reasoned that because stock market investors will eventually sell their shares to others, they are more concerned about other people's valuation of a company than the company's true worth. The best stock investors, in his view, are those who are good at out-guessing mass psychology. He believed that movements in stock prices often reflect irrational waves of *optimism and pessimism*, which he called the “*animal spirits*” of investors.

The two views of the stock market persist to this day. The stock market often fluctuates for no good reason, and because the stock market influences the aggregate demand for goods and services, these fluctuations are a source of short-run economic fluctuations.

2.5 The Stock Market and Tobin's q

Many economists see a link between fluctuations in investment and fluctuations in the stock market. The term **stock** refers to shares in the ownership of corporations, and the **stock market** is the market in which these shares are traded. Stock prices tend to be high when firms have many opportunities for profitable investment, because these profit opportunities mean higher future income for the shareholders. Thus, stock prices reflect the incentives to invest.

The Nobel Prize-winning economist James Tobin proposed that firms base their investment decisions on the ratio, which is now called **Tobin's q** :

$$q = \frac{\text{Market Value of Installed Capital}}{\text{Replacement Cost of Installed Capital}}$$

The numerator of Tobin's q is the value of the economy's capital as determined by the stock market. The denominator is the price of that capital if it were purchased today.

Tobin reasoned that net investment should depend on whether q is greater or less than 1. If q is greater than 1, then the stock market values installed capital at more than its replacement cost. In this case, managers can raise the market value of their firms' stock by buying more capital.

Conversely, if q is less than 1, the stock market values capital at less than its replacement cost. In this case, managers will not replace capital as it wears out.

The q theory and the neoclassical model two are closely related. To see the relationship, note that Tobin's q depends on current and future expected profits from installed capital. If the marginal product of capital exceeds the cost of capital, then firms are earning profits on their installed capital. These profits make the firms more desirable to own, which raises the market value of these firms' stock, implying a high value of q . Similarly, if the MPK falls short of the cost of capital, then firms are incurring losses on their installed capital, implying a low market value and a low value of q .

The advantage of Tobin's q as a measure of the incentive to invest is that it reflects the expected future profitability of capital as well as the current profitability. For example, suppose that Congress legislate a reduction in the corporate income tax beginning next year. This expected fall in the corporate tax means greater profits for the owners of capital. These higher expected profits raise the value of stock today, raise Tobin's q , and therefore encourage investment today. Thus, Tobin's q theory of investment emphasizes that investment decisions depend *not only on current economic policies but also on policies expected to prevail in the future*.

Chapter Three: Money Demand and Money Supply

3.1 What is money?

When we say that a person has a lot of money, we usually mean that he or she is wealthy. By contrast, economists use the term money in a more specialized way. To an economist, money does not refer to all wealth but only to one type of it: money is the stock of assets that can be readily used to make transactions. Roughly speaking, the currency in the hands of the public make up the nation's stock of money.

3.1.1 The Functions of Money

Money has three purposes. It is a store of value, a unit of account, and a medium of exchange.

As a store of value, money is a way to transfer purchasing power from the present to the future. If I work today and earn \$100, I can hold the money and spend it tomorrow, next week, or next month. Of course, money is an imperfect store of value: if prices are rising, the amount you can buy with any given quantity of money is falling. Even so, people hold money because they can trade the money for goods and services at some time in the future.

As a unit of account, money provides the terms in which prices are quoted and debts are recorded. Microeconomics teaches us that resources are allocated according to relative prices—the prices of goods relative to other goods—yet stores post their prices in dollars and cents. A car dealer tells you that a car costs \$20,000, not 400 shirts (even though it may amount to the same thing). Similarly, most debts require the debtor to deliver a specified number of dollars in the future, not a specified amount of some commodity. Money is the yardstick with which we measure economic transactions.

As a medium of exchange, money is what we use to buy goods and services. “This note is legal tender of all debts, public and private” is printed on the U.S. dollar “payable to the bearer on demand is printed on ethiopian birr”. When we walk into stores, we are confident that the shopkeepers will accept our money in exchange for the items they are selling. The ease with which money is converted into other things—goods and services—is sometimes called money's liquidity.

To better understand the functions of money, try to imagine an economy without it: a barter economy. In such a world, trade requires the double coincidence of wants—the unlikely

happens to two people each having a good that the other wants at the right time and place to make an exchange. A barter economy permits only simple transactions.

Money makes more indirect transactions possible. A professor uses her salary to buy books; the book publisher uses its revenue from the sale of books to buy paper; the paper company uses its revenue from the sale of paper to pay the lumberjack; the lumberjack uses his income to send his child to college; and the college uses its tuition receipts to pay the salary of the professor. In a complex, modern economy, trade is often indirect and requires the use of money.

3.1 The Types of Money

Money takes many forms. Money that has no intrinsic value is called fiat money because it is established as money by government decree, or fiat.

Although fiat money is the norm in most economies today, most societies in the past have used for money a commodity with some intrinsic value. Money of this sort is called commodity money. The most widespread example of commodity money is gold. When people use gold as money (or use paper money that is redeemable for gold), the economy is said to be on a gold standard. Gold is a form of commodity money because it can be used for various purposes—jewelry, dental fillings, and so on as well as for transactions. The gold standard was common throughout the world during the late nineteenth century.

3.1. Money supply and monetary expansion mechanism

The quantity of money available is called the **money supply**. The control over the money supply is called **monetary policy**.

There are Different measures of money supply. Not all of them are widely used and the exact classifications depend on the country.

Narrow money: Referred to as M1- normally include coins and notes in circulation and other money equivalents that are easily convertible into cash. It is the sum of currency out side banks and demand deposits . $M1 = C + D$

Broad money : referred to as M2. It is also known as money and quasy money . it is the sum of currency out side banks, demand deposits , time deposits , foriegn currency deposits of resident sectors

Diffrence between saving, demand and time deposits

Saving deposit : interst bearing account that can be with drawn up on demand at no cost . The depositer needs facilities like ATM , online banking and credit card. There is nof fixed time to with draw money from banks

Demand deposits : can be uded for transaction . Usually non interst bearning account and Can be with drawn on demand . There is No fixed time period to with draw money

Time deposits : Are deposited in banks for a fixed period of time – usually 1 to 5 years . if one with draw money before maturity then he has to pay penalty. The rate of interst offred by banks on time deposits is higher than saving accounts . One does not need facilities like ATM and online banking

The role of banks in money supply

$$M = C + D \text{ -----}(1)$$

Since the money supply includes demand deposits- the banking system plays an important role

Preliminaries

- ❖ **Reserves:** the portion of deposits that banks have not lent.
- ❖ To a bank-**liabilities** include deposits and **assets** include reserves and outstanding loans
- ❖ **100-percent-reserve banking:** a system in which banks hold all deposits as reserves.
- ❖ **Fractional-reserve banking:** a system in which banks hold a fraction of their deposits as reserves.

Scenario1: No banks

With no banks: $D = 0$ and $M = C = \text{USD}1000$.

Scenario 2: 100 Percent Reserve Banking

Initially

$$M = \text{USD}1000.$$

$$C = \text{USD } 1000$$

$$D = \text{USD } 0$$

$$M = \text{USD } 1000.$$

Now suppose households deposit the USD 1000 at First bank.

After the deposit:

$$C = \text{USD } 0$$

$$D = \text{USD } 1000$$

100 Percent Reserve Banking has no impact on size of money supply.

SCENARIO 3: Fractional-Reserve Banking

Suppose banks hold 20 Percent of deposits in reserve- making loans with the rest. First bank will make USD 800 in loans.

The money supply now equals USD 1800: The depositor still has USD 1000 in demand deposits but now the borrower holds USD 800 in currency.

Thus in a fractional-reserve banking system- banks create money

Finding the total amount of money:

Original deposit	= USD1000	
+	First bank lending	= USD 800
+	Second bank lending	= USD 640
+	Third bank lending	= USD 512
+	Other lending...	

Total money supply = $(1/rr) \times \text{USD}1000$. Where rr = ratio of reserves to deposits. In our example, $rr= 0.2$, so $M= \text{USD} 5000$

Money creation in the banking system

A fractional reserve banking system creates money, but it doesn't create wealth: bank loans give borrowers some new money and an equal amount of new debt.

A model of the money supply

- The monetary base, $B = C + R$ -----(2)

Which controlled by the central bank

$$M = C + D \text{ and } B = C + R$$

$$M/B = C+D/C+R \text{ ----- (3)}$$

Now, divide both the nominator and denominator by D

$$M/B = [C/D + D/D] / [C/D + R/D]$$

$$M/B = [cr + 1] / [cr + rr]$$

$$M = \{[cr + 1] / [cr + rr]\}B$$

$$M = \{[cr + 1] / [cr + rr]\}B = m \times B$$

Where, $m = cr+1/cr+rr$ is the money multiplier

- The **reserve-deposit ratio**, $rr = R/D$. depends on regulations & bank policies
- The **currency-deposit ratio**, $cr = C/D$. depends on households' preferences

Each dollar of the monetary base produces m dollars of money. Because the monetary base has a multiplied effect on the money supply and that is why the monetary base known as high-powered money.

If $rr < 1$, then $m > 1$

If monetary base changes by ΔB , then $\Delta M = m \times \Delta B$. m is called the **money multiplier**. The money supply is proportional to the monetary base. So, an increase in B increases M m -fold. The lower the reserve-deposit ratio, the more loans banks make and the higher is the money multiplier. The lower the currency deposit ratio, the fewer dollars of the monetary base the public holds as currency and the lower is the money multiplier

Exercise; suppose households decide to hold more of their money as currency and less in the form of demand deposits.

1. Determine impact on money supply and explain the intuition for your result.

Answer: Impact of an increase in the currency-deposit ratio $\Delta cr > 0$. An increase in cr increases the denominator proportionally more than the numerator. So m falls, causing M to fall too. If households deposit less of their money, then banks can't make as many loans, so the banking system won't be able to "create" as much money.

Instruments of monetary policy

Open market operations: The purchase or sale of government bonds by the Central bank.

How it Works: If The central Bank buys bonds from the public, it pays with new Currency, increasing B and therefore M .

Reserve requirements: Central Bank regulations require banks to hold a minimum reserve-deposit ratio.

How it Works: Reserve requirements affect rr and thus m : If CB reduces reserve requirements, then banks can make more loans and "create" more money from each deposit.

The discount rate: The interest rate that the Fed charges on loans it makes to banks.

How it Works: When banks borrow from the CB, their reserves increase, allowing them to make more loans and “create” more money. The CB can increase **B** by lowering the discount rate to induce banks to borrow more reserves from the CB.

Which instrument is used most often?

Open market operations: Most frequently used.

Changes in reserve requirements: Least frequently used

Changes in the discount rate: Largely symbolic; the CB is a “lender of last resort,” does not usually make loans to banks on demand.

Why the CB can not precisely control M

$M = \{[cr + 1] / [cr + rr]\} B = m \times B$. Households can change cr , causing m and M to change. Banks often hold **excess reserves** (reserves above the reserve requirement). If banks change their excess reserves, then rr , m and M change.

Example: $B = \text{Birr } 500 \text{ billion}$, $cr = 0.6$ and $rr = 0.1$: Find the money supply?

$$m = (0.6 + 1) / (0.6 + 0.1) = 2.3$$

$$M = 2.3(500) = \text{Birr } 1,150 \text{ billion}$$

Money Demand

The **demand for money** is the desired holding of financial assets in the form of money: that is, cash or bank deposits. There are two types of Money demand theories: Portfolio theories and Transactions theories.

A Simple Portfolio Theory

Stress the store of value function. Posit that money demand depends on risk/return of money & alternative assets. According to these theories, people hold money as part of their portfolio of assets. The key insight is that money offers a different combination of risk and return than other assets. In particular, money offers a safe (nominal) return, whereas the prices of stocks and bonds may rise or fall.

Thus, some economists have suggested that households choose to hold money as part of their optimal portfolio.

Portfolio theories predict that the demand for money should depend on the risk and return offered by money and by the various assets households can hold instead of money. In addition, money demand should depend on total wealth, because wealth measures the size of the portfolio to be allocated among money and the alternative assets. For example, we might write the money demand function as

$$(M/P)^d = L(r_s, r_b, \pi^e, W)$$

Where: r_s is the return on stocks, r_b is the return on bonds, π^e is the expected inflation rate and W is wealth. The demand for money is negatively related with return on stocks, bonds, expected inflation rate and positively related with wealth.

 From the standpoint of portfolio theories, we can view our money demand function, $L(i, Y)$, as a useful simplification. First, it uses real income Y as a proxy for real wealth W . Second, the only return variable it includes is the nominal interest rate, which is the sum of the real return on bonds and expected inflation (that is, $i = r_b + \pi^e$). According to portfolio theories, however, the money demand function should include the expected returns on other assets as well.

 Portfolio theories are more plausible as theories of money demand if we adopt a broad measure of money. The broad measures include many of those assets that dominate currency and checking accounts. **M2**, for example, includes savings accounts and money market mutual funds. When we examine why people hold assets in the form of **M2**, rather than bonds or stock, the portfolio considerations of risk and return may be paramount. Hence, although the portfolio approach to money demand may not be plausible when applied to **M1**, it may be a good theory to explain the demand for **M2**.

3.2 Money demand –portfolio and transaction theories

Theories of money demand that emphasize the role of money as a medium of exchange are called **transactions theories**.

These theories acknowledge that money is a dominated asset and stress that people hold money, unlike other assets, to make purchases. These theories best explain why people hold narrow

measures of money, such as currency and checking accounts, as opposed to holding assets that dominate them, such as savings accounts or Treasury bills.

✎ Transactions theories of money demand take many forms, depending on how one models the process of obtaining money and making transactions. All these theories assume that money has the cost of earning a low rate of return and the benefit of making transactions more convenient. People decide how much money to hold by trading off these costs and benefits.

The Baumol-Tobin Model of Cash Management

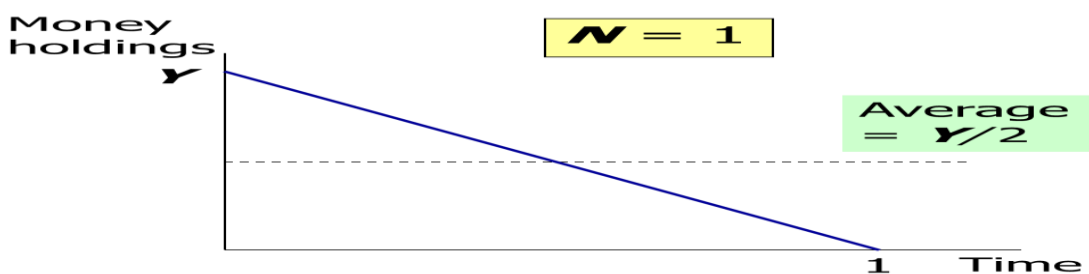
It analyzes the costs and benefits of holding money. The benefit of holding money is convenience: people hold money to avoid making a trip to the bank every time they wish to buy something. The cost of this convenience is the forgone interest they would have received had they left the money deposited in a savings account that paid interest.

To see how people trade off these benefits and costs, consider a person who plans to spend Y dollars gradually over the course of a year. (For simplicity, assume that the price level is constant, so real spending is constant over the year.) How much money should he hold in the process of spending this amount? That is, what is the optimal size of average cash balances?

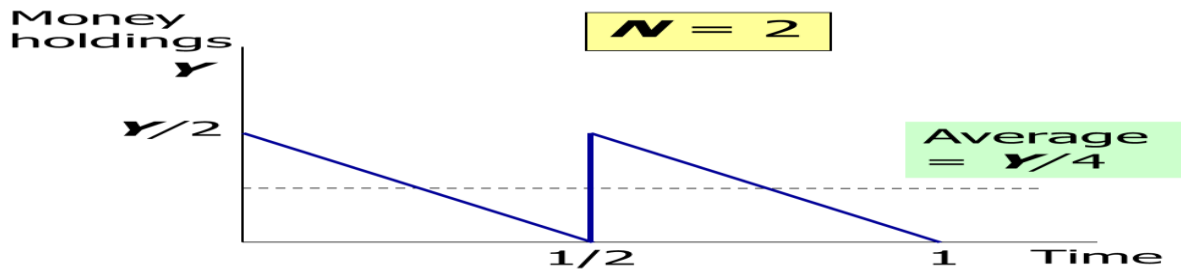
📖 **Notation:** Y = Total spending, done gradually over the year, i = interest rate on savings account, N = number of trips consumer makes to the bank to withdraw money from savings account, F = cost of a trip to the bank (e.g., if a trip takes 15 minutes and consumer's wage = \$12/hour, then $F = \$3$)

Money holdings over the year:

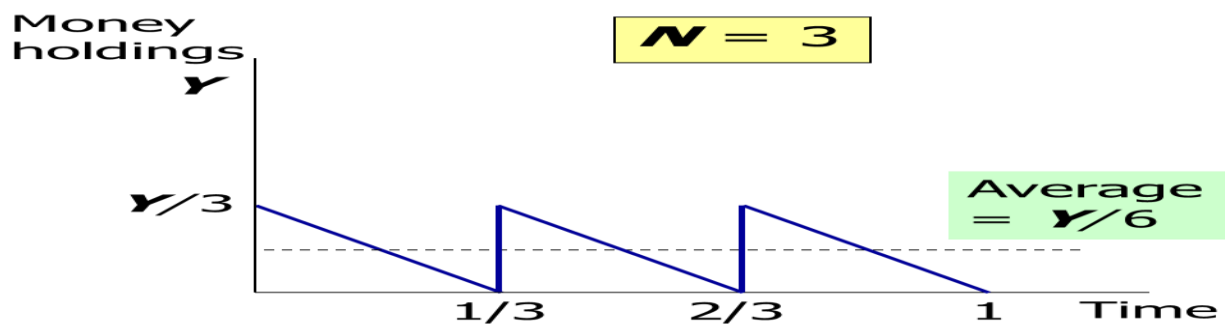
A. If the individual decides to withdraw once throughout the year



B. If the individual decides to make two trip ($N=2$)

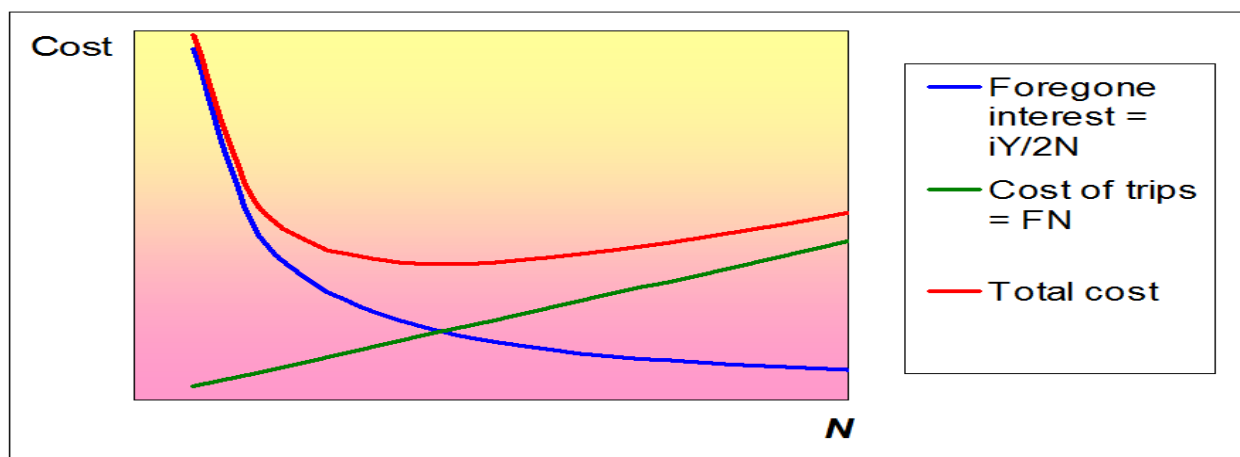


C. The individual making N trip to the bank.



The Cost of Holding Money

- In general, average money holdings = $Y/2N$
- Foregone interest = $i \times (Y/2N)$
- Cost of N trips to bank = $F \times N$
- **Total Cost** = $i \times (Y/2N) + F \times N$



- Given Y , i , and F , consumer chooses N to minimize total cost

Finding the cost-minimizing N

$$\text{Total Cost} = i \times (Y/2N) + F \times N$$

Take the derivative of total cost with respect to N , and then set it equal to zero:

$$\Delta \text{Total Cost} / \Delta N = -iY/2N^2 + F = 0.$$

Solve for the cost-minimizing N^* .

$$N^* = [iY/2F]^{0.5}$$

The Money Demand Function

$$\text{The cost-minimizing } N^* = [iY/2F]^{0.5}$$

To obtain the money demand function, plug N^* into the expression for average money holdings:

$$\text{Average Money Holdings} = Y/2N^* = [YF/2i]^{0.5}$$

Money demand depends positively on Y and F , and negatively on i .

$$\text{The Baumol-Tobin money demand function: } (M/P)^d = [YF/2i]^{0.5}$$

B-T shows how F affects money demand. B-T implies that the income elasticity of money demand = 0.5, interest rate elasticity of money demand = -0.5

One implication of the Baumol-Tobin model is that any change in the fixed cost of going to the bank F alters the money demand function, that is, it changes the quantity of money demanded for any given interest rate and income. It is easy to imagine events that might influence this fixed cost.

✎ The spread of automatic teller machines, for instance, reduces F by reducing the time it takes to withdraw money. Similarly, the introduction of Internet banking reduces F by making it easier to transfer funds among accounts. On the other hand, an increase in real wages increases F by increasing the value of time. And an increase in banking fees increases F directly. Thus, although the Baumol-Tobin model gives us a very specific money demand function, it does not give us reason to believe that this function will necessarily be stable over time.

Financial Innovation, Near Money, and the Demise of the Monetary Aggregates

Traditional macroeconomic analysis groups assets into two categories: those used as a medium of exchange as well as a store of value (currency, checking accounts) called “money.” and those used only as a store of value (stocks, bonds, savings accounts) called “near money”.

Near money is nonmonetary assets that have acquired some of the liquidity of money.

↗ In recent years it has become more difficult to use in monetary and nonmonetary asset. In part because of the deregulation of banks and other financial institutions, and in part because of improved computer technology, the past two decades have seen rapid financial innovation. Monetary assets such as checking accounts once paid no interest; today they earn market interest rates and are comparable to nonmonetary assets as stores of value.

The existence of near money complicates monetary policy by making the demand for money unstable. One response to this problem is to use a broad definition of money that includes near money

Chapter Four: Theories of Labor Market

4.1 Definition of unemployment

The participation of labors in different productive activities helps in bringing wellbeing of the economy as a whole. Labors come from household sector of the economy and a larger proportion of the aggregate demand arises from the household sector or their income. This is again determined by the fact that whether workers or members of households are employed or not. Thus, their income depends on employment and their ability to purchase in turn determines national output. This is why every country policy makers (governments) are concerned with solving or reducing the problem of unemployment as much as possible. Let us begin this section with definition of the concept of unemployment.

Unemployment refers to a situation where workers of the working age could not find job while they are ready to work at prevailing market wage rate. The problem of unemployment and its intensity is usually measured in terms of unemployment rate. This is because talking about the actual number of unemployed people makes no sense for countries of different population size. Unemployment of 10,000 people in a given country with total population of 500,000 people is much compared to unemployment of 10,000 people in a country with total population of 500,000,000 people. Under this condition, we can say there is unemployment problem in the former where as there is no significant unemployment problem in the later. This is why it is better to measure unemployment problem in terms of unemployment rate.

Unemployment rate is defined as the percentage or the proportion of the labour force that is unemployed. Labour force (L) is the sum of both employed (E) and unemployed people of working age and working or ready to work (U). This can be given by the following equation:

$$L = E + U \dots\dots\dots 1$$

Note that labour force and working age population are different. Labour force does not include part of the working age population which are not working or looking for work, that is, the working age population that is not in the labour market. The related concept is labour force participation rate. Labour force participation is the percentage of adult or working age population who are in the labour force. These two concepts are summarized by the following equations:

$$\text{Unemployment Rate} = \left(\frac{\text{Numbers of Unemployed}}{\text{Labor Force}} \right) \times 100$$

$$\text{Labor Force participation} = \left(\frac{\text{Labor Force}}{\text{Adult Population}} \right) \times 100$$

The concept of natural rate of unemployment is used as reference to the existence of unemployment problem. **Natural rate of unemployment** is the average rate of unemployment around which an economy fluctuates given by $U/L = n$ at that point. It is related to the rate at which workers lose job (job loss rate) and the rate at which jobless workers find or get job (job finding rate). This can be more elaborated as follows using the steady state unemployment rate. **Steady-state unemployment rate** is the point or condition in which the number of workers leaving or losing job are equal to the number of unemployed getting or finding job given by the following equation:

TRANSITIONS BETWEEN EMPLOYMENT AND UNEMPLOYMENT

The steady state condition

The labor market is in steady state, or long-run equilibrium, if the unemployment rate is constant. **The steady-state condition is:**

$$fU = sE \dots \dots \dots 2$$

🔔 **Number of employed people who lose or leave their jobs = Number of unemployed people who find jobs**

Where 'f' is rate of job finding

's' is rate of job separation or loss

'U' unemployed population

'E' Employed population

Rearranging equation (1) into $E = L - U$ and substituting in (2), we obtain the following relation:

$$fU = s(L - U) \quad \text{Dividing both sides by } L$$

$$fU/L = (s/L)(L - U)$$

$fU/L = s(1 - U/L)$ Dividing both sides by 'f' we obtain

$$U/L = s/(s + f) = n \text{ -----3}$$

This relation is obtained as follows:

$$fU/L = s(1 - U/L)$$

$$fU/L = s - sU/L \quad \text{Multiplying both sides by } (1/f)$$

$$U/L = s/f - (s/f)(U/L) \quad \text{Rearranging gives}$$

$$U/L + (s/f)(U/L) = s/f$$

$$U/L[1 + (s/f)] = s/f$$

$$U/L[(f + s)/f] = s/f \quad \text{multiply both sides by } f/(s + f), \text{ we obtain:}$$

$$U/L = (s/f) [f/(s + f)] \quad \text{Letting } U/L = n$$

$$n = [s/(s + f)]$$

Where $U/L = n$ is the natural rate of unemployment.

The policy implication of this is that any policy aimed at reducing the natural rate of unemployment should either reduce the rate of job separation (losing job), or increase the rate of job finding (f). Moreover, any policy that affects the rate of job separation (s) or job finding (f) also affects the natural rate of unemployment 'n'.

Examples:

- 1) If 90% of labour force of a given country with 10 million labour force population are employed on average, find the unemployment rate of the country.
- 2) If 3% of workers are on average leaving or losing their job and on average 40% of unemployed workers and workers newly joining labour market are finding or getting job, then what is the natural rate of unemployment?
- 3) In question number '1' if the size of the adult population of the country is 12 million, what is the labour force participation rate?

Solutions:

$$\begin{aligned}
 1) \quad & L = E + U \\
 & L = 90\%(L) + U \\
 & L - 0.9(L) = U \\
 & 0.1L = U \\
 & \underline{U/L = 0.1 = 10\%}
 \end{aligned}$$

$$\begin{aligned}
 2) \quad & \text{Given: job separation rate (s) = 3\% = 0.03} \\
 & \text{Job finding rate (f) = 40\% = 0.4} \\
 & U/L = n = [s/(s + f)] \quad \text{proved above} \\
 & = [0.03/(0.03 + 0.4)] \\
 & = 0.03/0.43 \\
 & = 0.0698 \\
 & \underline{n = 6.98\%}
 \end{aligned}$$

$$\begin{aligned}
 3) \quad & \text{Labor Force participation} = \left(\frac{\text{Labor Force}}{\text{Adult Population}} \right) \times 100 \\
 & = (10 \text{ million} / 12 \text{ million}) \times 100 \\
 & = 83.33\%
 \end{aligned}$$

Let us consider next the major types and/or classifications of unemployment.

4.3 Causes and Types of Unemployment

Depending on the causes or sources of the unemployment and the duration of the unemployment, we can divide unemployment into different categories. In this respect, economists divide unemployment into three or four major categories. These are:

- i. Frictional unemployment
- ii. Structural unemployment

- iii. Cyclical unemployment and/or
- iv. Seasonal unemployment

i. Frictional unemployment

Frictional unemployment is a type of unemployment usually caused by constant changes in the labour market. It occurs due to two reasons. The first reason is when employers are not aware of the available workers and their job qualifications. The second reason is when workers are not fully aware of the jobs being offered. The basic cause of frictional unemployment is thus lack of information flow among workers and employers called imperfect information. So for workers this is the period of unemployment until they get a job and the duration of unemployment is equal to the time it takes workers to search for a job. New graduating students from colleges and universities or training institutions are usually frictionally unemployed. This is the period when such people look for vacancies and apply to different offices getting interviewed and so on.

The change in composition of demand among economic sectors, industries or regions called **sectoral shift**. Due to this shift, the demand for workers also shift and some workers have to leave some sectors, industries or regions and look for jobs and join some other sectors which always involve frictional unemployment. The main characteristics of frictional unemployment are that:

- It affects large number and wide range of people.
- It tends to be of short period.
- Certain amount of frictional unemployment is unavoidable.

One of the policy options to solve such unemployment is improving labour market information (e.g. establishment of information office about workers and vacancies). Note that in trying to reduce frictional unemployment, some policies inadvertently increase the amount of frictional unemployment. One example of such policy is unemployment insurance. In this case, people will be reluctant in looking for job as soon as possible since they can collect some money because of their unemployment and prefer to stay for certain period after unemployment.

ii. Structural Unemployment

Structural unemployment occurs due to the structural changes in the economy. These changes eliminate some jobs while they create some new jobs for people with new skill level. The skill

sets take time to develop and hence some people lose their job simply because they do not have the new required skill(s).

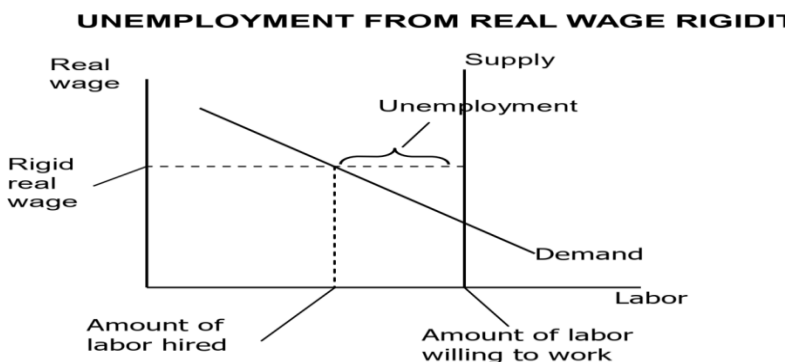
✎ This problem arises from mismatch between the types of jobs that are available and type of job seekers. Such mismatch may be related to skill, education level, geographical area, age, etc. For instance, some skills may no longer be demanded. For instance, typing machines are replaced by computers. In this case type writers who do not have computer skill would lose their job and potentially become unemployed.

Some of the characteristics of this type of unemployment are that:

- It tends to be concentrated among certain group of people who are adversely affected by technological change.
- It tends to be long lasting (e.g. it takes time to the victims until they train themselves under new situation or new technology).

In this respect one of the policy options used to reduce such unemployment is training workers and improving labour mobility. Structural unemployment also arises from real wage rigidity leading to failure of the labour market to adjust to equilibrium or the point where demand for and supply of labour are equal. See the following figure.

Figure 1: Real Wage and Structural Unemployment



Wage rigidity is the failure of wages to adjust to the point where the demand for labour equals its supply. If the real wage is stuck above the equilibrium level given by the point of intersection between labour supply and labour demand curves point 'e', then the supply of labour exceeds its demand. This results in unemployment equal to the distance between point 'a' and point 'b'.

Such amount of unemployment resulting from wage rigidity and job rationing is also known as structural unemployment. During this situation, workers are unemployed not because they cannot find job that suits them or their skills, but at the prevailing wage rate the supply of labour exceeds its demand. Note here that all factors that lead to wage rigidity also lead to structural unemployment.

Cyclical Unemployment

Cyclical unemployment occurs due to general downturn in the business activities including production and demand for the products and services. During recession business conditions, only few goods are produced and for such low production, only few employment opportunities would be available. Employers are therefore, obliged to lay-off workers and cut back employment. As we have tried to explain above, unemployment rate fluctuates around a line known as '**natural rate of unemployment**'. It is a rate where there is no cyclical unemployment or when all the unemployment is frictional and structural ones. Generally, there is always some amount of unemployment and economists very frequently use the term full employment. Full employment occurs when the unemployment rate is equal to the natural rate of unemployment. However, note that the concept of full employment does not mean that all workers are employed.

Cyclical unemployment is also known as **demand deficient unemployment**. Cyclical unemployment is the result of insufficient aggregate demand in the economy to generate enough jobs for those seeking them. It occurs during cyclical contraction of an economy (recession). The policy instrument to solve this problem is fiscal policy (for instance increasing government expenditure and reducing tax rates) and/or monetary policy (such as reducing interest rate and increasing money supply).

iii. Seasonal Unemployment

Seasonal unemployment is the type of unemployment that arises from a decline in the economic activity in some seasons (particular time in a year) and in some sectors. Therefore, seasonal unemployment results from fluctuations in demand for labour in these sectors and/or seasons. The special characteristic of this type of unemployment is that fluctuation can take a regular course of action and can be anticipated so that workers also make their own plan to move to particular sector in specific seasons to avoid such unemployment. For instance, workers can seek

job in agricultural sector during the first season of cultivation and during harvest time. In other seasons the demand for labour in agriculture becomes low and as a result workers would look for job in other sectors.

 In conclusion, with frictional and structural unemployment there could be enough jobs, but it is difficult to match job seekers with job vacancies. With cyclical unemployment, on the other hand, there are not enough jobs for job seekers. The duration of cyclical unemployment can be in between frictional and structural unemployment if the recession is handled and solved quickly (not as short duration as frictional and not as long duration as structural unemployment). Otherwise, it will take longer duration than both other types of unemployment.

Note that sometimes concepts such as underemployment, disguised unemployment and open unemployment are used to demonstrate different degrees of unemployment. Underemployment refers to the people who work below their capacity. For instance, a person may be employed for only 5 hours a day while he/she can and want to work more than that. Disguised unemployment is the case where the worker is employed but adding nothing to the output. Open unemployment represents the formal definition we have discussed above where the person has no job at all while he/she is ready to work at the prevailing market wage rate.

Once we understood the concept of unemployment, we can now discuss the relationship between employment and output of the economy. It is easy to see that these two variables are inversely related; as unemployment increases output also decreases. As a result, one can say that the largest single cost of unemployment is loss of production. People who do not work do not produce. Then, the cost of output lost becomes very high and this is the reason why many economists studied the relationship between these two important variables. The relationship between income or output and unemployment is first considered and explained by Arthur Okun and we treat the concept under Okun's Law. Unemployment has relation with other macroeconomic variables like inflation. We demonstrate this relation using Phillips curve. Now let us begin this discussion of the relationship between unemployment and other macroeconomic variables with Okun's Law and then we will pass to the concept of the Phillips curve.

Okun's Law

Arthur Okun studied the relationship between unemployment and output over the business cycle. In his research and the law (Okun's Law), he predicted that 1% percent change in unemployment rate costs 2% of GDP; i.e. GDP growth falls by 2%. This can be put in general as follows:

$$\text{Percentage change in Real GDP} = \text{growth in real GDP (without impact of unemployment)} - 2 \times (\text{change in unemployment rate})$$

For instance if real GDP is growing by 3% and unemployment rate rises from 6% to 8%, then, after accounting for impact of unemployment, the real GDP growth would be as follows:

$$\begin{aligned} \text{Percentage change in Real GDP} &= 3\% - 2\% (8\% - 6\%) \\ &= -1\% \end{aligned}$$

Since Okun's Law is a law which describes a relationship between growth of real output and change in the unemployment rate it can be roughly summarized by the following relationships:

$$\Delta U = -X(Y_a - Y_t)$$

$$\Delta U = -X(Y_a - Y_t)$$

Where ΔU is change in unemployment

X is a positive constant representing the magnitude, in which unemployment declines due to a percentage output growth,

Y_a is actual growth rate of output, and

Y_t is trend output growth rate.

This general equation simply shows the implied negative relationships between the two variables (growth of real output and change in the unemployment rate).

Phillips Curve

The Australian economist, A.W. Phillips plotted data on unemployment rates and rate of change in wage rates between 1861 and 1957 in UK. He first found that there was more or less stable relationship between these two variables. Later he and other economists extended the concept to

cover larger range of data and found out that there is negative relationship between unemployment rates and rate of change in wage rates. Then they termed the curve showing this relationship as Phillips Curve. The Phillips Curve is a downward sloping curve that shows an inverse relationship between the inflation rate and the unemployment rate.

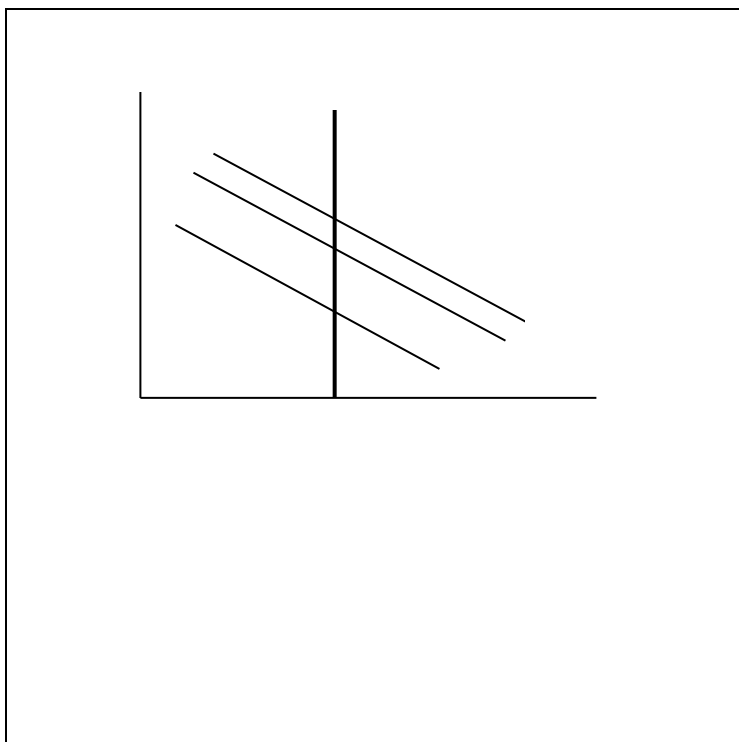
The explanation for this inverse relationship is simple; when unemployment rate is very low (employment is high), then workers have the market power to push up wages. Higher wage rate again implies that the cost of production is high and that sellers charge high prices. When unemployment rate is very high, then workers do not have much bargaining power; rather they would be ready to accept lower wage to get job. Due to this, the pressure on prices also remains low.

Such relationship between inflation and unemployment generated the idea of policy trade off. The Phillips Curve rapidly becomes a cornerstone of macroeconomic policy analysis. It suggests that policymakers could choose different combinations of unemployment and inflation rates. If the world works the way Phillips Curve suggests then policy makers must choose inflation and unemployment combinations along Phillips Curve.

This is because the curves suggest that less unemployment can always be attained by incurring more inflation so that the inflation rate can always be reduced by incurring the costs of more unemployment. In other words, the curve suggests that there is a trade-off between policies intended to reduce inflation and that intended to reduce unemployment. That is we could have only low unemployment as long as we put up with high inflation.

✎ However, after 1960, the Phillips Curve became unstable and people started questioning about the trade-off. In mid-1960s, economists like Milton Friedman and Edmund Phelps suggested that the unemployment rate and rate of inflation are vertical straight line in long-run. In long run there is always natural rate of unemployment (N). In Figure 3 below, PCs curves represent short run Phillips curves.

Figure 2: Long Run Phillips Curve



The implication of the vertical long run in Philips Curve is that there is no change or **deviation in unemployment rate from the natural rate is only of short-run**. An increase above the natural rate of unemployment or a fall below it would be adjusted in the long run.

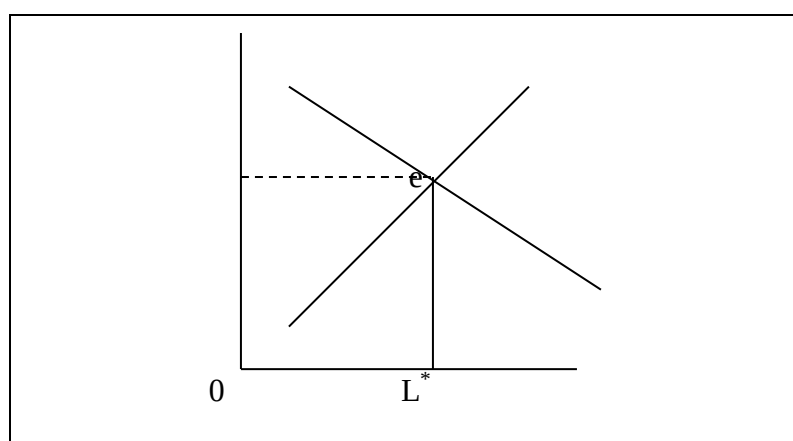
On the contrary, it is expected that inflation has positive impact on the labour market. Some economists argue that some level of inflation may make labour markets work better. They say that it “greases the wheels” of the labour markets. This is because with some level of inflation, the wage rate increases and workers get initiated simply by increased nominal wage rate.

4.4 Classical and Neoclassical Microeconomic model of the labour market

Both the demand for and supply of labour depend on the real wage, W/P . Real wage measures the amount of goods that can be bought with a reward of an hour of work or received wage income. According to classical economists the real wage adjusts to the point where demand for labour equals its supply. The argument of these economists is that during low demand for labour unemployment will be high.

✎ During such period there is no reason why workers do not cut their wage rate to get job. That means workers would be ready to cut their wage or to work at lower wage rate during high unemployment. Similarly, during high employment or low unemployment, which is the case of high demand for labor or when the demand for labor is higher than its supply, employers are ready to pay higher wage rate. This means that wage rate increases to the point where the demand for labor equals its supply. Due to this the labor market is always in equilibrium given by the following graph in Figure 7.4.

Figure 3: Classical Frictionless Labour market equilibrium



At market clearing condition, ‘e’, there is only natural or voluntary unemployment. The equilibrium is said to be frictionless because the classical economists believe that the market always clears. Underlying the classical model is the assumption that there is no unemployment. In equilibrium, everyone who wants to work is working.

✎ But there is always some unemployment. That level of unemployment is accounted for by labour market frictions, which occur because the labour market is always in a state of flux. Some people are moving and changing jobs; other people are looking for job for the first time; some firms are expanding and hiring new workers while others have lost business and are obliged to reduce employment by firing workers.

Frictional unemployment

Frictional unemployment is the unemployment that exists as a result of individuals’ shifting between jobs and looking for new jobs. It is also known as natural rate of unemployment and it

occurs even when the labour market is in equilibrium. The frictionless classical model is an idealized case and is based on the idea that:

- wages and prices are fully flexible.
- no costs to either workers in finding jobs or firms in increasing or reducing their labour force.
- firms behave competitively.
- a firm is expected to sell all its products at prevailing prices.

✎ The policy implication of this is that in the case of closed economy, the real wage rate remains unchanged. And any policy change for instance increasing money supply will result only in increased price level. (It does not affect output). The closed economy is equivalent to short run economic condition and thus, policies will be ineffective. Even if the economy is open, if both domestic and foreign prices change proportionally (with equal percent), the policy does not affect employment and output.

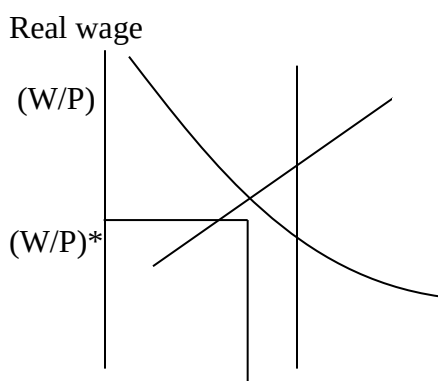
✎ In this case, the labour supply curve shifts to L_s in panel 'a'. However, such change in equal percentage is a very rare case because of existence of purchasing power parity. That is, a given amount of currency does not purchase the same item after changing to another currency because exchange rates does not frequently adjust to such changes.

It is important to discuss here the new Keynesian view of the labour market which contradicts with the position of the classical economists.

4.5 New – Keynesian view of unemployment

Keynesian economists say that the market does not clear because of commodity price rigidity in commodity market and real wage rigidity in the labour market for several reasons some which are mentioned and explained next to the figure below.

Figure 6: Keynesian Labour market Equilibrium



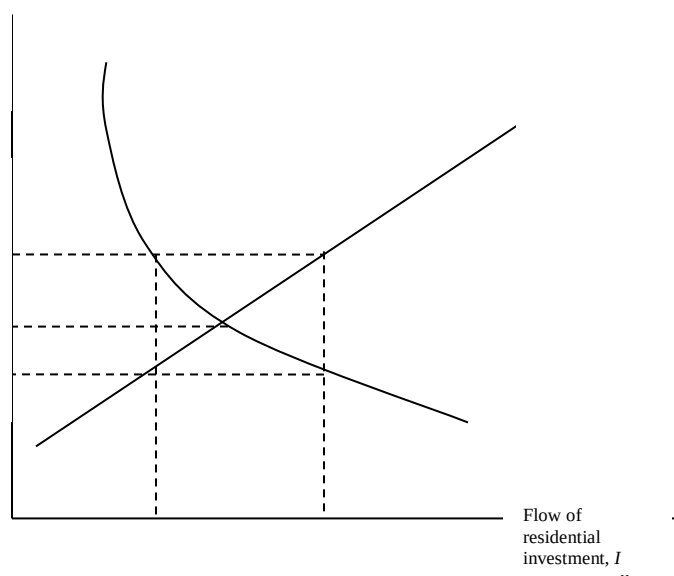
0 L^* L_f Labour Supply

For new Keynesians, the real wage can be (fixed) at $(W/P)^*$ and employment will be at L_1 . Thus the difference between L_f and L^* is wait unemployment or involuntary unemployment. This arises from wage rigidities and job rationing. However, what are the reasons for real wage rigidities? The following are some of the major reasons.

1) Minimum wage rate legislation

This is the result of government policy of welfare protection or of labour- union pressure. It represents the case where the real wage rate is fixed above the market clearing one. See the following figure (figure 7.8). When minimum wage is fixed at $(W/P)_{\min}$, the difference between L_1 and L_2 , (L_2-L_1) , will be involuntary unemployment. This amount of unemployment exists even people are willing to work at wage rate below minimum wage rate $(W/P)_{\min}$ for instance, at $(W/P)^*$.

Figure 7.8: Minimum wage rate



If the minimum wage is set at a point less than equilibrium wage rate, (for instance at w^1), this will not be a constraining factor.

2) Monopoly power of labour unions

Sometimes firms might be interested to pay high wage (which is larger than the equilibrium wage rate) to avoid unionization. Wage tends to be rigid because of interest conflict between insiders and outsiders. Outsiders (employers and unemployed workers) argue for or accept low wage while insiders (employed workers) argue for high wage.

3) Efficiency wage hypothesis

Firms prefer to pay above the equilibrium wage rate for several reasons. Some of the reasons are:

- a. Nutrition condition: Better paid worker is well-fed, healthy and efficient. And this in turn reduces absenteeism from work.
- b. To reduce labour turn over: By paying higher wage firms can avoid or reduce frequent recruitment costs, advertisement cost and training costs. This is because higher wage rates make workers prefer to stay with the high paying firm.
- c. To keep quality workers and improve work effort: Qualified workers seek for high paying firms. High payment also avoids shirking constraints.

4) Imperfect information

Imperfect information about changes in prices in the commodity market may lead to lag in change of wage rate in the labour market. Thus, because of lack of information, the real wage rate may also fail to adjust to the market clearing one.

Before concluding this section let us explain the reaction of labour supply or workers during the different phases of the business cycle (cyclical recession and cyclical boom).

Labour supply and Business Cycle

Labour supply responds not only to changes in economic opportunities over a worker's life time, but may also show adjustments to changes in labour opportunities induced by business cycle. In this respect, two major hypotheses are developed namely *added worker effect* hypothesis and *discouraged worker effect* hypothesis.

The *added worker effect* hypothesis suggests that secondary workers (who lack commitment to the labour market such as students, house wives, retired persons, etc) will tend to become labour market participant during economic recessions; and then withdraw their labour during prosperity. The *added worker effect* thus implies that the labour force participation rate of secondary workers has a *counter cyclical* trend (i.e. it moves in opposite direction to the business cycle). This is because, for primary workers, there are no enough jobs during recession.

The *discouraged workers effect* hypothesis, on the other hand, argues that unemployed primary workers finds so difficult to get work that they eventually are discouraged and stop looking for work. As a result the labour force participation rate (labour supply) declines. So the supply of primary workers has a *pro-cyclical* trend (it falls during recession and increases during boom).

Chapter Five: Macroeconomic Policy Debates

5.1 Should policy be active or passive

Policymakers in the federal government view economic stabilization as one of their primary responsibilities. To many economists the case for active government policy is clear and simple. Recessions are periods of high unemployment, low incomes, and increased economic hardship. The model of aggregate demand and aggregate supply shows how shocks to the economy can cause recessions. It also shows how monetary and fiscal policy can prevent recessions by responding to these shocks. These economists consider it wasteful not to use these policy instruments to stabilize the economy.

Other economists are critical of the government's attempts to stabilize the economy. These critics argue that the government should take a hands-off approach to macroeconomic policy. At first, this view might seem surprising. If our model shows how to prevent or reduce the severity of recessions, why do these critics want the government to refrain from using monetary and fiscal policy for economic stabilization? To find out, let's consider some of their arguments.

Lags in the Implementation and Effects of Policies

Economic stabilization would be easy if the effects of policy were immediate. Making policy would be like driving a car: policymakers would simply adjust their instruments to keep the economy on the desired path. Making economic policy, however, is less like driving a car than it is like piloting a large ship. A car changes direction almost immediately after the steering wheel is turned. By contrast, a ship changes course long after the pilot adjusts the rudder, and once the ship starts to turn, it continues turning long after the rudder is set back to normal. A novice pilot is likely to over steer and, after noticing the mistake, overreacts by steering too much in the opposite direction. The ship's path could become unstable, as the novice responds to previous mistakes by making larger and larger corrections.

Like a ship's pilot, economic policymakers face the problem of long lags. Indeed, the problem for policymakers is even more difficult, because the lengths of the lags are hard to predict. These long and variable lags greatly complicate the conduct of monetary and fiscal policy.

Economists distinguish between two lags in the conduct of stabilization policy: the inside lag and the outside lag. The **inside lag** is the time between a shock to the economy and the policy action responding to that shock. This lag arises because it takes time for policymakers first to recognize that a shock has occurred and then to put appropriate policies into effect. The **outside lag** is the time between a policy action and its influence on the economy. This lag arises because policies do not immediately influence spending, income, and employment.

A long inside lag is a central problem with using fiscal policy for economic stabilization. This is especially true in the United States, where changes in spending or taxes require the approval of the president and both houses of Congress. The slow and cumbersome legislative process often leads to delays, which make fiscal policy an imprecise tool for stabilizing the economy. This inside lag is shorter in countries with parliamentary systems, such as the United Kingdom, because there the party in power can often enact policy changes more rapidly.

Monetary policy has a much shorter inside lag than fiscal policy, because a central bank can decide on and implement a policy change in less than a day, but monetary policy has a substantial outside lag. Monetary policy works by changing the money supply and thereby interest rates, which in turn influence investment. But many firms make investment plans far in advance. Therefore, a change in monetary policy is thought not to affect economic activity until about six months after it is made.

The long and variable lags associated with monetary and fiscal policy certainly make stabilizing the economy more difficult. Advocates of passive policy argue that, because of these lags, successful stabilization policy is almost impossible. Indeed, attempts to stabilize the economy can be destabilizing. Suppose that the economy's condition changes between the beginning of a policy action and its impact on the economy. In this case, active policy may end up stimulating the economy when it is overheated or depressing the economy when it is cooling off. Advocates of active policy admit that such lags do require policymakers to be cautious. But, they argue,

these lags do not necessarily mean that policy should be completely passive, especially in the face of a severe and protracted economic downturn.

Some policies, called **automatic stabilizers**, are designed to reduce the lags associated with stabilization policy. Automatic stabilizers are policies that stimulate or depress the economy when necessary without any deliberate policy change. For example, the system of income taxes automatically reduces taxes when the economy goes into a recession, without any change in the tax laws, because individuals and corporations pay less tax when their incomes fall. Similarly, the unemployment insurance and welfare systems automatically raise transfer payments when the economy moves into a recession, because more people apply for benefits. One can view these automatic stabilizers as a type of fiscal policy without any inside lag.

The Difficult Job of Economic Forecasting

Because policy influences the economy only after a long lag, successful stabilization policy requires the ability to predict accurately future economic conditions. If we cannot predict whether the economy will be in a boom or a recession in six months or a year, we cannot evaluate whether monetary and fiscal policy should now be trying to expand or contract aggregate demand. Unfortunately, economic developments are often unpredictable, at least given our current understanding of the economy.

One way forecasters try to look ahead is with **leading indicators**. A leading indicator is a data series that fluctuates in advance of the economy. A large fall in a leading indicator signals that a recession is more likely.

Another way forecasters look ahead is with macro econometric models, which have been developed both by government agencies and by private firms for forecasting and policy analysis. These large-scale computer models are made up of many equations, each representing a part of the economy. After making assumptions about the path of the exogenous variables, such as monetary policy, fiscal policy, and oil prices, these models yield predictions about unemployment, inflation, and other endogenous variables.

Keep in mind, however, that the validity of these predictions is only as good as the model and the forecasters' assumptions about the exogenous variables.

5.2 Ignorance, Expectations, and the Lucas Critique

The prominent economist Robert Lucas once wrote, “As an advice-giving profession we are in way over our heads.” Even many of those who advise policymakers would agree with this assessment. Economics is a young science, and there is still much that we do not know. Economists cannot be completely confident when they assess the effects of alternative policies. This ignorance suggests that economists should be cautious when offering policy advice.

Although economists’ knowledge is limited about many topics, Lucas has emphasized the issue of how people form expectations of the future. Expectations play a crucial role in the economy because they influence all sorts of economic behavior.

For instance, households decide how much to consume based on expectations of future income, and firms decide how much to invest based on expectations of future profitability. These expectations depend on many things, including the economic policies being pursued by the government. Thus, when policymakers estimate the effect of any policy change, they need to know how people’s expectations will respond to the policy change. Lucas has argued that traditional methods of policy evaluation—such as those that rely on standard macro econometric models—do not adequately take into account this impact of policy on expectations.

This criticism of traditional policy evaluation is known as the **Lucas critique**. An important example of the Lucas critique arises in the analysis of disinflation.

As you may recall from Chapter 13, the cost of reducing inflation is often measured by the sacrifice ratio, which is the number of percentage points of GDP that must be forgone to reduce inflation by 1 percentage point. Because these estimates of the sacrifice ratio are often large, they have led some economists to argue that policymakers should learn to live with inflation, rather than incur the large cost of reducing it.

According to advocates of the rational-expectations approach, however, these estimates of the sacrifice ratio are unreliable because they are subject to the Lucas critique.

Traditional estimates of the sacrifice ratio are based on adaptive expectations, that is, on the assumption that expected inflation depends on past inflation. Adaptive expectations may be a

reasonable premise in some circumstances, but if the policymakers make a credible change in policy, workers and firms setting wages and prices will rationally respond by adjusting their expectations of inflation appropriately. This change in inflation expectations will quickly alter the short-run tradeoff between inflation and unemployment. As a result, reducing inflation can potentially be much less costly than is suggested by traditional estimates of the sacrifice ratio.

The Lucas critique leaves us with two lessons. The narrow lesson is that economists evaluating alternative policies need to consider how policy affects expectations and, thereby, behavior. The broad lesson is that policy evaluation is hard, so economists engaged in this task should be sure to show the requisite humility.

The Historical Record

In judging whether government policy should play an active or passive role in the economy, we must give some weight to the historical record. If the economy has experienced many large shocks to aggregate supply and aggregate demand, and if policy has successfully insulated the economy from these shocks, then the case for active policy should be clear. Conversely, if the economy has experienced few large shocks, and if the fluctuations we have observed can be traced to inept economic policy, then the case for passive policy should be clear. In other words, our view of stabilization policy should be influenced by whether policy has historically been stabilizing or destabilizing. For this reason, the debate over macroeconomic policy frequently turns into a debate over macroeconomic history.

Yet history does not settle the debate over stabilization policy. Disagreements over history arise because it is not easy to identify the sources of economic fluctuations. The historical record often permits more than one interpretation.

The Great Depression is a case in point. Economists' views on macroeconomic policy are often related to their views on the cause of the Depression. Some economists believe that a large contractionary shock to private spending caused the Depression. They assert that policymakers should have responded by stimulating aggregate demand. Other economists believe that the large fall in the money supply caused the Depression. They assert that the Depression would have been avoided if the Fed had been pursuing a passive monetary policy of increasing the money supply

at a steady rate. Hence, depending on one's beliefs about its cause, the Great Depression can be viewed either as an example of why active monetary and fiscal policy is necessary or as an example of why it is dangerous.

5.3 Should Policy Be Conducted by Rule or by Discretion?

A second topic of debate among economists is whether economic policy should be conducted by rule or by discretion. Policy is conducted by rule if policymakers announce in advance how policy will respond to various situations and commit themselves to following through on this announcement. Policy is conducted by discretion if policymakers are free to size up events as they occur and choose whatever policy seems appropriate at the time.

The debate over rules versus discretion is distinct from the debate over passive versus active policy. Policy can be conducted by rule and yet be either passive or active. For example, a passive policy rule might specify steady growth in the money supply of 3 percent per year. An active policy rule might specify that

$$\text{Money Growth} = 3\% + (\text{Unemployment Rate} - 6\%).$$

Under this rule, the money supply grows at 3 percent if the unemployment rate is 6 percent, but for every percentage point by which the unemployment rate exceeds 6 percent, money growth increases by an extra percentage point. This rule tries to stabilize the economy by raising money growth when the economy is in a recession. We begin this section by discussing why policy might be improved by a commitment to a policy rule. We then examine several possible policy rules.

Distrust of Policymakers and the Political Process

Some economists believe that economic policy is too important to be left to the discretion of policymakers. Although this view is more political than economic, evaluating it is central to how we judge the role of economic policy. If politicians are incompetent or opportunistic, then we may not want to give them the discretion to use the powerful tools of monetary and fiscal policy.

Incompetence in economic policy arises for several reasons. Some economists view the political process as erratic, perhaps because it reflects the shifting power of special interest groups. In

addition, macroeconomics is complicated, and politicians often do not have sufficient knowledge of it to make informed judgments.

This ignorance allows **charlatans** to propose incorrect but superficially appealing solutions to complex problems. The political process often cannot weed out the advice of charlatans from that of competent economists.

Opportunism in economic policy arises when the objectives of policymakers conflict with the well-being of the public. Some economists fear that politicians use macroeconomic policy to further their own electoral ends. If citizens vote on the basis of economic conditions prevailing at the time of the election, then politicians have an incentive to pursue policies that will make the economy look good during election years. A president might cause a recession soon after coming into office to lower inflation and then stimulate the economy as the next election approaches to lower unemployment; this would ensure that both inflation and unemployment are low on Election Day. Manipulation of the economy for electoral gain, called the **political business cycle**, has been the subject of extensive research by economists and political scientists.

Distrust of the political process leads some economists to advocate placing economic policy outside the realm of politics. Some have proposed constitutional amendments, such as a balanced-budget amendment, that would tie the hands of legislators and insulate the economy from both incompetence and opportunism.

The Time Inconsistency of Discretionary Policy

If we assume that we can trust our policymakers, discretion at first glance appears superior to a fixed policy rule. Discretionary policy is, by its nature, flexible. As long as policymakers are intelligent and benevolent, there might appear to be little reason to deny them flexibility in responding to changing conditions.

Yet a case for rules over discretion arises from the problem of **time inconsistency** of policy. In some situations policymakers may want to announce in advance the policy they will follow in order to influence the expectations of private decision makers. But later, after the private decision makers have acted on the basis of their expectations, these policymakers may be tempted to renege on their announcement. Understanding that policymakers may be inconsistent

over time, private decision makers are led to distrust policy announcements. In this situation, to make their announcements credible, policymakers may want to make a commitment to a fixed policy rule.

Time inconsistency is illustrated most simply in a political rather than an economic example—specifically, public policy about negotiating with terrorists over the release of hostages. The announced policy of many nations is that they will not negotiate over hostages. Such an announcement is intended to deter terrorists: if there is nothing to be gained from kidnapping hostages, rational terrorists won't kidnap any. In other words, the purpose of the announcement is to influence the expectations of terrorists and thereby their behavior.

But, in fact, unless the policymakers are credibly committed to the policy, the announcement has little effect. Terrorists know that once hostages are taken, policymakers face an overwhelming temptation to make some concession to obtain the hostages' release. The only way to deter rational terrorists is to take away the discretion of policymakers and commit them to a rule of never negotiating. If Policy makers were truly unable to make concessions; the incentive for terrorists to take hostages would be largely eliminated. The same problem arises less dramatically in the conduct of monetary policy.

Consider the dilemma of a Federal Reserve that cares about both inflation and unemployment. According to the Phillips curve, the tradeoff between inflation and unemployment depends on expected inflation. The Fed would prefer everyone to expect low inflation so that it will face a favorable tradeoff. To reduce expected inflation, the Fed might announce that low inflation is the paramount goal of monetary policy. But an announcement of a policy of low inflation is by itself not credible.

Once households and firms have formed their expectations of inflation and set wages and prices accordingly, the Fed has an incentive to renege on its announcement and implement expansionary monetary policy to reduce unemployment.

People understand the Fed's incentive to renege and therefore do not believe the announcement in the first place. Just as a president facing a hostage crisis is sorely tempted to negotiate their release, a Federal Reserve with discretion is sorely tempted to inflate in order to reduce

unemployment. And just as terrorists discount announced policies of never negotiating, households and firms discount announced policies of low inflation.

The surprising outcome of this analysis is that policymakers can sometimes better achieve their goals by having their discretion taken away from them. In the case of rational terrorists, fewer hostages will be taken and killed if policymakers are committed to following the seemingly harsh rule of refusing to negotiate for hostages' freedom. In the case of monetary policy, there will be lower inflation without higher unemployment if the Fed is committed to a policy of zero inflation.

(This conclusion about monetary policy is modeled more explicitly in the appendix to this chapter.) The time inconsistency of policy arises in many other contexts. Here are some examples:

- To encourage investment, the government announces that it will not tax income from capital. But after factories have been built, the government is tempted to renege on its promise to raise more tax revenue from them.
- To encourage research, the government announces that it will give a temporary monopoly to companies that discover new drugs. But after a drug has been discovered, the government is tempted to revoke the patent or to regulate the price to make the drug more affordable.
- To encourage good behavior, a parent announces that he or she will punish a child whenever the child breaks a rule. But after the child has misbehaved, the parent is tempted to forgive the transgression, because punishment is unpleasant for the parent as well as for the child.
- To encourage you to work hard, your professor announces that this course will end with an exam. But after you have studied and learned all the material, the professor is tempted to cancel the exam so that he or she won't have to grade it.

In each case, rational agents understand the incentive for the policymaker to renege, and this expectation affects their behavior. And in each case, the solution is to take away the policymaker's discretion with a credible commitment to a fixed policy rule.

Rules for Monetary Policy

Even if we are convinced that policy rules are superior to discretion, the debate over macroeconomic policy is not over. If the Fed were to commit to a rule for monetary policy, what rule should it choose? Let's discuss briefly three policy rules that various economists advocate.

Some economists, called **monetarists**, advocate that the Fed keep the money supply growing at a steady rate. The quotation at the beginning of this chapter from

Milton Friedman—the most famous monetarist—exemplifies this view of monetary policy. Monetarists believe that fluctuations in the money supply are responsible for most large fluctuations in the economy. They argue that slow and steady growth in the money supply would yield stable output, employment, and prices. Although a monetarist policy rule might have prevented many of the economic fluctuations we have experienced historically, most economists believe that it is not the best possible policy rule. Steady growth in the money supply stabilizes aggregate demand only if the velocity of money is stable. But sometimes the economy experiences shocks, such as shifts in money demand that cause velocity to be unstable. Most economists believe that a policy rule needs to allow the money supply to adjust to various shocks to the economy. A second policy rule that economists widely advocate is nominal GDP targeting.

Under this rule, the Fed announces a planned path for nominal GDP. If nominal GDP rises above the target, the Fed reduces money growth to dampen aggregate demand. If it falls below the target, the Fed raises money growth to stimulate aggregate demand. Because a nominal GDP target allows monetary policy to adjust to changes in the velocity of money, most economists believe it would lead to greater stability in output and prices than a monetarist policy rule.

A third policy rule that is often advocated is inflation targeting. Under this rule, the Fed would announce a target for the inflation rate (usually a low one) and then adjust the money supply when the actual inflation deviates from the target. Like nominal GDP targeting, inflation targeting insulates the economy from changes in the velocity of money. In addition, an inflation target has the political advantage that it is easy to explain to the public.

Notice that all these rules are expressed in terms of some nominal variable—the money supply, nominal GDP, or the price level. One can also imagine policy rules expressed in terms of real variables. For example, the Fed might try to target the unemployment rate at 5 percent. The

problem with such a rule is that no one knows exactly what the natural rate of unemployment is. If the Fed chose a target for the unemployment rate below the natural rate, the result would be accelerating inflation. Conversely, if the Fed chose a target for the unemployment rate above the natural rate, the result would be accelerating deflation. For this reason, economists rarely advocate rules for monetary policy expressed solely in terms of real variables, even though real variables such as unemployment and real GDP are the best measures of economic performance.

7.3 Conclusion: Making Policy in an Uncertain World

In this chapter we have examined whether policy should take an active or passive role in responding to economic fluctuations and whether policy should be conducted by rule or by discretion. There are many arguments on both sides of these questions. Perhaps the only clear conclusion is that there is no simple and compelling case for any particular view of macroeconomic policy. In the end, you must weigh the various arguments, both economic and political, and decide for yourself what kind of role the government should play in trying to stabilize the economy.

For better or worse, economists play a key role in the formulation of economic policy. Because the economy is complex, this role is often difficult. Yet it is also inevitable. Economists cannot sit back and wait until our knowledge of the economy has been perfected before giving advice. In the meantime, someone must advise economic policymakers. That job, difficult as it sometimes is, falls to economists.

The role of economists in the policymaking process goes beyond giving advice to policymakers. Even economists cloistered in academia influence policy indirectly through their research and writing. In the conclusion of *The General Theory*, John Maynard Keynes wrote that the ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed, the world is ruled by little else. Practical men, who believe themselves to be quite exempt from intellectual influences, are usually the slaves of some defunct economist.

Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back. This is as true today as it was when Keynes wrote it in 1936—except now that academic scribbler is often Keynes himself.

Chapter Six: Macroeconomic Growth Model

6.1 Introduction: Modeling Economic Growth

Economic growth is an increase in the production of economic goods and services in one period of time compared with a previous period. It can be measured in nominal or real (adjusted to remove inflation) terms. Traditionally, aggregate economic growth is measured in terms of gross national product (GNP) or gross domestic product (GDP), although alternative metrics are sometimes used.

In simplest terms, economic growth refers to an increase in aggregate production in an economy, which is generally manifested in a rise in national income.¹

Often, but not necessarily, aggregate gains in production correlate with increased average marginal productivity. That leads to an increase in incomes, inspiring consumers to open up their wallets and buy more, which means a higher material quality of life and standard of living.

Economic growth is understood as a dynamic process in which each economy establishes interactions between the different productive factors that, together with economic policies, allows the generation of a greater quantity of goods and services produced, which improves the wellbeing of the population.

A model of economic growth is based on economic theory to establish basic fundamental assumptions that allow proposing an interaction between the factors of production in order to explain the determinants of economic growth. Based on exogenous and endogenous growth theories, different models have been proposed to explain economic growth including the human factor, capital accumulation and technological change among other factors of production.

6.2 Theories of economic Growth

6.2.1 The Harrod–Domar Model

Following the publication of Keynes's *General Theory* in 1936, some economists sought to dynamize Keynes's static short-run theory in order to investigate the long-run dynamics of capitalist market economies. Roy Harrod (1939, 1948) and Evsey Domar (1946, 1947) independently developed theories that relate an economy's rate of growth to its capital stock.

It is one of the simplest growth theories, which extends the simple short run Keynesian model into the long-run. It is an extension of Keynesian short-term analysis of full employment and income theory and provides “a more comprehensive long period theory of output”

While Keynes emphasized the impact of investment on aggregate demand, Harrod and Domar emphasized how investment spending also increased an economy's productive capacity (a supply-side-effect).

A major strength of the Harrod–Domar model is its simplicity.

The model assumes

- a. The economy is closed and that there is no government economic activity.
- b. There are only two factors of production, labor (L) and capital (K), and in our simple version of the model, there is *no technical progress*.
- c. Labor is homogenous and exogenous, measured in its own units and grows at the constant natural rate of growth.
- d. The production exhibits constant returns to scale.
- e. Saving (s) is a fixed proportion of income (Y) that is $S = sY$ where s is both the average and marginal propensity of save. Investment (I) is autonomous and there is no depreciation.

- f. The potential level of national income (Y) is proportional to the quantity of capital and to quantity of labor. Thus, a given technology exhibiting fixed factor proportions (constant capital–labour ratio, K/L) and a fixed capital–output ratio (K/Y).

Assuming a two-sector economy (households and firms), we can write the simple national income equation as:

$$Y_t = C_t + S_t$$

Where $Y_t = GDP$, $C_t =$ consumption and $S_t =$ saving.

Equilibrium in this simple economy requires:

$$I_t = S_t$$

Substituting (2) into (1) yields:

$$Y_t = C_t + I_t$$

According to the model the *growth of real GDP* is assumed to be proportional to the share of investment spending (I) in GDP and for the growth of an economy, net additions to the capital stock are required. For example, \$3 of capital is always necessary to produce an annual \$1 stream of GDP. The accumulation of the capital stock over time is given in equation:

$$K_{t+1} = (1 - \delta) K_t + I_t$$

The relationship between the size of the total capital stock (K) and total GDP (Y) is known as the capital–output ratio ($K/Y = k$) and is *assumed fixed*. Given that we have defined $k = K/Y$, it also follows that $k = \Delta K / \Delta Y$ (where $\Delta K / \Delta Y$ is the incremental capital–output ratio, or ICOR). Assume that total new investment is determined by total savings. And assume that total saving is some fraction (s) of GDP (Y), as shown in equation:

$$S_t = sY_t$$

Since $K = kY_t$ and $I_t = S_t$, it follows that we can rewrite equation for accumulation of the capital stock as:

$$kY_{t+1} = (1 - \delta) kY_t + sY_t$$

Dividing through by k , simplifying, and subtracting Y_t from both sides of equation yields:

$$Y_{t+1} - Y_t = \left[\frac{s}{k} - \delta \right] Y_t$$

Dividing through by Y_t gives us equation:

$$[Y_{t+1} - Y_t]/Y_t = (s/k) - \delta$$

Here $[Y_{t+1} - Y_t]/Y_t$ is the growth rate of GDP (G), we can write the Harrod–Domar growth equation as:

$$\Delta Y/Y = (s/k) - \delta$$

This simply states that the growth rate of GDP is jointly determined by the savings ratio (s) divided by the capital–output ratio (k). The higher the savings ratio and the lower the capital–output ratio and depreciation rate, the faster will an economy grow. In the discussion that follows we will ignore the depreciation rate and consider the Harrod–Domar model as being represented by the equation:

$$\Delta Y/Y(G) = (s/k)$$

The model says that in the absence of government, the growth rate of national income will be directly or positively related to the savings ratio (i.e., the more an economy is able to save—and invest—out of a given GDP, the greater the growth of that GDP will be) and inversely or negatively related to the economy’s capital-output ratio (i.e., the higher k is, the lower the rate of GDP growth will be)

The economic logic of the model is for an economy to grow; it must save and invest a certain proportion of their GDP. The more they can save and invest, the faster they can grow. But the actual rate at which they can grow for any level of saving and investment, how much additional output can be had from an additional unit of investment, can be measured by the inverse of the capital-output ratio, k , because this inverse, $1/k$, is simply the output-capital or output-investment ratio.

It follows that multiplying the rate of new investment, $s = I/Y$, by its productivity, $1/k$, will give the rate by which national income or GDP will increase.

In a general way, technological progress can be expressed in the Harrod-Domar context as a decrease in the required *capital-output ratio*, giving more growth for a given level of investment.

Constraint to or obstacle to development

According to Harrod-Domar growth model, one of the most fundamental strategies of economic growth is simply to *increase the proportion of national income saved* (i.e., not consumed). For example, if we assume that the national capital-output ratio in some less developed country is,

say, 3% and the aggregate net saving ratio is 6% of GDP, it follows from Equation 10 that this country can grow at a rate of 2% per year because

$$\frac{\Delta Y}{Y} = G = (s/k) = 2\%$$

Now if the national net savings rate can somehow be increased from 6% to, say, 15%—through some combination of increased taxes, foreign aid, and general consumption sacrifices—GDP growth can be increased from 2% to 5% because now

$$\frac{\Delta Y}{Y} = G = (s/k) = 5\%$$

Countries that were able to save 15% to 20% of GDP could grow (“develop”) at a much faster rate than those that saved less. Moreover, this growth would then be self-sustaining. The mechanisms of economic growth and development, therefore, are simply a matter of increasing national savings and investment.

The main obstacle to or constraint on development, according to this theory, is the relatively *low level of new capital formation in most poor countries*.

But if a country wanted to grow at, say, a rate of 7% (target rate of GDP growth (G)*) per year and if it could not generate savings and investment at a rate of 21% of national income (assuming that k , the final aggregate capital-output ratio, is 3) but could only manage to save 15%, it could seek to fill this “savings gap” of 6% either through private saving or foreign aid.

If domestic sources of finance were inadequate to achieve the desired growth target, then foreign aid could fill the ‘savings gap’ (Riddell, 1987). Aid requirements (A_r) would simply be calculated as $s^* - s = A_r$ (Chenery and Strout, 1966).

Weakness of Harrod–Domar model

A *major weakness* of the Harrod–Domar approach is the assumption of a fixed capital–output ratio. The inverse of k ($1/k$) is the productivity of investment. Unfortunately, the productivity of investment is not a given, but it is determined by the efficiency of the policy framework and the incentive structures within which investment decisions are taken.

Furthermore, the model also assumed that aid inflows would go into investment one to one. But it soon became apparent that inflows of foreign aid, with the objective of closing the savings gap, did not necessarily boost total savings. Aid does not go into investment one to one. Indeed, in

many cases inflows of aid led to a reduction of domestic savings together with a decline in the productivity of investment (Griffin, 1970; White, 1992).

A further weakness of the Harrod–Domar framework is the assumption of zero substitutability between capital and labour (that is, production function with a *fixed factor proportions*). This is a ‘crucial’ but inappropriate assumption for a model concerned with *long-run* growth. The problem arises from the assumption of a production function with an *inflexible technology*. In the Harrod–Domar model the capital–output ratio (K/Y) and the capital–labour ratio (K/L) are assumed to be constant. This means to maintain equilibrium K and Y must always grow at the same rate. However, because the model also assumes a constant capital–labour ratio (K/L), K and L must also grow at the same rate.

Thus, whenever K and L do not grow at the same rate, the economy falls off its equilibrium ‘knife-edge’ growth path. However, the evidence is overwhelming that this property does not fit well with the actual experience of growth

6.2.2 Neoclassical Growth Model

✍ After the contributions of Solow (1956, 1957) and Swan (1956), the neoclassical model became the dominant approach to the analysis of growth, at least within academia. The Solow growth model shows how saving, population growth, and technological progress affect the level of an economy’s output and its growth over time.

The key assumptions of the Solow model are:

- i. The economy consists of one sector producing one type of commodity that can be used for either investment or consumption purposes;
- ii. The economy is closed to international transactions and the government sector is ignored;
- iii. All output that is saved is invested;
- iv. Since the model is concerned with the long run, the economy is always producing its potential (natural) level of total output;
- v. Solow abandons the Harrod–Domar assumptions of a fixed capital–output ratio (K/Y) and fixed capital–labour ratio (K/L);
- vi. The rate of technological progress, population growth and the depreciation rate of the capital stock are all determined exogenously.

Given these assumptions we can concentrate on developing the three key relationships in the Solow model, namely, the *production function*, the *consumption function* and the *capital accumulation* process.

The production function

The Solow growth model is built around the neoclassical aggregate production function and focuses on the *proximate* causes of growth:

$$Y = A_t F(K, L)$$

Where Y is real output, K is capital, L is the labour input and A_t is a measure of technology (that is, the way that inputs to the production function can be transformed into output) which is exogenous and depend on time. Sometimes, A_t is called ‘total factor productivity’. In the neoclassical theory of growth, technology is assumed to be a *public good*. Applied to the world economy this means that every country is assumed to share the same stock of knowledge which is freely available; that is, all countries have access to the same production function.

But many economists disagree with this approach and insist that there are significant technology gaps between nations. However, to progress with our examination of the Solow model we will continue to treat technology as a public good.

For simplicity, let us begin by first assuming a situation where there is no technological progress. Let us concentrate on the relationship between output per worker and capital per worker. We can therefore rewrite as:

$$Y = F(K, L)$$

The aggregate production function given satisfies the following three conditions.

- First, for all values of $K > 0$ and $L > 0$, the function exhibits positive but diminishing marginal returns with respect to both capital and labour; that is, $\partial F / \partial K > 0$, $\partial^2 F / \partial K^2 < 0$, $\partial F / \partial L > 0$, and $\partial^2 F / \partial L^2 < 0$.
- Second, the production function exhibits constant returns to scale such that $F(\lambda K, \lambda L) = \lambda Y$; that is, raising inputs by λ will also increase aggregate output by λ . Letting $\lambda = 1/L$ yields $Y/L = F(K/L)$. This assumption allows the equation to be written down in intensive form as:

$$y = f(k), \text{ where } f'(k) > 0, \text{ and } f''(k) < 0 \text{ for all } k$$

Where y is output per worker (Y/L) and k = capital per worker (K/L)

The equation states that output per worker is a positive function of the capital–labour ratio and exhibits diminishing returns. The key assumption of constant returns to scale implies that the economy is sufficiently large, so that the size of the economy, in terms of the labour force, has no influence on output per worker.

- Third, as the capital–labour ratio approaches infinity ($k \rightarrow \infty$) the marginal product of capital (MPK) approaches zero; as the capital–labour ratio approaches zero the marginal product of capital tends towards infinity ($MPK \rightarrow \infty$).

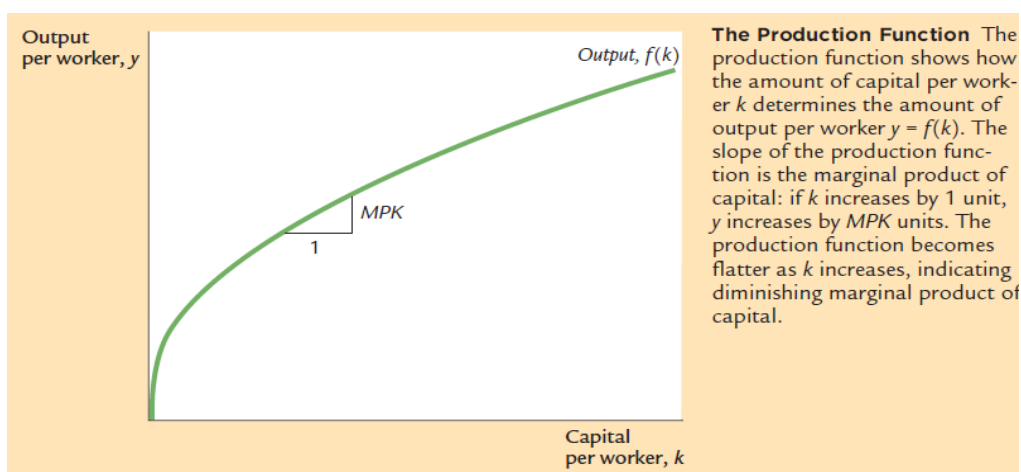


Figure 1: The neoclassical aggregate production function

As the diagram illustrates, for a given technology, any country that increases its capital–labour ratio (more equipment per worker) will have a higher output per worker. However, because of diminishing returns, the impact on output per worker resulting from capital accumulation per worker (capital deepening) will continuously decline. Thus for a given increase in k , the impact on y will be much greater *where capital is relatively scarce* than in economies where capital is relatively abundant.

The slope of the production function measures the marginal product of capital, where $MPK = f(k + 1) - f(k)$. It shows how much extra output a worker produces when given an extra unit of capital. In the Solow model the MPK should be much higher in developing countries compared to developed countries. In an open economy setting with no restrictions on capital mobility, we should therefore expect to see, *ceteris paribus*, capital flowing from rich to poor

countries, attracted by higher potential returns, thereby accelerating the process of capital accumulation.

The consumption function

Since output per worker depends positively on capital per worker, we need to understand how the capital–labour ratio evolves over time. In a closed economy aggregate output = aggregate income and comprises two components, namely, consumption (C) and investment (I) = Savings (S). Therefore we can write equation for income as:

$$Y = C + I$$

or equivalently

$$Y = C + S$$

Here $S = sY$ is a simple savings function where s is the fraction of income saved and $1 > s > 0$.

We can rewrite as:

$$Y = C + sY$$

Given the assumption of a closed economy, private domestic saving (sY) must equal domestic investment (I).

The capital accumulation process

A country's capital stock (K_t) at a point in time consists of plant, machinery and infrastructure. Each year a proportion of the capital stock wears out. Countering this tendency for the capital stock to decline is a flow of investment spending each year (I_t) that adds to the capital stock. Therefore, given these two forces, we can write an equation for the evolution of the capital stock of the following form:

$$K_{t+1} = I_t + (1 - \delta)K_t = sY_t + K_t - \delta K_t$$

Rewriting the above equation in per worker terms yields equation:

$$K_{t+1}/L = sY_t/L + K_t/L - \delta K_t/L$$

Deducting K_t/L from both sides of the above equation gives us:

$$K_{t+1}/L - K_t/L = sY_t/L - \delta K_t/L$$

In the neoclassical theory of growth the accumulation of capital evolves according to the following equation, which is the fundamental differential equation of the Solow model:

$$\dot{k} = \Delta k = sf(k) - \delta k$$

where $\dot{k} = \Delta k = K_{t+1}/L - K_t/L$ is the change of the capital input per worker, and $sf(k) = sy = sY_t/L$ is saving (investment) per worker. The $\delta k = \delta K_t/L$ term represents the ‘investment requirements’ per worker in order to keep the capital–labour ratio constant. The steady-state condition in the Solow model is given as:

$$sf(k^*) - \delta k^* = 0$$

Thus, in the steady state $sf(k^*) = \delta k^*$; that is, investment per worker is just sufficient to cover depreciation per worker, leaving capital per worker constant.

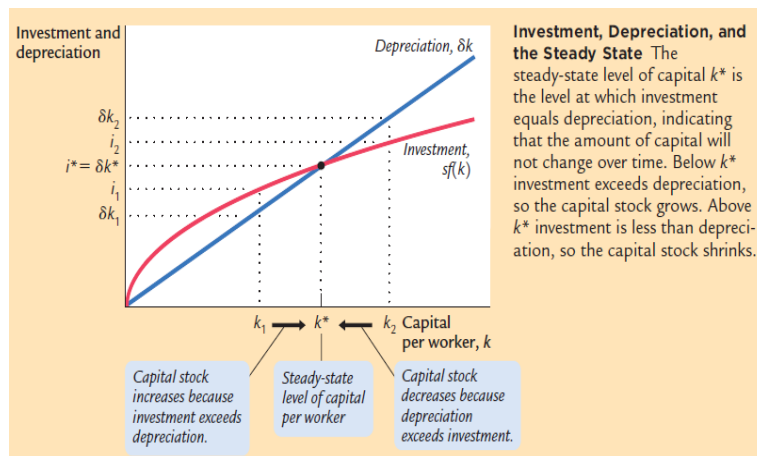


Figure 3 Investment, Depreciation and Steady- state

How Saving Affects Growth

Consider what happens to an economy when its saving rate increases. The economy is assumed to begin in a steady state with saving rate s_1 and capital stock k_1^* . When the saving rate increases from s_1 to s_2 , the $sf(k)$ curve shifts upward. At the initial saving rate s_1 and the initial capital stock k_1^* , the amount of investment just offsets the amount of depreciation. Immediately after the saving rate rises, investment is higher, but the capital stock and depreciation are unchanged. Therefore, investment exceeds depreciation. The capital stock will gradually rise until the economy reaches the new steady state k_2^* , which has a higher capital stock and a higher level of output than the old steady state.

§ The Solow model shows that the saving rate is a key determinant of the steady-state capital stock. *If the saving rate is high, the economy will have a large capital stock and a high level of*

output in the steady state. If the saving rate is low, the economy will have a small capital stock and a low level of output in the steady state.

☞ This conclusion sheds light on many discussions of fiscal policy. A government budget deficit can reduce national saving and crowd out investment. The long-run consequences of a reduced saving rate are a *lower capital stock and lower national income*. This is why many economists are critical of persistent budget deficits.

According to Solow model higher saving leads to faster growth, but only temporarily. An increase in the rate of saving raises growth only until the economy reaches the new steady state. If the economy maintains a high saving rate, it will maintain a large capital stock and a high level of output, but it will not maintain a high rate of *growth forever*. Policies that alter the steady-state growth rate of income per person are said to have a *growth effect*; By contrast, a higher saving rate is said to have a *level effect*, because only the *level of income per person* influenced by the saving rate in the steady state.

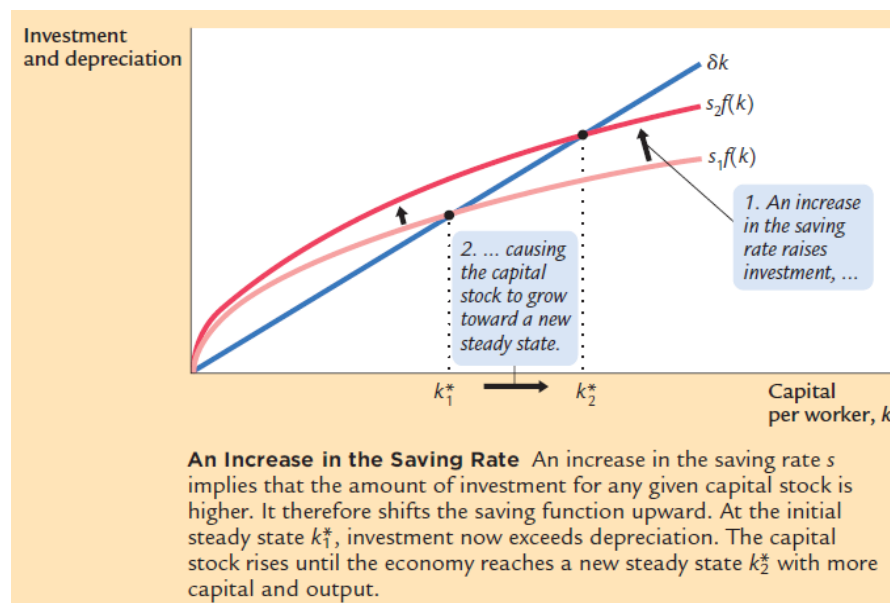


Figure 3: An Increase in the Saving Rate

The Golden Rule Level of Capital

So far, we have used the Solow model to examine how an economy's rate of saving and investment determines its steady-state levels of capital and income. This analysis might lead you

to think that higher saving is always a good thing because it always leads to greater income. Yet suppose a nation had a saving rate of 100 percent. That would lead to the largest possible capital stock and the largest possible income. But if all of this income is saved and none is ever consumed, what good is it?

Comparing Steady States

Assume that a policymaker can set the economy's saving rate at any level. By setting the saving rate, the policymaker determines the economy's steady state. What steady state should the policymaker choose?

🔔 The policymaker's goal is to maximize the well-being of individual who make up a society. Individuals themselves do not care about the amount of capital in the economy or amount of output. They care about the amount of goods and services they can consume. Thus, a benevolent policymaker would want to choose the steady-state with the highest level of consumption. The steady-state value of k that maximizes consumption is called the **Golden Rule level of capital** and is denoted k_{Gold}^* .

To identify whether an economy is at the Golden Rule level, we must first determine steady-state consumption per worker. Then we can see which steady state provides the most consumption. To find steady-state consumption per worker, we begin with the national income accounts identity

$$c = y - i$$

Then we substitute steady-state values for output and investment. Steady-state output per worker is $f(k^*)$, where k^* is the steady-state capital stock per worker. Furthermore, because the capital stock is not changing in the steady state, investment equals depreciation δk^* . Substituting $f(k^*)$ for y and δk^* for i , we can write steady-state consumption per worker as

$$c = f(k^*) - \delta k^*$$

Accordingly, steady-state consumption is what's left of steady-state output after paying for steady-state depreciation. This equation shows that an increase in steady-state capital has two opposing effects on steady-state consumption. On the one hand, more capital means more output.

On the other hand, more capital also means that more output must be used to replace capital that is wearing out.

Figure 4 graphs $f(k^*)$ and δk^* as a function of the steady-state capital stock. Steady-state consumption is the gap between output and depreciation. This figure shows that there is one level of the capital stock, the Golden Rule level k_{Gold}^* that maximizes consumption.

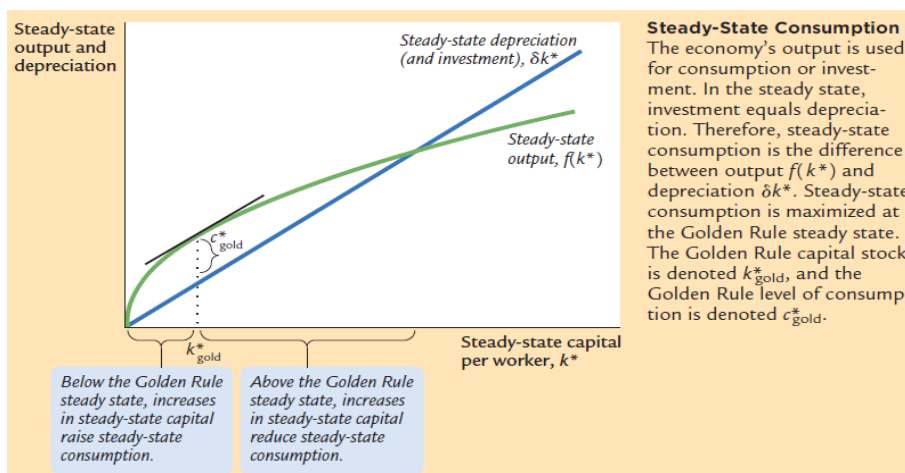


Figure 4: Steady-State and Consumption

We must keep in mind that higher levels of capital affect both output and depreciation. If the capital stock is *below* the Golden Rule level, an increase in the capital stock raises output more than depreciation, so consumption rises. In this case, the production function is steeper than the δk^* line, so the gap between these two curves, which equals consumption, grows as k^* rises. By contrast, if the capital stock is *above* the Golden Rule level, an increase in the capital stock reduces consumption, because the increase in output is smaller than the increase in depreciation. In this case, the production function is flatter than the δk^* line, so the gap between the curves, consumption, shrinks as k^* rises. At the Golden Rule level of capital, the production function and the δk^* line have the same slope, and consumption is at its greatest level.

We can now derive a simple *condition that characterizes the Golden Rule level of capital*. Recall that the slope of the production function is the marginal product of capital MPK . The slope of the δk^* line is δ . Because these two slopes are equal at k_{Gold}^* , the Golden Rule is described by the

equation $MPK = \delta$. At the Golden Rule level of capital, the marginal product of capital equals the depreciation rate.

To make the point somewhat differently, suppose that the economy starts at some steady-state capital stock k^* and that the policymaker is considering increasing the capital stock to $k^* + 1$. The MPK , the amount of extra output from this increase in capital, would be $f(k^* + 1) - f(k^*)$. The amount of extra depreciation from having 1 more unit of capital is the depreciation rate δ . Thus, the net effect of this extra unit of capital on consumption is $MPK - \delta$. If $MPK - \delta > 0$, then increases in capital increase consumption, so k^* must be below the Golden Rule level. If $MPK - \delta < 0$, then increases in capital decrease consumption, so k^* must be above the Golden Rule level. Therefore, the following condition describes the Golden Rule:

$$MPK - \delta = 0$$

As we will see, a policymaker can use this condition to find the Golden Rule capital stock for an economy.

Keep in mind that the economy does not automatically gravitate toward the Golden Rule steady-state. If we want any particular steady-state capital stock, such as the Golden Rule, we need a particular saving rate to support it. If the saving rate is higher than the one used in this figure, the steady-state capital stock will be too high. If the saving rate is lower, the steady-state capital stock will be too low. In either case, steady-state consumption will be lower than it is at the Golden Rule steady state.

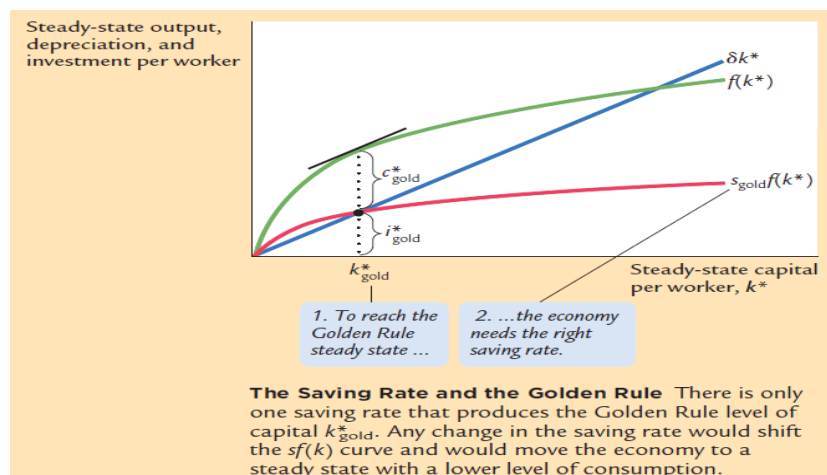


Figure 5: The Saving Rate and the Golden Rule

Population growth and steady-state capital accumulation

Extending the model to allow for growth of the labour force is relatively straightforward. In the Solow model it is assumed that the participation rate is constant, so that the labour force grows at a constant proportionate rate equal to the exogenously determined rate of growth of population = n . Because $k = K/L$, population growth, by increasing the supply of labour, will reduce k . Therefore population growth has the same impact on k as depreciation. The fundamental differential equation now becomes:

$$\dot{k} = \Delta k = sf(k) - (n + \delta)k$$

We can think of the expression $(n + \delta)k$ as the ‘required’ or ‘break-even’ investment necessary to keep the capital stock per unit of labour (k) constant.

In order to prevent k from falling, some investment (δk) is required to offset depreciation. Some investment is also required because the quantity of labour is growing at a rate = n . This is the (nk) . Hence the capital stock must grow at rate $(n + \delta)$ just to hold k steady. When investment per unit of labour is greater than required for break-even investment, then k will be rising and in this case the economy is experiencing ‘capital deepening’. Given the structure of the Solow model the economy will approach a steady state where actual investment per worker, $sf(k)$, equals break-even investment per worker, $(n + \delta)k$. In the steady state the change in capital per worker $\dot{k} = 0$, although the economy continues to experience ‘capital widening’, the extension

of existing capital per worker to additional workers. Using * to indicate steady-state values, we can define the steady state as:

$$sf(k^*) = (n + \delta)k^*$$

Figure 4 captures the essential features of the Solow model. In the top panel of Figure 4 the curve $f(k)$ graphs a well-behaved intensive production function; $sf(k)$ shows the level of savings per worker at different levels of the capital–labour ratio (k); the linear relationship $(n + \delta)k$ shows that break-even investment is proportional to k .

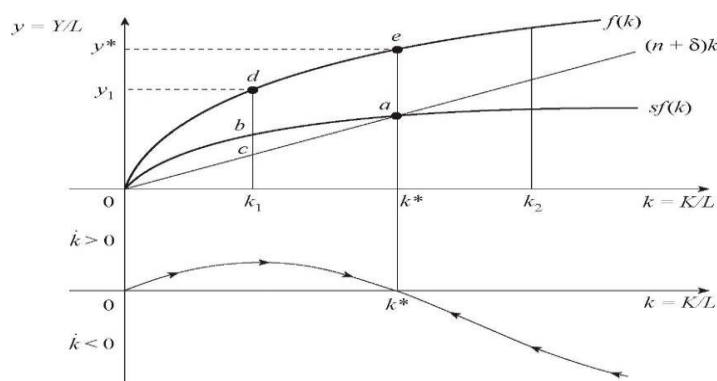


Figure 6: The Solow Growth Model

At the capital–labour ratio k_1 , savings (investment) per worker (point b) exceed required investment (point c) and so the economy experiences capital deepening and k rises. At k_1 consumption per worker is indicated by $d - b$ and output per worker is y_1 . At k_2 , because $(n + \delta)k > sf(k)$ the capital–labour ratio falls, capital becomes ‘shallower’ (Jones, 1975). The steady state balanced growth path occurs at k^* , where investment per worker equals break-even investment. Output per worker is y^* and consumption per worker is $e - a$. In the bottom panel of Figure 4 the relationship between $\dot{k}$ (the change of the capital–labour ratio) and k is shown with a phase diagram. When $\dot{k} = \Delta K > 0$, k is rising; when $\dot{k} = \Delta K < 0$, k is falling.

In the steady state equilibrium, shown as point a in the top panel of Figure 4, output per worker (y^*) and capital per worker (k^*) are constant. However, although there is no intensive growth in the steady state, there is extensive growth because population (and hence the labour input = L) is growing at a rate of n per cent per annum. Thus, in order for $y^* = Y/L$ and $k^* = K/L$ to remain constant, both Y and K must also grow at the same rate as population.

It can be seen from Figure 4 that the steady state level of output per worker will increase (*ceteris paribus*) if the rate of population growth and/or the depreciation rate are reduced (a downward pivot of the $(n + \delta)k$ function), and vice versa.

✎ The steady state level of output per worker will also increase (*ceteris paribus*) if the savings rate increases (an upward shift of the $sf(k)$ function), and vice versa. Of particular importance is the prediction from the Solow model that an increase in the savings ratio *cannot permanently increase the long-run rate of growth*. A higher savings ratio does temporarily increase the growth rate during the period of transitional dynamics to the new steady state and it *also permanently increases the level of output per worker*. Of course the period of transitional dynamics may be a long historical time period and level effects are important and should not be undervalued (see Solow, 2000; Temple, 2003).

The Effects of Population Growth

- Population growth alters the basic Solow model in three ways.

First, it brings us closer to explaining sustained economic growth. In the steady state with population growth, capital per worker and output per worker are constant. Because the number of workers is growing at rate n , however, *total* capital and *total* output must also be growing at rate n . Hence, although population growth cannot explain sustained growth in the standard of living (because output per worker is constant in the steady state), it can help explain sustained growth in total output.

Second, population growth gives us another explanation for why some countries are rich and others are poor. Consider the effects of an increase in population growth. Figure 5 shows that an increase in the rate of population growth from n_1 to n_2 reduces the steady-state level of capital per worker from k_1^* to k_2^* . Because k^* is lower and because $y^* = f(k^*)$, the level of output per worker y^* is also lower.

Thus, the Solow model predicts that countries with higher population growth will have lower levels of GDP per person. Notice that a change in the population growth rate, like a change in the saving rate, has a level effect on income per person but does not affect the steady-state growth rate of income per person.

Finally, population growth affects our criterion for determining the Golden Rule (consumption-maximizing) level of capital. To see how this criterion changes, note that consumption per worker is

$$c = y - i.$$

Because steady-state output is $f(k^*)$ and steady-state investment is $(\delta + n)k^*$, we can express steady-state consumption as

$$c^* = f(k^*) - (\delta + n)k^*.$$

Using an argument largely the same as before, we conclude that the level of k^* that maximizes consumption is the one at which

$$MPK = \delta + n,$$

or equivalently,

$$MPK - \delta = n.$$

In the Golden Rule steady state, the marginal product of capital net of depreciation equals the rate of population growth.

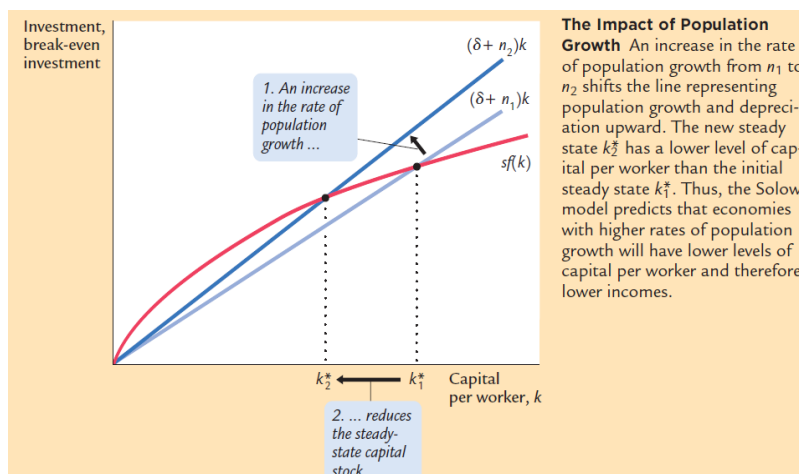


Figure 7 the impact of population growth

Technological Progress and Solow Model

So far, Solow model has assumed that the production function is not changing. In other words, the relationship between the inputs of capital and labor and the output of goods and services are not changing. Now let us assume that technological progress is exogenous which over time expands society's ability to produce.

To incorporate technological progress, let us take the production function which relates total capital K and total labor L to total output Y .

$$Y = f(K, L)$$

We now write the production function as

$$Y = f(K, L \times E),$$

Where, E is a new variable called the **efficiency of labor**. The efficiency of labor is nothing but society's knowledge about production methods. This is due to the fact that as the available technology improves, the efficiency of labor increases. For example, we can compare the efficiency of labors before and after the inventions of machines. Certainly we can find that now labors are more efficient than olden times. The efficiency of labor also rises when there are improvements in the health, education, or skills of the labor force.

The term $L \times E$ measures the number of *effective workers*. It takes into account the number of workers L and the efficiency of each worker E . This new production function states that total output Y depends on the number of units of capital K and on the number of effective workers $L \times E$. Increases in the efficiency of labor E are, in effect, like increases in the labor force L .

The simplest assumption about technological progress is that it causes the efficiency of labor E to grow at some constant rate ' g '. If we assume that $g = 0.05$, then each unit of labor becomes 5 percent more efficient each year. This form of technological progress is called *labor augmenting*, and ' g ' is called the rate of **labor augmenting technological progress**. Because the labor force L is growing at rate ' n ', and the efficiency of each unit of labor E is growing at rate ' g ', the number of effective workers $L \times E$ is growing at rate $n + g$.

The effect of Labor augmenting technological progress has similar affect like in case of population growth. In case of population growth we analyzed the economy in terms of output per labour and allowed the number of labors to rise over time. Now we are analyzing the economy in terms of output per effective worker and allow the number of effective workers to rise. $k = K/L \times E$ and $y = Y/L \times E$

In order to do this, we have to modify our notation slightly.

$k = \frac{K}{L \times E}$ Which stand for capital per effective labour and similarly

$y = \frac{Y}{L \times E}$ This stands for output per effective labour .With these minor modifications, we can again write $y = f(k)$.

Our analysis of the economy proceeds just as it did when we examined population growth. The equation showing the evolution of k over time now changes to

$$\dot{k} = sy - (\delta + \eta + g)k$$

$$\Delta k = sy - (\delta + \eta + g)k$$

As before, the change in the capital stock Δk equals investment sy minus depreciation along with population growth and efficiency rate $(\delta + n + g)k$. Now, only difference is that $k = \frac{K}{L \times E}$, depreciation line includes three terms i.e., δk is needed to replace depreciating capital, nk is needed to provide capital for new workers, and gk is needed to provide capital for the new “effective workers” created by technological progress.

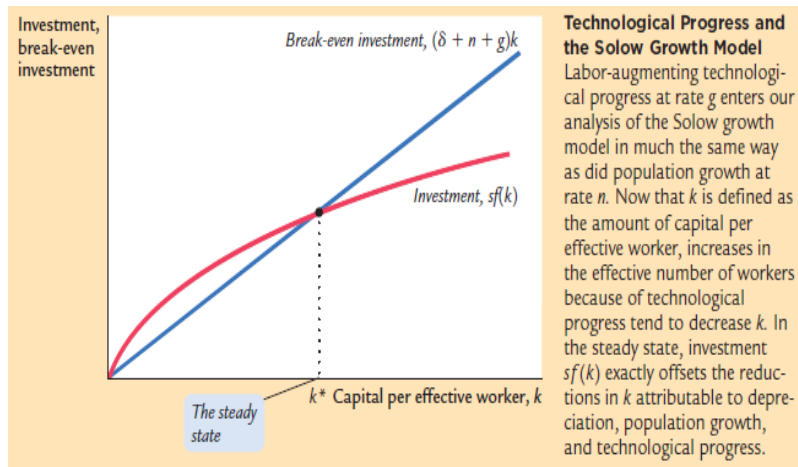


Figure 8 Technological Progresses and the Solow Growth Model

Now that k is defined as the amount of capital per effective worker, increase in the number of effective workers because technological progress tends to decrease k . In the steady state, investment $sf(k)$ exactly offsets the reductions in k attributable to depreciation, population growth, and technological progress.

☞ We have already seen three cases in Solow model i.e. with capital accumulation, with population growth and with technological progress. The key concepts of Solow model are as follows.

- ☞ The growth rate of output in steady state is exogenous and independent of savings rate.
- ☞ Although an increase in the savings rate does not affect the steady state growth rate, it does increase the steady state level of income by increasing the capital labour ratio.
- ☞ When we allow for productivity growth i.e. technological improvement, we can show that if there is steady state, the steady state growth rate of output remains exogenous.
- ☞ The final prediction of the neo-classical theory is about convergence. If two countries have the same rate of population growth, the same savings rate and access to the same production function, they will eventually reach the same level of income. Following this idea one can conclude that poor countries are poor because they have less capital, but if they save at the same rate as rich countries and have access to same technology, they may eventually catch up.

6.3 The Endogenous Growth Theory

One goal of growth theory is to explain the persistent rise in living standards that we observe in most parts of the world. The Solow growth model shows that such persistent growth must come from technological progress. But where does technological progress come from? In the Solow model, it is just assumed!

The Basic Model

To illustrate the idea behind endogenous growth theory, let's start with a particularly simple production function:

$$Y = AK,$$

where Y is output, K is the capital stock, and A is a constant measuring the amount of output produced for each unit of capital. Notice that this production function does not exhibit the property of diminishing returns to capital. One extra unit of capital produces A extra units of output, regardless of how much capital there is. This absence of diminishing returns to capital is the key difference between this endogenous growth model and the Solow model.

As before, we assume a fraction s of income is saved and invested. We therefore describe capital accumulation with an equation similar to those we used previously:

$$\Delta K = sY - \delta K$$

This equation states that the change in the capital stock (ΔK) equals investment (sY) minus depreciation (δK). Combining this equation with the $Y = AK$ production function, we obtain, after a bit of manipulation,

$$\Delta Y/Y = \Delta K/K = sA - \delta$$

This equation shows what determines the growth rate of output $\Delta Y/Y$. Notice that, as long as $sA > \delta$, the economy's income grows forever, even without the assumption of exogenous technological progress. Thus, a simple change in the production function can alter dramatically the predictions about economic growth. In the Solow model, saving leads to growth temporarily, but diminishing returns to capital eventually force the economy to approach a steady state in which growth depends only on exogenous technological progress. By contrast, in this endogenous growth model, saving and investment can lead to persistent growth.

But is it reasonable to abandon the assumption of diminishing returns to capital? The answer depends on how we interpret the variable K in the production function $Y = AK$. If we take the traditional view that K includes only the economy's stock of plants and equipment, then it is natural to assume diminishing returns. Giving 10 computers to a worker does not make that worker 10 times as productive as he or she is with one computer.

Advocates of endogenous growth theory, however, argue that the assumption of constant returns to capital is more palatable if K is interpreted more broadly. The endogenous growth model views *knowledge as a type of capital*.

✎ Clearly, knowledge is an important input into the economy's production, both its production of *goods and services* and its production of *new knowledge*. Compared to other forms of capital, however, it is less natural to assume that knowledge exhibits the property of diminishing returns. (Indeed, the increasing pace of scientific and technological innovation over the past few centuries has led some economists to argue that there are increasing returns to knowledge.) If we accept the view that knowledge is a type of capital, then this endogenous growth model with its assumption of constant returns to capital becomes a more plausible description of long-run economic growth.

A Two-Sector Model

Although the $Y = AK$ model is the simplest example of endogenous growth, the theory has gone well beyond this. One line of research has tried to develop models with more than one sector of production in order to offer a better description of the forces that govern technological progress. To see what we might learn from such models, let's sketch out an example.

The economy has two sectors, which we can call *manufacturing firms* and *research universities*. Firms produce goods and services, which are used for consumption and investment in physical capital. Universities produce a factor of production called "knowledge," which is then freely

used in both sectors. The economy is described by the production function for firms, the production function for universities, and the capital-accumulation equation:

$$\begin{aligned} Y &= F[K, (1 - u)LE] && \text{(Production function in manufacturing firms),} \\ \Delta E &= g(u)E && \text{(production function in research universities),} \\ \Delta K &= sY - \delta K && \text{(capital accumulation),} \end{aligned}$$

where u is the fraction of the labor force in universities (and $1 - u$ is the fraction in manufacturing), E is the *stock of knowledge* (which in turn determines the efficiency of labor), and g is a function that shows *how the growth in knowledge depends on the fraction of the labor force in universities*. The rest of the notation is standard. As usual, the production function for the manufacturing firms is assumed to have *constant returns to scale*: if we double both the amount of physical capital (K) and the effective number of workers in manufacturing $[(1 - u)LE]$, we *double the output* of goods and services (Y).

This model is a cousin of the $Y = AK$ model. Most important, this economy exhibits constant (rather than diminishing) returns to capital, as long as capital is broadly defined to include knowledge. In particular, if we double both physical capital K and knowledge E , then we double the output of both sectors in the economy. As a result, like the $Y = AK$ model, this model can generate persistent growth without the assumption of exogenous shifts in the production function. Here *persistent growth arises endogenously because the creation of knowledge in universities never slows down*.

At the same time, however, this model is also a *cousin of the Solow growth model*. If u , the fraction of the labor force in universities, is held constant, then the efficiency of labor E grows at the constant rate $g(u)$. This result of constant growth in the efficiency of labor at rate g is precisely the assumption made in the Solow model with technological progress. Moreover, the manufacturing productions function and the capital-accumulation equation, also resembles the rest of the Solow model. As a result, for any given value of u , this endogenous growth model works just like the Solow model.

There are two key decision variables in this model. As in the Solow model, the fraction of output used for saving and investment, s , determines the steady-state stock of physical capital. In addition, the fraction of labor in universities, u , determines the growth in the stock of knowledge. Both s and u affect the level of income, although *only u affects the steady-state growth rate of income*. Thus, this model of endogenous growth takes a small step in the direction of showing which societal decisions determine the rate of technological change.

The Microeconomics of Research and Development

The two-sector endogenous growth model just presented takes us closer to understanding technological progress, but it still tells only a rudimentary story about the creation of knowledge.

If one thinks about the process of research and development for even a moment, three facts become apparent. First, although knowledge is largely a public good (that is, a good freely available to everyone), much research is done in *firms that are driven by the profit motive*. Second, research is profitable because innovations give firms *temporary monopolies*, either because of the patent system or because there is an advantage to being the first firm on the market with a new product. Third, when one firm innovates, other firms build on that innovation to produce the next generation of innovations.

These (essentially microeconomic) facts are not easily connected with the (essentially macroeconomic) growth models we have discussed so far. Some endogenous growth models try to incorporate these facts about research and development. Doing this requires modeling both the decisions that firms face as they engage in research and the interactions among firms that have some degree of monopoly power over their innovations. Going into more detail about these models is beyond the scope of this book, but it should be clear already that *one virtue of these endogenous growth models is that they offer a more complete description of the process of technological innovation*.

One question these models are designed to address is whether, from the standpoint of society as a whole, private profit-maximizing firms tend to engage in too little or too much research. In other words, *is the social return to research* (which is what society cares about) *greater or smaller than the private return* (which is what motivates individual firms)? It turns out that, as a theoretical matter, there are effects in both directions. On the one hand, when a firm creates a new technology, it makes other firms *better off* by giving them a base of knowledge on which to build in future research. As Isaac Newton famously remarked, “If I have seen farther than others, it is because I was *standing on the shoulders of giants*.” On the other hand, when one firm invests in research, it can also make other firms *worse off* if it does little more than being the first to discover a technology that another firm would have invented in due course. This duplication of research effort has been called the “stepping on toes” effect. Whether firms left to their own devices do too little or too much research depends on whether the positive “standing on shoulders” externality or the negative “stepping on toes” externality is more prevalent.

Although theory alone is ambiguous about whether research effort is more or less than optimal, the empirical work in this area is usually less so. Many studies have suggested the “standing on shoulders” externality is important and, as a result, the social return to research is large, often in excess of 40 percent per year. This is an impressive rate of return, especially when compared to the return to physical capital. In the judgment of some economists, this finding justifies substantial government subsidies to research.

The Process of Creative Destruction

In his 1942 book *Capitalism, Socialism, and Democracy*, economist Joseph Schumpeter suggested that economic progress comes through a process of “creative destruction.” According

to Schumpeter, the driving force behind progress is the *entrepreneur* with an idea for a new product, a new way to produce an old product, or some other innovation. When the entrepreneur's firm enters the market, it has some degree of monopoly power over its innovation; indeed, it is the prospect of monopoly profits that motivates the entrepreneur. The entry of the new firm is good for consumers, who now have an expanded range of choices, but it is often bad for incumbent producers, who may find it hard to compete with the entrant.

If the new product is sufficiently better than old ones, the incumbents may even be driven out of business. Over time, the process keeps renewing itself. The entrepreneur's firm becomes *an incumbent*, enjoying high profitability until its product is displaced by another entrepreneur with the next generation of innovation.

History confirms Schumpeter's thesis that there are winners and losers from technological progress. For example, in England in the early nineteenth century, an important innovation was the invention and spread of machines that could produce textiles using unskilled workers at low cost. This technological advance was good for consumers, who could clothe themselves more cheaply. Yet skilled knitters in England saw their jobs threatened by new technology, and they responded by organizing violent revolts. The rioting workers, called *Luddites*, smashed the weaving machines used in the wool and cotton mills and set the homes of the mill owners on fire (a less than creative form of destruction). Today, the term "Luddite" refers to anyone who opposes technological progress.

A more recent example of creative destruction involves the retailing giant Wal-Mart. Although retailing may seem like a relatively static activity, in fact it is a sector that has seen sizable rates of technological progress over the past several decades. Through better inventory-control, marketing, and personnel management techniques, for example, Wal-Mart has found ways to bring goods to consumers at lower cost than traditional retailers. These changes benefit consumers, who can buy goods at lower prices, and the stockholders of Wal-Mart, who share in its profitability. But they adversely affect small mom-and pop stores, which find it hard to compete when a Wal-Mart opens nearby.

Faced with the prospect of being the victims of creative destruction, incumbent producers often look to the political process to stop the entry of new, more efficient competitors. The original Luddites wanted the British government to save their jobs by restricting the spread of the new textile technology; instead, Parliament sent troops to suppress the Luddite riots. Similarly, in recent years, local retailers have sometimes tried to use local land-use regulations to stop Wal-Mart from entering their market. The cost of such entry restrictions, however, is to slow the pace of technological progress. In Europe, where entry regulations are stricter than they are in the United States, the economies have not seen the emergence of retailing giants like Wal-Mart; as a result, productivity growth in retailing has been much lower.

Schumpeter's vision of how capitalist economies work has merit as a matter of economic history. Moreover, it has inspired some recent work in the theory of economic growth. One line of endogenous growth theory, pioneered by economists Philippe Aghion and Peter Howitt, builds on Schumpeter's insights by modeling technological advance as a process of entrepreneurial innovation and creative destruction.

Chapter Seven: African Macroeconomics

7.1 Basic Features of African Economies

With the exception of South Africa and the countries of North Africa, all of which have diversified production systems, the economy of most of Africa can be characterized as under developed. Africa as a whole has abundant natural resources, but much of its economy has remained predominantly agricultural, and subsistence farming still engages more than 60 percent of the population.

Until the beginning of the 20th century this system of farming relied on simple tools and techniques, as well as on traditional organization of the family or community for its labour. Because of poor transport and communications, production was largely for domestic use. There was little long-distance trade, and wage labour was virtually unknown. The small size and vast heterogeneity of polities at that time also made exchanges very limited. There were, however, notable exceptions, especially in western Africa, where for many centuries societies had engaged in long-distance trade and had elaborate exchange and craft facilities, communications, and a political infrastructure to maintain their trade routes.

Africa experienced considerable economic development during the 20th century, and, while this provided many benefits, it also gave rise to a number of serious problems. The first significant changes occurred under colonial rule in the first half of the century: wage labour was introduced, transportation and communications were improved, and resources were widely developed in the colonial territories. The legacy of this, however, has been that the export of two or three major agricultural products or minerals—such as peanuts, petroleum, or copper—has come to provide most of the foreign-exchange earnings for nearly all African countries. Fluctuations in the prices of these commodities have made the economies of these countries vulnerable and fragile. The situation has been exacerbated in countries in the marginal dry land zones, where the increasing frequency of drought conditions have undermined agricultural productivity.

The second major change was the vigorous promotion of industrial development, often with foreign assistance, that took place in the two decades (1960–80) following the political independence of most African countries. The political fragmentation of the continent, however,

also became a major constraint to industrial growth, because it created numerous small markets. Consequently, most African countries became saddled with excess industrial capacity, coupled with enormous foreign debts incurred in large part to build this capacity.

In nearly all African countries a poor economic situation has been aggravated by rapid population growth, which has kept per capita gross domestic product low or in some cases caused it to decline. Thus, any hope for improving economic conditions in most of Africa rests on two factors: population control within individual countries to give their economies the chance to grow; and the organization of groups of states into regional economic blocs in order to create internal markets large enough to sustain growth.

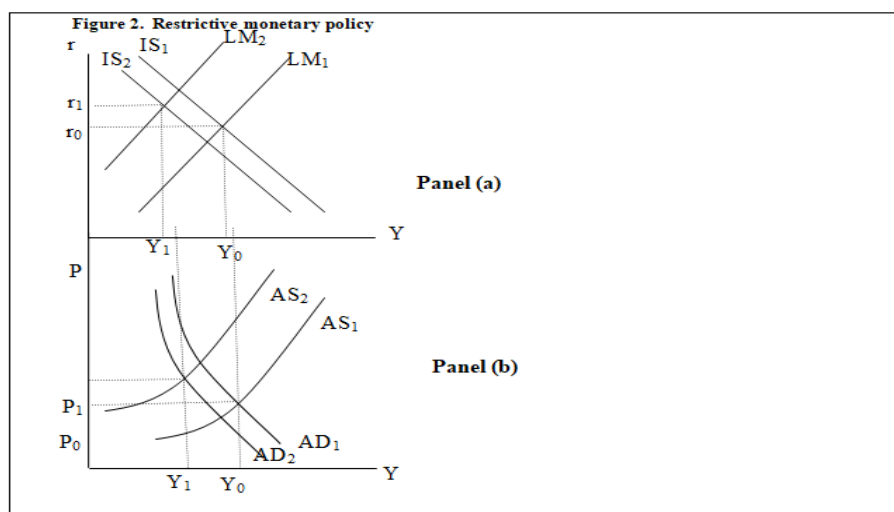
7.2 The Applicability of Conventional Theories to African Economies

Let us now come back to our earlier question as to whether conventional short-run stabilization theories are applicable in African economies as well. Conventionally, contractionary monetary policy helps to reduce price level (control inflation); devaluation improves balance of payment (BOP), and so on. But do these keep true for Africa? In line with the unique features of the IS-LM curves in Africa let us see the effectiveness of classical economists' policy prescriptions:

- ✓ Tight/contractionary monetary policy
- ✓ Tight/ contractionary fiscal policy and
- ✓ Devaluation

i. Contractionary Monetary Policy

In principle, according to conventional theory, the standard result of contractionary monetary policy is to reduce output and prices. Let us employ the IS-LM model to check whether standard results hold true for Africa. We may remember that since investment demand and money demand are less sensitive to changes in interest rate (inelastic) both IS and LM curves are steeper for less developed countries (African countries). The impact of monetary contraction is depicted in Figure 2



As the money supply is reduced, the LM curve shifts back from LM_1 to LM_2 in panel **a** and aggregate supply curve shifts back from AS_1 to AS_2 in panel **b**. Reduced money supply increases interest rate and thus investment demand will be reduced leading to a backward shift in IS curve from IS_1 to IS_2 . Since people move to curb market the shift in IS is smaller than the shift in LM curve.

So aggregate demand curve also shifts back from AD_1 to AD_2 in panel **b**, which is less than the shift in aggregate supply resulting in higher price level and smaller output level.

Thus, in less developed countries such as African countries the effect of contractionary monetary policy is to increase price and reduce output. The effect on price level is opposite to the conventional one. Please remember investment is less sensitive to interest rate change in less developed countries because in LDCs people use money for transaction purpose not that much for speculation or investment purpose.

In summary, if we introduce the characteristics of LDCs, contractionary monetary policy ends up with inflation as opposed to the conventional theory.

Restrictive Fiscal Policy

To investigate the effect of contractionary fiscal policy in less developed countries (Africa), let's use the IS-LM and AD-AS models. The conventional results of restrictive fiscal policy are also reduced output and reduced price.

While restrictive monetary policy is likely to increase both unemployment and inflation in the LDC in the short run, restrictive fiscal policy is likely to be more successful in reducing the price level without the costs of a major recession.

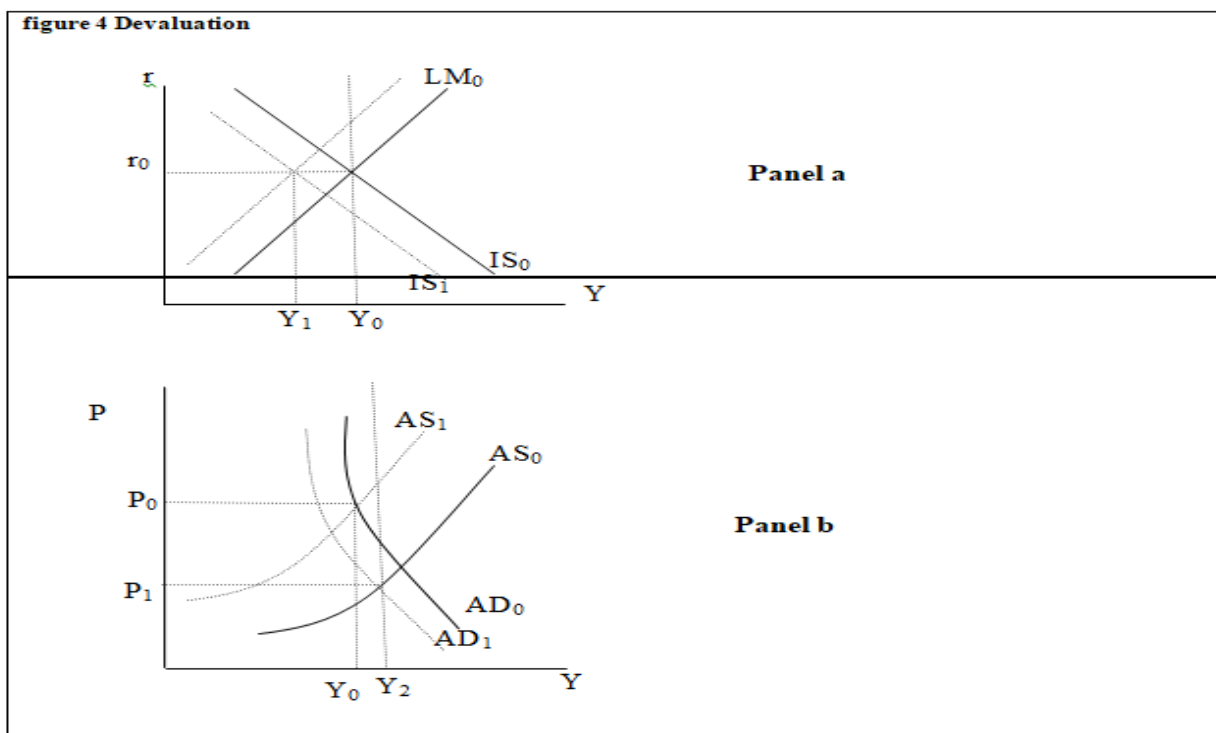
A fiscal contraction (for instance reduced government expenditure or increased tax) shifts IS curve in panel **a** from IS_0 to IS_1 and this shifts the aggregate demand curve in panel **b** down from AD_0 to AD_1 . The reduced interest rate reduces the cost of borrowing leading to increased output and so shifting aggregate supply curve from AS_0 to AS_1 .

Devaluation

The difference between the LDCs and more developed nations models are perhaps most evident when considering devaluation. In the context of developed countries, devaluation encourage export and so the IS curve shifts upward (to the right). But in less developed countries because of their agrarian economic structure to increase the amount to be exported (for instance coffee) may take several years. That is, there is supply side rigidity and devaluation doesn't encourage export that much. On the other hand, since such economy depends highly on imported inputs, devaluation (increased exchange rate) increases cost of import and hence production. Thus the aggregate supply curve shifts backward and so lower output level. In this case increased import of factors of production dominates (exceeds) an increase in export (if any).

$$Y \downarrow = C + I + G + X \uparrow - M \uparrow$$

Another effect of devaluation is through currency substitution. As exchange rate increases (devaluation) demand for money decreases and so LM curve shifts back reducing the aggregate supply (output).



- ❖ Several macroeconomic theory and policy options have been developed over long period of time. Very recently different researches have been developed new theories or new versions of earlier theories have been developed some being reactions to, some being supports of or complements to, and/or evaluation of earlier macroeconomic theories and policies. In this unit, two major such issues have been discussed. These are the recent macroeconomic policy debate by the new classical economists and applicability of the conventional macroeconomics policies to less developed economies such as Africa.
- ❖ The new classical economists argue that government policy instruments or intervention are generally ineffective. These economists have explained their policy positions using three common hypotheses namely, rational expectation, continuous market clearing assumption and aggregate supply hypothesis. The policy implications of the new classical economists lie in the explanation of the concepts such as policy effectiveness hypothesis, output and employment cost of reducing inflation, dynamic time inconsistency and policy credibility, Lucas critique, and so on all of them explaining that government policy is ineffective.
- ❖ The second section focuses on the peculiarities of aggregate supply and aggregate demand in Africa and its impact on the direct application of standard macroeconomic theories for an African economy. In African economy the aggregate supply function is entirely different

from the advanced countries. This is due to the fact that industrial structure is oligopoly with import substitution policy. Labor market on the other hand is highly imperfect which is characterized by premature labor unions and significant wage rigidities accompanied by excess labor supply. The implications of these characteristics of African economies on aggregate supply are that the aggregate supply is flat.

- ❖ LM curve is steeper in developing countries resulting from low interest elasticity of money demand. In most cases money serves for transaction purpose not for speculation in LDCS. The IS also steeper in LDCS because of low short run elasticity of investment demand. The problem of investors is not interest rate but the problem is whether to find the fund or not. The interaction of the two above (steeper LM and steeper IS curves) leads to steeper aggregate demand.
- ❖ For classical economists the government should simply follow tight monetary and fiscal policies (restriction of government intervention) and by extension their suggestion for Africa's economy is for governments be less interventionist. For structural economists Africa's macroeconomic problems are direct results of the structure of the economy and hence we need government intervention for structural transformation.
- ❖ Moreover, when standard macroeconomic theories are applied to Africa's economy is the result could be very different. For example, contractionary monetary policy helps to control inflation. But due to the abovementioned unique features of the economy, it is very likely for such a policy to increase price and reduce output. Likewise, it is possible for contractionary monetary policy to result in reduced price and increased output level, contrary to the standard macro models.
- ❖ In the context of developed countries, devaluation encourages export but in less developed countries because of supply side rigidities devaluation doesn't encourage export that much while it makes imports expensive. As a result, the outcome of a devaluation measure is contrary to its effect in the developed, at least in the short run.

References

N. Gregory Mankiw, 2007, Macroeconomics 4th edition Worth Publishers USA

William H. Branson, 2006 Macroeconomic Theory and Policy

Eduard Shapiro, 2007. Macroeconomic Analysis 5th edition New Delhi

Olivier Blanchard, 1997. Macroeconomics Prentice Hall Inc. New jersey USA

Frederick Mishkin, 1998. The Economics of Money, Banking and Financial Markets 5th edition USA