Economic Fluctuations

History doesn't repeat itself but it often rhymes. – Mark Twain

1 Introduction

In this chapter, we study economic fluctuation, which may be characterized by an alternating cycle of recessions and expansions. For business owners and managers weighing investment decisions, "macro traders" who want to anticipate major economic trends and policy changes, politicians facing forthcoming elections, or technocrats in some branches of the government (e.g., the central bank), economic fluctuations are almost all they care about macroeconomics.

Indeed, macroeconomics was born because of the Great Depression, the most severe recession in the twentieth century. During the Great Depression, the public found the classical approach to economic policy making no longer acceptable. The classical approach was to do nothing and wait for the economy to rebound. Classical economists would argue that the market was self-correcting and that the downturn was part of the self-correction from the previous boom. However deep the recession was, classical economists would say, the economy would get back to its potential level with full employment in the long run. But in the long run, as Keynes quipped, we would be all dead. The short-term pain of the recessions, particularly the mass unemployment, could no longer be ignored by the increasingly more democratic governments in the world.

The Great Depression prompted John Maynard Keynes to propose a theory that explained why the Depression happened and what the government could do in such circumstances. In his magnum opus, *The General Theory of Employment, Interest, and Money*, Keynes not only made a convincing diagnosis and prescription but also revolutionized the way we think about the economy as a whole. Treating the economy as a whole, Keynes introduced aggregate concepts such as aggregate demand, aggregate supply, effective demand, and so on. Only after the Keynesian revolution did people find it necessary to have a separate discipline within economics – macroeconomics, different from microeconomics, which mainly deals with individual behavior.

In this chapter, we first give an empirical overview of the business cycles using Chinese and US data. Then we introduce a group of Keynesian models that rely on the sticky-price assumption. The Keynesian approach came into being as a group of neoclassical Keynesian economists, including John Hicks and Paul Samuelson, formalized Keynes writing into mathematical models with classical flavor. Samuelson, in particular, transformed the economics profession by expressing both Keynesian macroeconomics and classical microeconomics in rigorous mathematical models. Although the Keynesian models have largely disappeared from advanced textbooks, they continue to occupy the central position in intermediate macroeconomic textbooks. The resilience is due to, first, the models' simplicity and, second, their continued relevance to the real world.

The Keynesian models, however, lost some key ideas of Keynes. The sticky-price assumption, for example, is not central to Keynes. The capital market, in contrast, is central in the determination of investment and hence, employment, but it has no role in the Keynesian models. After the discussion of Keynesian models, we introduce the theory of Keynes himself and Hyman Minsky, an influential post-Keynesian economist.

2 Business Cycles

The business cycle is the fluctuation of output around the long-term trend of output potential. Other macroeconomic variables, such as unemployment and inflation rate, also fluctuate with the output.

Figure 1 shows a stylized business cycle of output and the corresponding growth rate. A cycle may start from a *peak*, where the growth rate of real output turns negative. From the peak, the economy falls in a recession. Eventually, the economy reaches the lowest point during the cycle, which we call *trough*. The trough marks the end of the recession and the beginning of an expansion. The economy continues to expand and eventually reaches another peak, finishing a business cycle and beginning a new one.

The real-world business cycles are something of great complexity. Figure 2 shows the time-series plots of the real GDP growth, the unemployment rate, and the inflation of the US. Several features stand out. First, real-world business cycles are highly irregular. Both the lengths of recessions and expansions vary over time.

Second, expansions are often long, and recessions (the shaded area) are often short-lived. In the past seven decades, the longest recession was the Great Recession (2008Q1–2009Q2), which lasted six quarters. It was followed by the longest expansion (2009Q3–2019Q4), which lasted 42 quarters. And it was followed by the shortest recession in history, the Covid recession in 2020, which lasted two months.

Third, related to the second point, the unemployment rate rises rapidly during recessions and declines relatively slowly during expansions. During the Great Recession, for example, it took only 18 months for the unemployment rate to climb from 5.0% in April 2008 to 10.0% in October 2009. But it took 71 months for the unemployment rate to fall back to 5.0% in September 2015. As a result, the press often talks about the so-called "jobless recovery."

Fourth, recessions differ in the severity of output loss or the unemployment problem. During the recent recession due to the Covid-19 pandemic, the US real GDP declined 9.3% from peak to trough and the unemployment rate shot up to

Figure 1: Stylized business cycles.

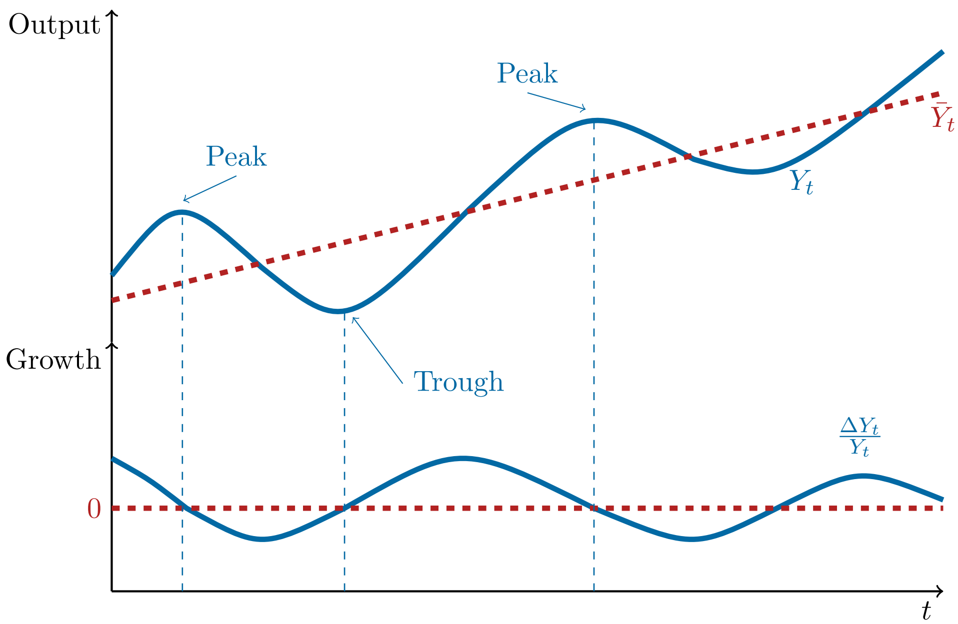
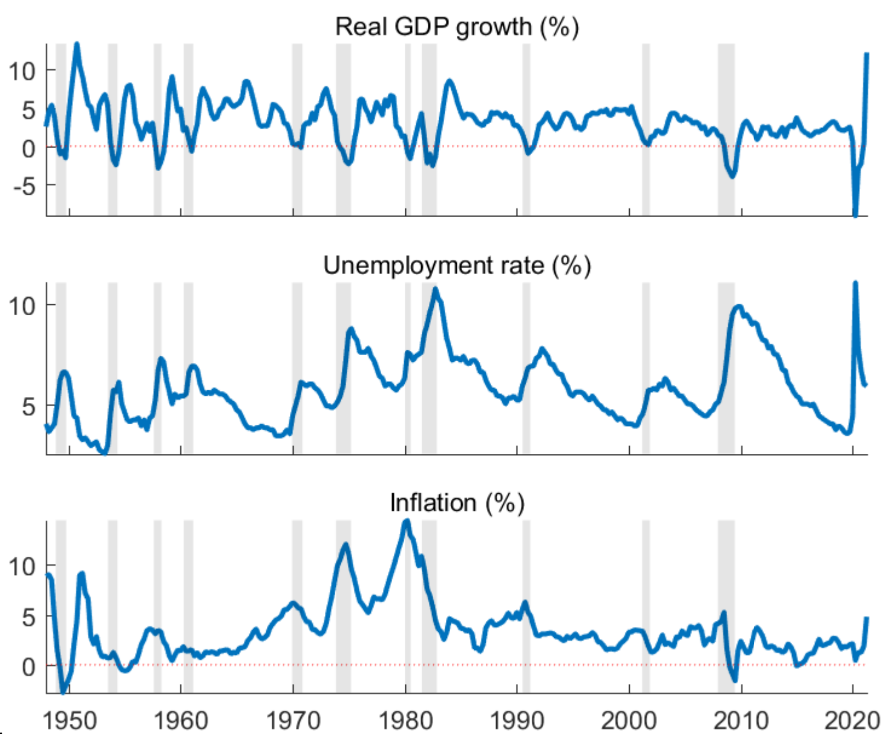


Figure 2: US business cycles.



†Data source: FRED.

14.8%. And some recessions are very mild. For example, the 2001 recession lasted only eight months, with an output loss of 0.3% and a maximum unemployment rate of 6.3%.

Fifth, the unemployment rate rarely *stays* at some level. During economic expansions, it declines almost continuously until reaching a tipping point, where it reverses its course and starts to rise rapidly. The tipping point often coincides with the onset of some crisis.

Sixth, the volatility of macroeconomic variables may also vary over time. For example, inflation was volatile in the 1970s and 1980s. Then the volatility of inflation became subdued until the outbreak of the Covid-19 pandemic. In particular, the period from the mid-1980s to 2007 was known as the Great Moderation, during which the volatility of economic fluctuations was much reduced in developed countries.

Finally, the unemployment rate is conversely related to real GDP growth. Okun's law often characterizes such a relationship between output and employment. One version of Okun's law is as follows:

$$Y_t - \bar{Y} = -\gamma(u_t - u^n), \tag{1}$$

where $Y_t - \bar{Y}$ is the output gap, u_t is the unemployment rate, u^n is the natural unemployment rate, and γ is a positive coefficient. If (1) holds, then we may deduce that the unemployment rate rises when the output declines.

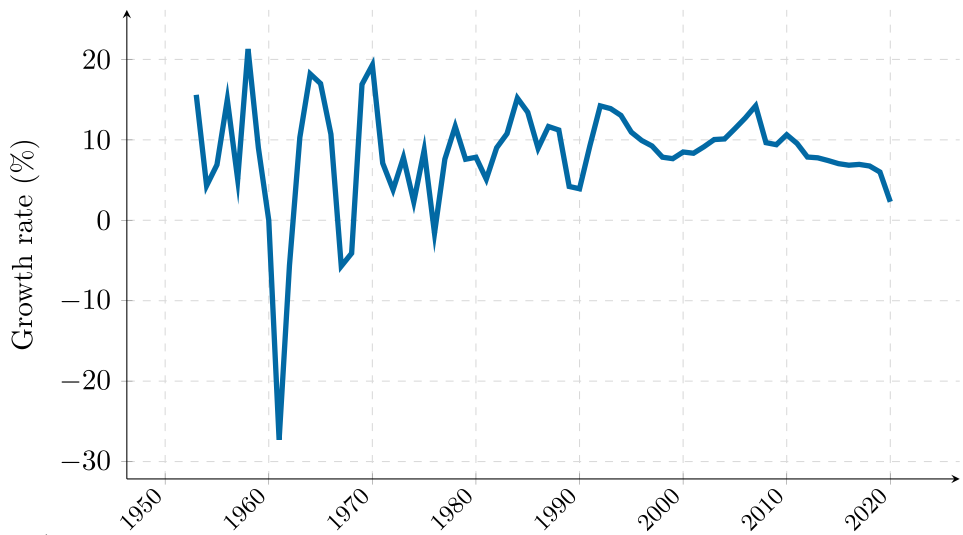
China's Economic Fluctuations

China's economic fluctuations are very different from those of developed countries. We often divide the post-war economic history of China into two periods, with a break in 1978, when the "Reform and Open-Up" began. We may call the first the pre-Reform era and the second the post-Reform era. We should understand, however, that the Reform was not a single event, but a process that roughly started from around 1978.

The output growth in the pre-Reform era was highly volatile (See Figure 3). The most salient fluctuation was during 1958–1962. The output increased by 21.3% when the Chinese government launched the Great Leap Forward campaign, the aim of which was to industrialize China in a big push. However, the campaign violated almost every principle of economics. For example, it violated both the division of labor and the economy of scale when the government mobilized farms to produce steel in small mills scattered in the countryside. The campaign soon turned into an economic disaster. The output growth ground to a halt in 1960 and declined 27.3% in 1961, when the campaign stopped. The decline in agricultural production, in particular, resulted in a devastating famine.

The "Reform and Open-Up" since 1978 has brought over 40 years of continued growth. The volatility of growth was much lower than in the pre-Reform era but much higher than in the developed countries. In the 1980s and early 1990s, the

Figure 3: Macroeconomic fluctuations in China.



†Data source: National Bureau of Statistics, China

economic fluctuations were mostly driven by the reform-and-retrenchment cycle. When the reformers ruled the day, they pushed for reform initiatives that relaxed the central control on investment and prices. But this led to inflation. The conservatives would then push back and used the device of central planning to cool the economy down. Once the economy was back in balance, the reformers would once again take initiatives, as the conservatives could not offer better alternatives. Thus the Reform and Open-Up continued in the fashion of "two steps forward, one step back."

Since the mid-1990s, the business cycles of China have become more and more synchronized with those in other economies, especially those of developing countries. The slowdown in the late 1990s and during 2008 and 2009, for example, coincided with the Asian Financial Crisis and the Global Financial Crisis, respectively. Note that the annual data may conceal short-term volatility. During the Global Financial Crisis, for example, China's real GDP growth slowed to 6.4% in the first quarter of 2009. Since then, however, the effects of the Four-Trillion-Stimulus program started to kick in and the economy rebounded forcefully, bringing the annual growth rate to 9.4%.

3 The Keynesian Cross

We first study a simple model called the *Keynesian Cross*. We use the model to understand that the economy may not produce the potential output in the short run. As a result, the economy may experience recessions and expansions.

We assume that the total expenditure of a closed economy is the sum of consumption expenditure, investment expenditure, and government expenditure:

$$TE = C + I + G,$$

where C is the consumption expenditure, I is the investment expenditure, and G is the government purchase. We further assume that the consumption expenditure is a function of disposable income:

$$C = C(Y - T).$$

where Y represents total income and T represents tax. We further assume that

$$0 < C'(\cdot) < 1 \quad \text{and} \quad C''(\cdot) < 0.$$

That is, the marginal propensity to consume (MPC) is positive, and MPC declines as the disposable income increases.

Every expenditure generates an equal amount of income on the other side of the transaction. As a result, the total expenditure has to be equal to the total income (Y):

$$Y = C(Y - T) + I + G. \tag{2}$$

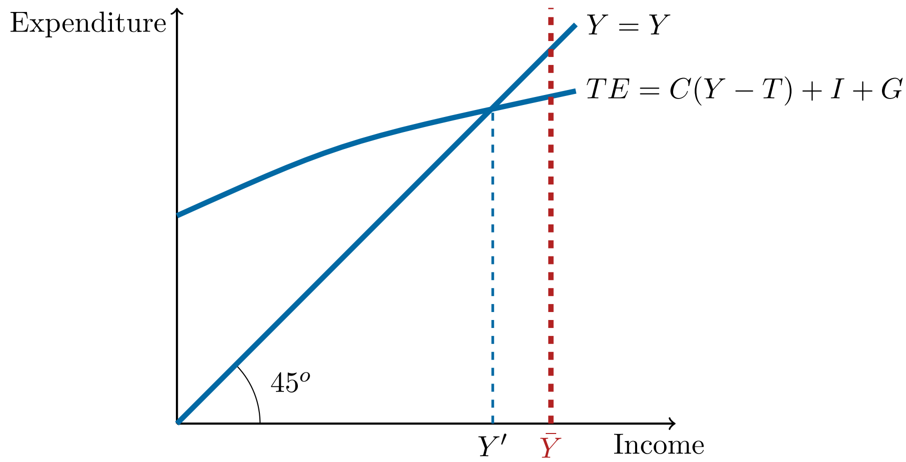
In this model, the only endogenous variable is Y, the total income or expenditure. Other variables, including I, are assumed to be exogenous. Simple as it is, the model gives us many insights about the economy as a whole. First, total expenditure determines total income. It is not intuitive at first sight since we, in everyday life, determine expenditure according to our income. But for the economy as a whole, it is almost a tautology to say that total expenditure determines total income since they are equal by definition.

As a result, if a substantial number of people choose to reduce consumption and increase savings, then the reduced consumption expenditure would be accompanied by a reduction of total income. If thrifty people save more by consuming less, they can maintain their income and increase their wealth. But if everyone saves more, the total income would decline, leading to reduced national wealth (via a financial crisis often associated with economic recessions). This apparent paradox is known as the *paradox of thrift*, a well-known example of the *fallacy of composition*.

Second, the equilibrium output (Y') may be below the output potential $\left(\bar{Y}\right)$, as shown in Figure 4. Note that in the Keynesian Cross model in (2), the only endogenous variable is Y, the equilibrium income or expenditure. To graphically represent the solution to (2), we first draw the right-hand side of (2) as a curve, $Y = C(Y - T) + I + G$. Then we draw the left-hand side of (2) as a 45-degree line ($Y = Y$). The solution of (2), the equilibirum total income or expenditure Y', is at the cross of the two lines.

Following Keynes, we may call Y' the *effective demand*. There is no reason, in a modern monetary economy, to believe that the aggregate demand always meets

Figure 4: The Keynesian Cross.



the aggregate supply at the potential output level $\left(\bar{Y}\right)$. Say's law, which states that "supply creates demand," is valid only in a barter economy. In an economy with money, sellers may choose to hold money (to save) instead of spending all their income generated from transactions. If a substantial portion of the population chooses to save more money, then the total expenditure would be lower than the economy can potentially produce.

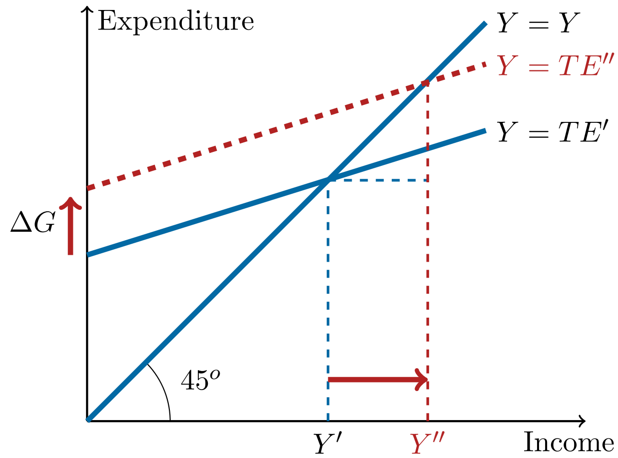
Third, to accommodate a higher income, the sum of investment and government expenditure must rise more than consumption, due to the assumption that MPC declines as the income increases. Suppose that the firms expand their production capacity in anticipation of future profit. If the increased output can be absorbed by increased aggregate demand, then total income would rise. However, since MPC declines as income increases, a larger share of the increased aggregate demand must be taken care of by investment or government expenditure.

Finally, an increase in investment or government expenditure would lead to a larger increase in total income. That is, investment or government expenditure has a *multiplier effect*. Intuitively, an increase in government expenditure generates more income, which leads to more consumption expenditure, which again generates more income, and so on.

To see this point more clearly, we may assume that the consumption function is linear, e.g., $C(Y - T) = C_0 + C_1(Y - T)$ with $C_1 < 1$, then the right-hand side of (2) is a straight line with a slope less than one (Figure 5). Now, if the government increases expenditure by ΔG, then the line TE shifts upward by ΔG, moving equilibrium income from Y' to Y''. The increase in total income is clearly larger than the initial increase in government expenditure.

Mathematically, (2) defines an implicit function $Y(G, T, I)$. Note that investment I is an exogenous variable in this model just like G and T. Using the implicit

Figure 5: The multiplier effect.



function theorem, we obtain

$$\frac{\partial Y}{\partial G} = \frac{1}{1 - C'(Y - T)} = \frac{1}{1 - \text{MPC}}.$$

Note that $C'(Y-T)$ is MPC, marginal propensity to consume. Since $0 < \text{MPC} < 1$ is less than one, $\partial Y/\partial G$ must be greater than one. It is for this reason that we call $\partial Y/\partial G = 1/(1 - \text{MPC})$ the *government expenditure multiplier*. Similarly, we have

$$\frac{\partial Y}{\partial T} = -\frac{C'(Y - T)}{1 - C'(Y - T)} = -\frac{\text{MPC}}{1 - \text{MPC}},$$

which we call the *tax multiplier*. In contrast to the government expenditure multiplier, the magnitude of the tax multiplier is not necessarily greater than one. If $\text{MPC} < 0.5$, indeed, then $|\partial Y/\partial T| < 1$.

An Example: Government Expenditure and Tax Multipliers

If we assume that the consumption function is linear, say,

$$C(Y - T) = 100 + 0.8(Y - T),$$

then $\text{MPC} = 0.8$, and

$$\frac{\partial Y}{\partial G} = \frac{1}{1 - \text{MPC}} = 5, \quad \frac{\partial Y}{\partial T} = -\frac{\text{MPC}}{1 - \text{MPC}} = -4.$$

The magnitude of the expenditure multiplier is always bigger than the tax multiplier. We can understand this point by noting that tax cuts affect disposable income, which in turn induces individuals to consume, raising aggregate demand. That is to say, the immediate effect of tax cuts on aggregate demand is moderated by MPC, which is always smaller than 1. Government spending (e.g., on public works), however, directly raises aggregate demand.

The multiplier effect is often used to justify fiscal stimulus. A multiplier greater than 1, in particular, is sometimes considered a necessary condition for government expenditures on public works. This is not necessarily the case, at least when the economy is in a deep recession and many resources (particularly the labor force) are idle. Government spending on public works would employ idle resources to build something useful, which is better than doing nothing. In other words, if opportunity costs of building public works are low, then building public works always raises aggregate welfare. A big multiplier effect is thus a plus, not a necessary condition.

We can easily extend the preceding analysis to an open economy, where part of the expenditure is on imported goods. In this case, an increase in G (even when all of it is spent on domestic goods) would bring less increase in total national income since not all income generated from transactions (i.e., G) goes to domestic households. This effect is often called *leakage* of fiscal stimulus. The leakage effect may lead to coordination failure on the world stage to combat global recessions: Every country is reluctant to implement stimulus packages that benefit other countries as well.

Note that the multiplier effect also works in the reverse direction. If the government cut spending, the reduction of national income would be more than the spending cut. If the government increases taxes, then it would lead to a reduction of national income that is proportional to the tax multiplier. If the investment or consumption expenditure declines, the immediate reduction of income would also start a chain reaction that results in a much bigger income reduction than the immediate one.

4 The IS-LM Model

In the Keynesian Cross model, investment is considered exogenous, and there is no role for the interest rate. As a result, there is no "crowding-out" effect predicted by classical models. More realistically, however, investment should be endogenous and it may depend on, among other factors, financing cost measured by interest rate. We now present the IS-LM model that endogenizes interest rate and investment. The IS-LM model was proposed by John Hicks in 1937, and it became the leading interpretation of Keynes's theory by Keynesian economists.

In this section, we first discuss the sticky-price assumption, then study the plain vanilla IS-LM model, which is for a closed economy. The next section extends the IS-LM model to analyze open economies.

4.1 Sticky Price

Before introducing the IS-LM model, we first discuss the sticky-price (or nominal rigidity) assumption, which states that prices (including wages) are sticky or rigid in the short term. It is a crucial assumption for the IS-LM models in this chapter.

First, the sticky-price assumption implies that the general price level can be treated as an *exogenous* variable, simplifying analysis. Once we make the assumption, however, it would be awkward to talk about inflation, another important macroeconomic variable. Inflation will always be zero if the sticky-price assumption holds. We will relax this assumption later, when we analyze inflation.

Second, under the sticky-price assumption, markets would not automatically clear. When a negative shock occurs to the aggregate demand, prices would fail to fully adjust, with a consequence that the supply side cannot operate at its full potential. Particularly, the downward wage rigidity, which has wide empirical support, would cause excess supply of labor and involuntary unemployment during recessions.

Third, under the sticky-price assumption, the classical dichotomy would fail and money would no longer be neutral. When money supply increases, for example, prices and wages would not rise proportionally. As a result, a change in money supply (monetary policy) would have some real effects.

4.2 The Model

IS stands for *investment* and *saving*. LM stands for *liquidity* and *money*. The IS-LM model consists of two equations corresponding to two equilibrium conditions: investment equals saving, and liquidity demand equals money supply.

4.2.1 IS Equation

The IS equation may be viewed as a generalization of the Keynesian Cross model in (2). Instead of treating the investment as exogenous, we assume that investment is a function of the *endogenous* real interest rate:

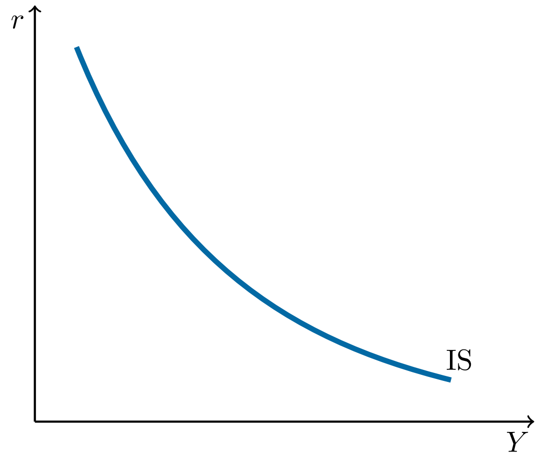
$$I = I(r).$$

And we assume that $I(r)$ is differentiable and $I'(r) < 0$. That is, a higher interest rate lowers the level of investment in the economy. The IS equation is then

$$Y = C(Y - T) + I(r) + G. \tag{3}$$

We may interpret the IS equation as an equilibrium in the market for goods and services. The right-hand side gives the total demand, and the left-hand side gives

Figure 6: The IS curve.



the total supply. Alternatively, we rewrite (3) as

$$I(r) = S(Y, T, G), \tag{4}$$

where $S(Y, T, G) = Y - C(Y - T) - G$ is the national savings. We may thus interpret the IS equation as an equilibrium in the financial market for loanable funds. We may read (4) as *investment equals savings*, hence the name IS.

Unlike the Keynesian Cross model, which has one equation for one endogenous variable, the IS equation contains two endogenous variables, Y and r. Hence the IS equation is not a standalone model. Any combination of Y and r satisfying (3) may be a possible equilibrium. All these combinations form the IS curve on the two-dimensional coordinate diagram with Y on the horizontal axis and r on the vertical axis.

IS Curve

In the IS equation, there are two endogenous variables, Y and r, and two exogenous variables, T and G. The IS equation defines an implicit function $Y(r, T, G)$. The implicit function characterizes how the total income changes with the real interest rate, given the fiscal condition expressed in T and G.

Or, equivalently, the IS equation defines an implicit function $r(Y, T, G)$, which is more suitable for graphs. We may represent the implicit function $r(Y, T, G)$ on the two-dimensional coordinate diagram in Figure 6, giving us the IS curve. Note that the location of the IS curve is determined by the exogenous variables. When exogenous variables change, the IS curve *shifts* on the diagram.

To see why the IS curve is downward sloping, note that a decline in r results in higher investment in (3). Consequently, as in the Keynesian Cross model, the total income or output Y would be higher as well.

Mathematically, the shape of the IS curve is determined by the slope, which is defined as

$$\frac{\partial r}{\partial Y} \equiv \lim_{\Delta Y \to 0} \frac{r(Y + \Delta Y, T, G) - r(Y, T, G)}{\Delta Y}.$$

This is a measure of how r changes with a unit change in Y, keeping T and G fixed. We can apply the implicit function theorem to obtain:

$$\frac{\partial r}{\partial Y} = -\frac{1 - C'(Y - T)}{-I'(r)} = \frac{1 - C'}{I'},$$

where C' and I' are shorthand for $C'(Y - T)$ and $I'(r)$, respectively. C' is nothing but MPC and I' measures the sensitivity of investment to the real interest rate. Since $0 < C' < 1$ and $I' < 0$, we must have a negative slope, $\partial r / \partial Y < 0$.

"Move Along" or "Shift"

Fixing exogenous variables (T and G), we have a fixed IS curve on the two-dimensional diagram. When there is a change in Y, r would also change. When this happens, we say that the equilibrium point *moves along* the IS curve.

Changing any of the exogenous variables (T or G), however, *shifts* the IS curve. Suppose there are N equilibrium points on the IS curve, $(r_i, Y_i), i = 1, \ldots, N$. Increasing the government expenditure by ΔG, for example, increases every Y_i by $\Delta G/(1 - MPC)$, keeping r_i fixed. That is, the new equilibrium points are $(r_i, Y_i + \Delta G/(1 - MPC))$. That is, the IS curve shifts to the right by $\Delta G/(1 - MPC)$.

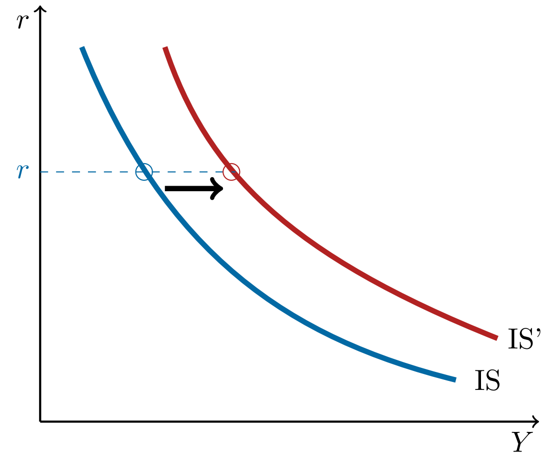
In this book, we most often analyze horizontal shifts: Fixing the endogenous variable on the vertical axis (r in the IS-LM model), we check whether an exogenous change makes the other endogenous variable on the horizontal axis (Y) bigger or smaller. If the exogenous change makes Y bigger (smaller), then we say the exogenous change shifts the curve to the right (left).

Equivalently, we may also analyze vertical shifts: Fixing Y, we check whether the exogenous change makes r higher or lower. If it is higher (lower), then the curve shifts up (down).

If a curve is vertical, then we can only analyze horizontal shifts. If a curve is horizontal, then we can only analyze vertical shifts.

For example, if the government increases spending by ΔG, then the IS curve would shift to the right. To see this, let's fix any real interest rate r, and analyze how the spending increase would affect Y. Fixing the interest rate means fixing the

Figure 7: The effect of a fiscal stimulus on the IS curve.



investment expenditure $I(r)$. As a result, we can apply the analysis of the Keynesian Cross, which tells us that the fiscal stimulus would increase Y by $\Delta G/(1 - MPC)$. Since this is true for every r, the increased government spending would shift the IS curve to the right (Figure 7).

Mathematically, note that (3) defines an implicit function $Y(r, T, G)$. To calculate the effect of a unit increase in G on Y, keeping r fixed, we calculate the partial derivative:

$$\frac{\partial Y(r, T, G)}{\partial G} \equiv \lim_{\Delta G \to 0} \frac{Y(r, T, G + \Delta G) - Y(r, T, G)}{\Delta G}.$$

Applying the implicit function theorem, we obtain:

$$\frac{\partial Y}{\partial G} = \frac{1}{1 - C'(Y - T)} = \frac{1}{1 - \text{MPC}},$$

which is exactly the government expenditure multiplier in the Keynesian Cross model. Note that the MPC in the preceding equation is a function of $(Y - T)$, the disposable income. If ΔG is small, then a ΔG-increase in government expenditure would increase Y by $\Delta G/(1 - \text{MPC})$, approximately. If, furthermore, the consumption function is linear, then the approximation is exact. That is, a ΔG-increase in government expenditure would shift the IS curve to the right by exactly $\Delta G/(1 - \text{MPC})$.

Similarly, we can calculate how a tax cut would shift the IS curve. We apply the implicit function theorem and obtain:

$$\frac{\partial Y}{\partial T} = -\frac{C'(Y - T)}{1 - C'(Y - T)} = -\frac{\text{MPC}}{1 - \text{MPC}},$$

which is exactly the tax multiplier. A tax cut by the amount $\Delta T < 0$ would shift the IS curve to the right by approximately $-\text{MPC}/(1 - \text{MPC})\Delta T$.

4.2.2 LM Equation

The LM equation characterizes money market equilibrium. We introduce an exogenous money supply M, which is assumed to be controlled by the central bank. And we introduce a function, $L(r, Y)$, to characterize the *real* demand for money (or liquidity). Denote the general price level by P, and the nominal demand for money is then $PL(r, Y)$.

We assume that a higher interest rate depresses money demand, since money (at least in the traditional sense) has zero interest rate. We also assume that a higher income boosts money demand, since higher income naturally leads to higher demand for transactions. In mathematical language, we assume that $L(r, Y)$ is decreasing in r and increasing in Y. If $L(\cdot, \cdot)$ is differentiable, then $L_1 \equiv \partial L/\partial r < 0$, $L_2 \equiv \partial L/\partial Y > 0$.

In money market equilibrium, we have $M = PL(r, Y)$, which says that money supply equals money demand. More conventionally, we write:

$$\frac{M}{P} = L(r, Y). \tag{5}$$

We may interpret the left-hand side of (5) as the *real* money supply. We may thus read (5) as *real money supply equals liquidity demand*, hence the name LM equation.

The LM equation contains two endogenous variables (Y and r) and two exogenous variables (M and P). The exogeneity of the general price level is justified by the sticky-price assumption.

Like the IS equation, the LM equation is not a standalone model. Any combination of Y and r satisfying (5) may be a possible equilibrium for the money market. All these combinations form the LM curve on the two-dimensional coordinate diagram with Y on the horizontal axis and r on the vertical axis.

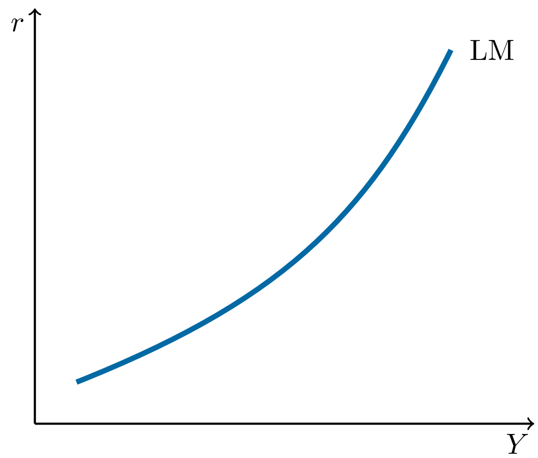
LM Curve

Like the IS equation, the LM equation also defines an implicit function $r(Y, M, P)$, giving us the LM curve. It characterizes how the real interest rate changes with the income, given the general price level and the money supply. We now analyze the shape of the curve and how the curve shifts when M/P changes.

First, the LM equation dictates that, given any M and P, a decline in r must accompany a decline in Y (Figure 8). Hence the LM curve is an upward-sloping curve. To understand this intuitively, note that a decline in income would reduce the demand for money. But the money supply is fixed. Hence the interest rate must decline to maintain money market equilibrium.

More precisely, we may apply the implicit function theorem and obtain the slope

Figure 8: The LM curve.



of the LM curve:

$$\frac{\partial r}{\partial Y} = -\frac{L_2}{L_1} > 0. \tag{6}$$

There are two interesting special cases: If $L_1 = \infty$, then the LM curve would be horizontal. On the other hand, if $L_1 = 0$, then the LM curve would be vertical.

Second, if the monetary authority increases M, then the LM curve would shift to the right. To see this, suppose that we fix any r and analyze how an increase in M would affect Y. Under the sticky-price assumption, the price level P does not change. Then an increase in M increases the *real money supply*, the left-hand side of (5). As a result, Y on the right-hand side must rise to make the equation hold.

More precisely, note that the LM equation in (5) also defines an implicit function $Y(r, M, P)$. We apply the implicit function theorem and obtain

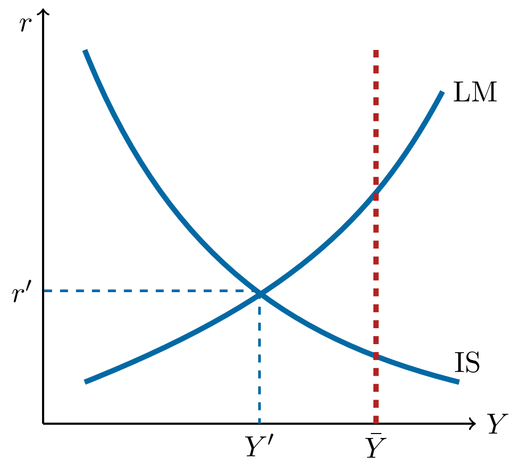
$$\frac{\partial Y(r, M, P)}{\partial M} \equiv \lim_{\Delta M \to 0} \frac{Y(r, M + \Delta M, P) - Y(r, M, P)}{\Delta M} = \frac{1}{PL_2} > 0.$$

That is, if the monetary authority increases the money supply by ΔM, then the LM curve would shift to the right by $\Delta M/(PL_2)$, approximately. If the money demand function $L(r, Y)$ is linear in Y, then the approximation would be perfect.

Note that in most cases, it suffices to study horizontal shift of curves as in preceding equation. When an upward-sloping curve (e.g., the LM curve) shifts to the right, it is equivalent to shifting upward. However, there are cases where it is necessary to study how a curve would shift vertically (e.g., if the curve is horizontal). If this is the case, we apply the implicit function theorem to the implicit function $r(Y, M, P)$ defined in (5) and obtain:

$$\frac{\partial r(Y, M, P)}{\partial M} \equiv \lim_{\Delta M \to 0} \frac{r(Y, M + \Delta M, P) - r(Y, M, P)}{\Delta M} = \frac{1}{PL_1} < 0. \tag{7}$$

Figure 9: The IS-LM model.



That is, if the monetary authority increases M, then the LM curve would shift downwards. Examining (6) and (7), we may find that $L_1 = \infty$ is an interesting special case: The LM curve is *horizontal*, and monetary expansion is unable to shift the LM curve downwards. We call this special case a *liquidity trap*, a situation where increases in the money supply fail to lower interest rates.

4.2.3 The IS-LM Model

The IS-LM model is composed of (3) and (5), both of which we rewrite here:

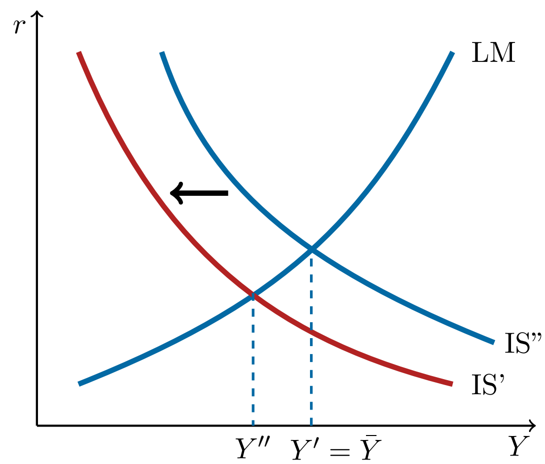
$$\begin{aligned} Y &= C(Y-T) + I(r) + G \\ \frac{M}{P} &= L(r, Y). \end{aligned}$$

The equilibrium of the economy is the solution to the preceding two equations, i.e., the point at which the IS curve and the LM curve cross (Figure 9). We may understand the point (Y', r') as the joint equilibrium in the financial market for the loanable funds and the money market. Once again, the equilibrium output Y' may or may not coincide with the output potential $\bar{Y}$.

4.3 The Demand-Shock Recession

Suppose that the economy is initially at the output potential ($\bar{Y}$ in Figure 10). If there is a negative shock to the IS curve, then the IS curve shifts to the left. As a result, the equilibrium output declines to $Y'' < \bar{Y}$. At the same time, the real interest rate falls, and, by Okun's law, the unemployment rate rises above the natural unemployment rate.

Figure 10: Demand-shock recession.



For example, a sudden decline in housing prices may depress household consumption, hence shifting the IS curve to the left. For another example, the crash of an over-leveraged stock market may dent investment sentiment, also sending the IS curve to the left.

4.4 Policy Analysis with IS-LM

When the economy falls into a recession, the modern government is invariably under pressure to do something. The government may increase expenditures on transportation or energy infrastructure, subsidize durable-goods consumption, and so on. The government may also cut tax for businesses and households, hoping to stimulate investment and consumption. Finally, the monetary authority may inject more liquidity into the financial market, pushing down interest rates, elevating asset prices, and stimulating investment.

We may apply the IS-LM model to analyze the effect of monetary and fiscal policy on the economy. Figure 11 illustrates the effect of expansionary fiscal policies. If there is a fiscal stimulus, in the form of either a spending increase or a tax cut, then the IS shifts to the right, bringing higher output and employment.

At the same time, however, the real interest rate rises, and the investment expenditure declines. We call this side-effect of fiscal stimulus the *crowding-out* effect. Unlike in the classical model, however, the crowding out here is *partial.* After the fiscal stimulus, the total output will rise, but by an amount less than in the Keynesian Cross model. Recall that the Keynesian Cross model takes investment as exogenous, independent of the interest rate, hence there is no crowd-out effect.

If there is a monetary stimulus, then the LM curve shifts to the right, bringing a lower interest rate and a higher output (Figure 12). The IS-LM model, indeed,

Figure 11: Policy analysis: fiscal stimulus.

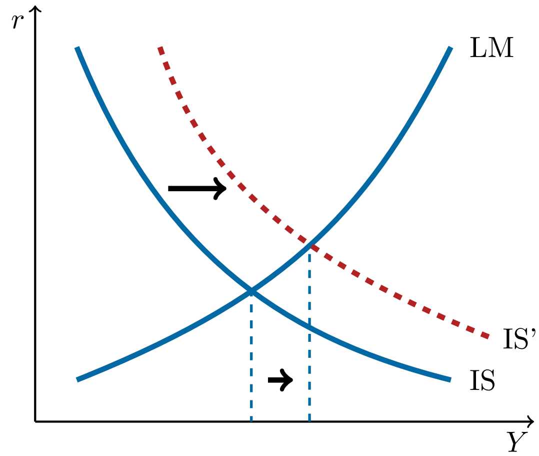
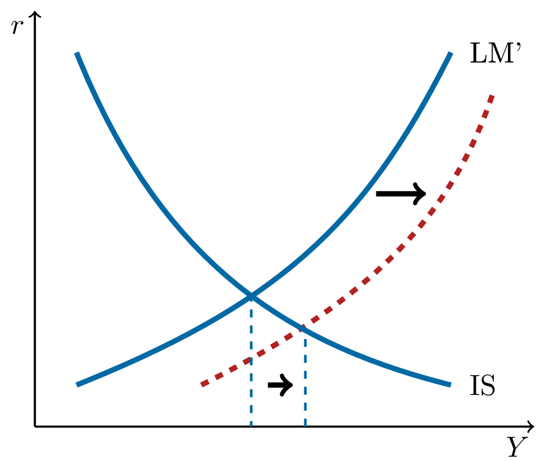


Figure 12: Policy analysis: monetary stimulus.



characterizes an interest-rate channel of monetary policy transmission:

$$M \uparrow \ \Rightarrow r \downarrow \ \Rightarrow I(r) \uparrow \ \Rightarrow \quad Y \uparrow .$$

Note that the direct effect of monetary stimulus (e.g., a massive liquidity injection) is to lower the nominal interest rate i. But under the sticky-price assumption, a lower nominal interest rate translates into a lower real interest rate, which businesses and households really care about. Note that there are other channels for monetary policies to affect the economy, which we will discuss later.

4.4.1 Mathematical Treatment

Mathematically, note that the IS-LM equations define a system of implicit functions:

$$\begin{aligned} Y &= Y(T, G, M, P), \\ r &= r(T, G, M, P). \end{aligned}$$

Note that these two implicit functions are the *solution* to the IS-LM equations, representing each endogenous variable using a function of exogenous variables. Given the values of the exogenous variables, the solution determines the equilibrium output and the real interest rate.

To analyze the effect of government policies on the output, say the effect of increasing government expenditure on the output ($\partial Y/\partial G$), we cannot use the implicit function theorem as usual, due to the fact that the implicit function $Y(T, G, M, P)$ is defined by a system of equations. There is a general version of the implicit function theorem for the system of equations, but it is difficult to remember and cumbersome to apply. Instead, we can take total differentiation of the IS-LM equations and obtain

$$\begin{aligned} (1 - C')dY - I'dr &= dG - C'dT \\ PL_2 dY + PL_1 dr &= dM, \end{aligned}$$

where $I' = I'(r)$, $L_1 = \partial L(r, Y)/\partial r$, $L_2 = \partial L(r, Y)/\partial Y$, and $C' = C'(Y - T)$ is MPC. The total differentiation transforms the system of nonlinear equations into a linear system of equations. Now the endogenous variables become dY and dr, which we may regard as changes in Y and r, respectively. And the exogenous variables are dG, dT, and dM. Under the sticky-price assumption, we have $dP = 0$.

We may write the linear system of equations in matrix form:

$$\begin{pmatrix} 1 - C' & -I' \\ PL_2 & PL_1 \end{pmatrix} \begin{pmatrix} dY \\ dr \end{pmatrix} = \begin{pmatrix} dG - C'dT \\ dM \end{pmatrix}.$$

Note that changes in the endogenous variables are on the left-hand side and changes in the exogenous variables are on the right-hand side. The preceding linear system of equations is all we need for policy analysis.

To analyze the effect of government expenditure (G), we hold both T and M fixed (that is, $dT = 0$, $dM = 0$). Using Cramer's rule, we obtain

$$\begin{aligned} \frac{dY}{dG} &= \frac{L_1}{L_1(1 - C') + I'L_2} > 0, \\ \frac{dr}{dG} &= -\frac{L_2}{L_1(1 - C') + I'L_2} > 0. \end{aligned}$$

To analyze the effect of tax (T), we hold both G and M fixed (that is, $dG = 0$, $dM = 0$). Using Cramer's rule, we obtain

$$\begin{aligned}\frac{dY}{dT} &= -\frac{L_1 C'}{L_1(1-C') + I'L_2} < 0, \\ \frac{dr}{dT} &= \frac{L_2 C'}{L_1(1-C') + I'L_2} < 0.\end{aligned}$$

To analyze the effect of monetary policy, we hold G and T fixed (that is, $dG = 0$, $dT = 0$). Using Cramer's rule, we obtain

$$\begin{aligned}\frac{dY}{dM} &= \frac{I'}{P\left(L_1(1-C') + I'L_2\right)} > 0, \\ \frac{dr}{dM} &= \frac{1-C'}{P\left(L_1(1-C') + I'L_2\right)} < 0.\end{aligned}$$

We can see that fiscal stimulus (increasing G or cutting T) would generally lead to a higher output (income) and a higher interest rate. And an expansionary monetary policy, such as quantitative easying (QE), would generally result in a lower interest rate and higher output (income). When the economy is in a recession, both policies may be helpful. However, there are two interesting special cases, where only one of them would work.

4.4.2 When Monetary Policy Doesn't Work

First, when $L_1 \equiv \partial L(r, Y)/\partial r = \infty$ (liquidity trap), then $dY/dM = 0$. That is, monetary policy does not affect output. But fiscal policies work:

$$\frac{dY}{dG} = \frac{1}{1-C'}, \quad \frac{dY}{dT} = -\frac{C'}{1-C'}.$$

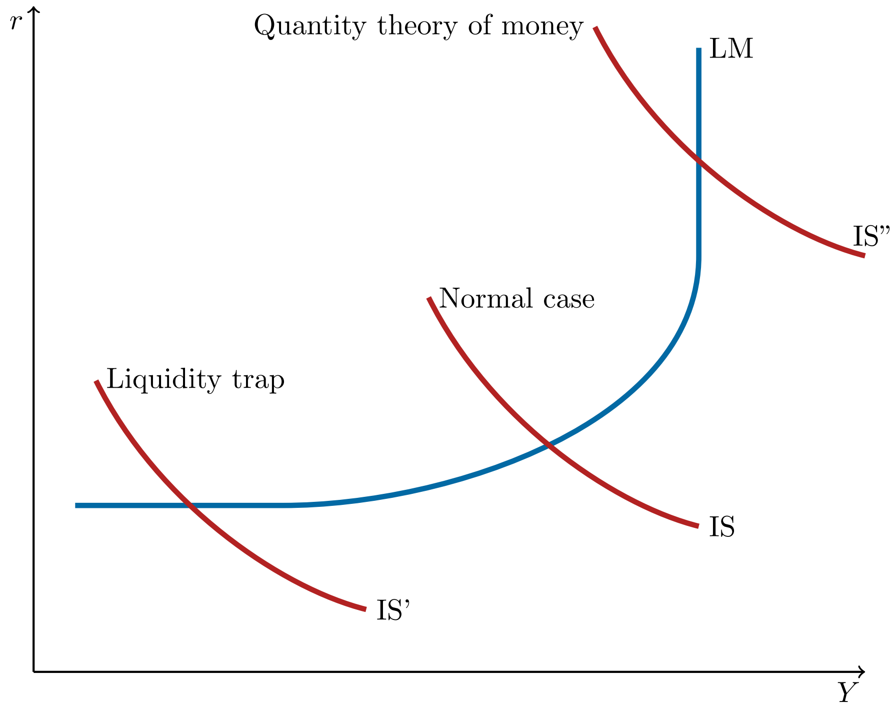
In this case, the government expenditure multiplier and the tax multiplier are exactly the same as in the Keynesian Cross model. To understand this, recall that the assumption $L(r, Y)/\partial r = \infty$ implies that the LM curve is horizontal and that monetary expansion is unable to shift the LM curve downwards. In this case, r becomes a constant and the IS-LM model reduces to the Keynesian Cross model.

4.4.3 When Fiscal Policies Don't Work

Second, when $L_1 \equiv L(r, Y)/\partial r = 0$, then $dY/dG = dY/dT = 0$. That is, fiscal policies do not affect output. But the monetary policy works in this case:

$$\frac{dY}{dM} = \frac{1}{PL_2} > 0.$$

Figure 13: IS-LM in three scenarios.



To understand this, note that the assumption $L(r, Y)/\partial r = 0$ makes the LM equation equivalent to the classical quantity theory of money, $(M/P = kY)$, where k is constant. As a result, the LM equation alone determines the total output Y. A monetary expansion directly leads to an expansion of Y.

Graphically, the assumption $L(r, Y)/\partial r = 0$ makes the LM curve vertical. In this case, the shifting of the IS curve does not affect the output. The fiscal policies do not work, and only monetary policy matters. The Monetarists (e.g., Milton Friedman) took such a position during the stagflation in the 1970s.

Figure 13 illustrates the application of the IS-LM models in three scenarios: the liquidity trap, the quantity theory of money, and the normal case. From this graph, we may say that the so-called Monetarist revolution did not refute Keynesian economics, but applied Keynesian economics to a particular situation.

4.5 Interest Rate Policy

Central banks in the real world have a dominant influence on short-term interest rates without targeting any level of money supply. A typical practice is to define a narrow corridor for the overnight interest rate at which banks lend to each other. For example, the US Federal Reserve sets a small range for the federal funds rate (FFR). As of October 31, 2020, the target range for FFR is 0.00–0.25%. For another example, the European Central Bank sets a rate on its deposit facility, which acts as the floor of the interest rate corridor, and a rate on its marginal lending facility, which acts as the ceiling. In addition, ECB also sets a rate on its main refinancing operations (MRO), in which banks can borrow liquidity from the Eurosystem against collateral on a weekly basis. As of October 31, 2020, the rate on the ECB deposit facility is -0.5%, the rate on the marginal lending facility is 0.25%, and the rate on MRO is 0%.

For the simplicity of analysis, we may assume that the central bank sets the interest rate. Money supply changes when the interest rate changes. To implement an interest rate cut, for example, the central bank may inject an adequate amount of liqudity, increasing the base money. The interest rate cut may also stimulate bank lending, thus expanding the broad money. In other words, the policy rate may be regarded as exogenous, while the money supply may be regarded as endogenous.

To model interest rate policy, we may modify the IS-LM as follows:

$$\begin{cases} Y &= C(Y-T)+I(r)+G \\ r &= \bar{r} \\ \frac{M}{P} &= L(r,Y), \end{cases}$$

where $\bar{r}$ is the exogenous policy rate. In this model, the money supply M is endogenous. Given the policy rate $\bar{r}$, the IS equation solely determines the total output or income. The money supply then automatically adjusts to meet the demand for money, which is determined by the policy rate and the equilibrium income. To see how a cut in the policy rate may expand the economy, see the upper panel of Figure 14. And the rate cut must be accompanied by an increase in money supply, as shown in the lower panel of Figure 14.

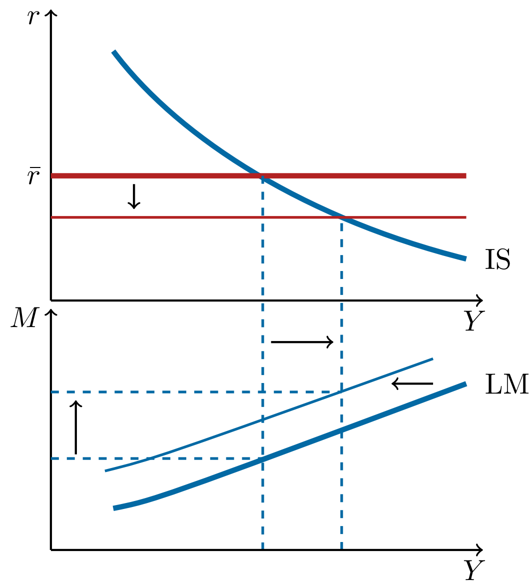
5 Open-Economy IS-LM Models

We now extend the IS-LM model to analyze open economies. We first discuss different types of exchange rate regimes, which make a difference in model analysis. Then we consider models for *small open economies*, whose excess saving or saving deficit does not affect the world interest rate. And we also consider a model for a *large open economy*. In the following, we assume free international capital flows, which implies a common real interest rate in the world. The common real interest rate is exogenous to small open-economy models but endogenous to the large open-economy model.

5.1 Exchange Rate Regime

An exchange rate regime is how a monetary authority manages its currency related to other currencies and the foreign exchange market. There are two major types: floating (or flexible) exchange rate and fixed (or pegged) exchange rate.

Figure 14: An interest-rate cut.



In a floating exchange rate regime, the exchange rate is allowed to fluctuate freely in the foreign exchange market. The monetary authority may or may not intervene with the foreign exchange market. If the monetary authority does not intervene with the foreign exchange market, we call it a *clean float*. Otherwise, we may call it *managed float* or *dirty float*.

In a fixed exchange rate regime, the currency of a small economy is pegging to a reserve currency (e.g., USD, Euro) at a particular exchange rate. And there are different forms of pegging in history, including the gold standard, monetary union (e.g., Euro area), currency boards, and so on. In a typical fixed exchange rate regime, the monetary authority stands ready to buy or sell the foreign currency at a pre-determined exchange rate. If the domestic currency tends to appreciate, the monetary authority has to "sell" domestic currency and buy the foreign currency in the foreign exchange market. For the monetary authority, "selling" domestic currency is injecting money into the economy, expanding the money supply. If the domestic currency faces depreciation pressure, the monetary authority has to do the opposite, "buy" domestic currency and sell the foreign currency. To buy domestic currency is to withdraw liquidity from the market and decrease the money supply.

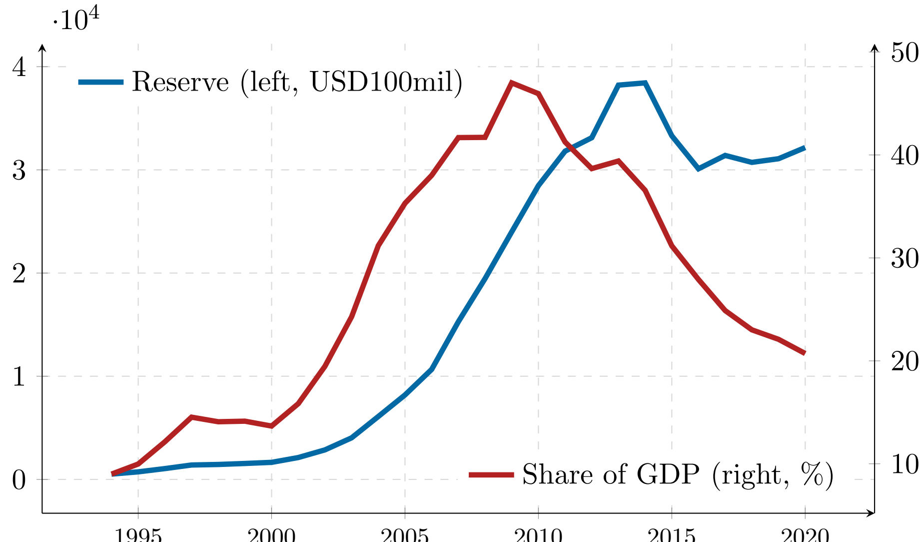
The advantage of a floating exchange rate is that the country can make monetary policies independently from other central banks in the world. As we can see in the preceding discussion, if the exchange rate is pegged to a reserve currency, the money supply has to change with the occurrences of appreciation/depreciation pressure. The monetary authority is thus unable to pursue objectives other than the defense of the fixed exchange rate. The disadvantage of the floating exchange rate is that the economy has to live with an exchange rate that is often excessively volatile.

The advantage of having a fixed exchange rate (often with a major trading partner) is mainly the stability of the exchange rate, which facilitates international trade and investment. For countries with a poor record of responsible monetary policy, fixing the exchange rate with a major foreign currency is a way of "importing" responsible monetary policy from abroad. Defending the exchange rate can often rally more political support than conducting responsible monetary policy.

The disadvantage of the fixed exchange rate is, first, the loss of the independence of monetary policy. Second, to defend the fixed exchange rate against speculative attacks, the country has to accumulate a substantial amount of *foreign exchange reserve*. The foreign exchange reserve must be liquid and safe, which necessarily implies a low return to reserve assets. If a country has a higher return on domestic investment, then maintaining a massive low-return foreign exchange reserve is costly to the welfare of the country.

From 1995 to 2005, China pegged RMB to the US dollar. To maintain the peg, and especially to keep RMB stable even after the "7.21 Reform" of 2005, China accumulated a massive foreign reserve. In absolute terms, China's foreign reserve reached an all-time high of USD 3.84 trillion in 2014, which amounts to 36.5% of GDP that year (Figure 15). Since the foreign reserve is mainly invested in safe assets with high liquidity, e.g., the US treasury bonds, the return to foreign reserve is very low. It is a curious fact that a fast-growing country like China, in which capital is presumably scarce and return to investment is high, lends a tremendous amount of money at generous terms to a developed country like the US. The welfare loss to China is obvious. And this is the price to pay for maintaining a stable exchange rate with respect to the US dollar.

Figure 15: China's foreign reserves.



†Data source: WIND

China's Exchange Rate Regime

The currency of China is called Renminbi (RMB), meaning "people's money." Interestingly, there are two (largely) segmented markets for RMB, the onshore RMB (CNY) market and the offshore RMB (CNH) market. CNH is a clean float and CNY is a managed float.

The managed float of CNY works as follows[1]. On each trading day, the People's Bank of China (PBC) announces central parity rates (CPR) for important currencies around 9:15 a.m. The most important CPR is the one for USDCNY, the exchange rate of CNY with respect to USD. CPR's for other currencies (e.g., JPY) are consistent with market rates (e.g., USDJPY).

CPR serves as the benchmark, and the market exchange rates can fluctuate within a trading band of 2% around CPR. Heavy trading of CNY exchange rates takes place in the interbank foreign exchange market, which opens at 9:30 a.m. and closes at 11:30 p.m.

In contrast, the offshore market (mostly in Hong Kong and Singapore) for CNH trades around the clock. And there is no trading band on CNH exchange rates. The trading volume of CNH, however, is minuscule compared to that of CNY in the onshore market.

CPR has become the main instrument for PBC to manage CNY exchange rates. Currently, PBC uses "the previous closing rate plus changes in the currency basket" to determine CPR[2]. The previous closing rate in the onshore market reflects the demand for and supply of foreign currencies. And the consideration of changes in the currency basket allows CNY to be relatively stable with respect to a basket of currencies.

5.2 Modeling a Small Open Economy with a Floating Exchange Rate

We now introduce a small open-economy model with a floating exchange rate. The model, together with the small open-economy model with a fixed exchange rate, is often called the Mundell–Fleming model.

We may imagine that there are only two economies in the world, one small and the other large. The large economy is much bigger than the small, to the extent that the large economy determines the world real interest rate, while the small economy has no influence at all on the interest rate. The small economy is a price taker.

5.2.1 The Model

The model is a close relative to IS-LM, consisting of two equations:

$$Y = C(Y-T) + I(r^*) + G + X(e) \tag{8}$$

$$\frac{M}{P} = L(r^*, Y), \tag{9}$$

where r^* is the world interest rate, and $X(e)$ is the net export, which is a decreasing function of the nominal exchange rate e. Under the sticky-price assumption, changes in the nominal exchange rate are equivalent to changes in the real exchange rate.

We may re-write the IS equation in (8) as

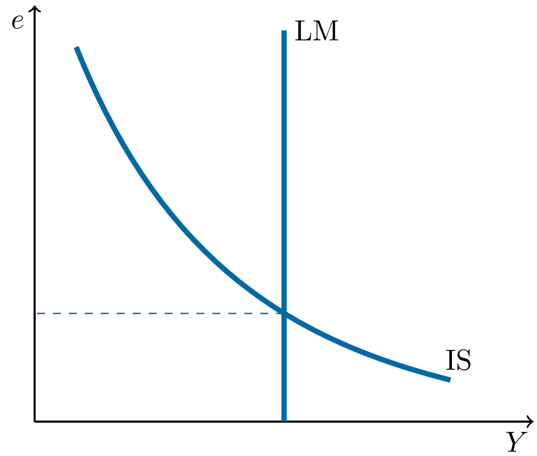
$$S(Y,T,G) - I(r^*) = X(e),$$

where $S(Y,T,G) \equiv Y - C(Y-T) - G$ is the national saving. Now we may understand the IS equation as an equilibrium condition in the foreign exchange market. The supply side of the foreign currency is $X(e)$, which is earned from net export. The demand side is the excess saving $S(Y,T,G) - I(r^*)$, which has to flow out of the country.

Note that the LM equation is the same as in the plain-vanilla IS-LM model, except that the interest rate is now exogenous. That is, the monetary policy of the small open economy cannot change the interest rate.

The two endogenous variables in the model are Y and e. The IS equation defines an implicit function $Y(e)$. Since a lower e corresponds to a higher Y (using the Keynesian Cross analysis), the IS curve is downward sloping. There is only one endogenous variable Y in the LM equation, which uniquely determines Y. Hence the LM curve is vertical. Figure 16 illustrates the joint equilibrium in the foreign exchange market and the money market.

Figure 16: IS-LM curves of the small open economy with floating exchange rate.



5.2.2 The Effect of Fiscal and Monetary Policies

Suppose that the economy is in a recession, producing less than the output potential. If the government conducts fiscal stimulus (e.g., increasing government expenditure), then the IS curve shifts to the right. But the LM curve is vertical. The rightward shift of the IS curve only produces an appreciation of the domestic currency, which fully "crowds out" the effect of the fiscal stimulus.

If the monetary authority conducts monetary stimulus, then the LM curve shifts to the right. This leads to the depreciation of the domestic currency and higher output and employment. Indeed, the small open-economy model with a floating exchange rate characterizes an *exchange rate channel* of monetary policy transmission:

$$M \uparrow \;\Rightarrow\; e \downarrow \;\Rightarrow\; X(e) \uparrow \;\Rightarrow\; Y \uparrow .$$

That is, although the monetary authority of the small open economy cannot cut the interest rate, it can still depreciate its currency to stimulate the economy.

We may also analyze another interesting scenario: The world interest rate r^* rises. This may happen when a large economy tightens its monetary policy. For the small open economy, the IS curve shifts to the left, and the LM curve shifts to the right, producing a depreciation of the domestic currency and somewhat counterintuitively, a higher output. The increased output is brought about by the currency depreciation, which stimulates exports. This experiment illustrates the merit of the flexible exchange rate, which acts as a buffer between global monetary conditions and the domestic economy. If the exchange rate is not flexible, then the *global tightening* (in the form of rising r^*) would transmit to the domestic economy.

5.3 Modeling Small Open Economy with a Fixed Exchange Rate

Now we consider a small open economy with a fixed exchange rate. To maintain the fixed exchange rate, the monetary authority of the small economy stands ready to sell the foreign currency using foreign reserve when there is depreciation pressure, or buy the foreign currency when there is appreciation pressure, at the fixed exchange rate. Selling foreign currency means withdrawing domestic currency from the economy, and buying foreign currency means issuing domestic currency into the economy. Since both operations would change the money supply, the money supply in the small open economy with a fixed exchange rate is necessarily *endogenous*.

5.3.1 The Model

The IS-LM model of the small open economy with a fixed exchange rate is given by the following sytem of equations:

$$Y = C(Y-T) + I(r^*) + G + X(e^*) \tag{10}$$

$$\frac{M}{P} = L(r^*, Y), \tag{11}$$

where r^* is the world interest rate and e^* is the fixed exchange rate. The endogenous variables are output (Y) and money supply (M), and the exogenous variables include r^*, e^*, T, G, and P.

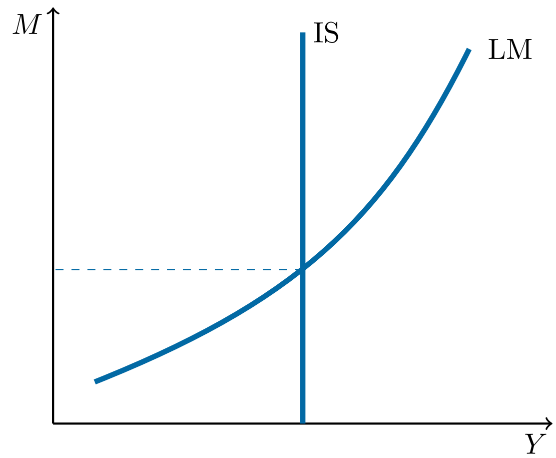
It is clear that the IS equation in (10) determines the level of output or income, Y, regardless of the money supply (M). As a result, if we draw the IS curve on the two-dimensional coordinate system with Y on the horizontal axis and M on the vertical axis, then the IS curve would be vertical (Figure 17). In the LM equation (11), higher M would correspond to higher Y. As a result, the LM curve is upward sloping (Figure 17).

5.3.2 The Effect of a Foreign-Demand Shock

Suppose that a major trade partner goes into recession and that the net export declines. For example, the US economy plunged into recession in 2008. It was a foreign-demand shock for many countries, including China.

If a small economy is initially at full employment, producing the potential output, then the foreign-demand shock would shift the IS curve to the left, causing a foreign-demand-shock recession. Note that economies with fixed exchange rates are more vulnerable to foreign-demand shocks. In an economy with a flexible exchange rate, currency depreciation may prevent the economy from going into recession, as shown in the analysis of the small-open-economy model with a floating exchange rate.

Figure 17: IS-LM curves of the small open economy with fixed exchange rate.



5.3.3 Case Study: China's Slowdown in 2008-2009

China's GDP in 2007 was about 26% of the US GDP. At the same time, although the RMB exchange rate was no longer pegged to USD, it was still very inflexible, so inflexible that it was called a *crawling peg*. Consequently, we may apply the small open-economy model with a fixed exchange rate to China's economy back then.

From 2005 to the first half of 2008, China maintained double-digit real GDP growth rates, thanks to the domestic investment boom and the US consumption boom. As the Global Financial Crisis erupted in the US, the US demand for Chinese goods took a nosedive. In the first quarter of 2009, the Chinese export to the US declined 14.8% compared with the same quarter in 2008. In the first quarter of 2009, the slowdown of the Chinese economy reached the bottom at 6.4%. Starting from the second quarter of 2009, however, the Chinese economy started to rebound vigorously, thanks to the massive *Four-Trillion-Stimulus* program. The export to the US did not start to rebound until the fourth quarter of 2009.

5.3.4 The Effect of a Fiscal Stimulus

Suppose that the economy is in a recession, producing less than the output potential. If the government conducts fiscal stimulus, say increasing the government expenditure, then the IS curve shifts to the right, raising both output and money supply. To understand why money supply would expand, recall that in the small open economy with a floating exchange rate, fiscal stimulus would tend to strengthen the domestic currency. To keep the domestic currency from appreciation, the monetary authority has to issue domestic currency and buy the foreign currency, hence the monetary expansion.

Note that the small open-economy model with a fixed exchange rate does not

Table 1: China's slowdown in 2008 and 2009.

Quarter	Real GDP growth (YoY growth, %)	Export to the US (YoY growth, %)
2008Q1	11.5	5.4
2008Q2	10.9	12.0
2008Q3	9.5	15.3
2008Q4	7.1	0.7
2009Q1	6.4	−14.8
2009Q2	8.2	−18.5
2009Q3	10.6	−16.7
2009Q4	11.9	0.6

† Data source: WIND.

predict any "crowding-out" effect: The increase in output is the same as the rightward shift of the IS curve. This is because both the interest rate and the exchange rate in the IS equation are fixed. The interest rate is the world interest rate, which the small economy cannot influence. The exchange rate is held constant by the monetary authority, which in effect cooperates with the fiscal authority by conducting a monetary expansion.

5.3.5 The Effect of Global Monetary Tightening

We may also analyze another interesting virtual experiment: Suppose that the world real interest rate rises. This may happen when the US starts to raise interest rates. Since the US dollar is the dominant funding currency in the international monetary system, the US monetary tightening is equivalent to a global monetary tightening.

In the small open-economy model with a fixed exchange rate, the global monetary tightening would shift the IS curve to the left and the LM curve to the right, resulting in lower output and money supply. That is, we have a combination of economic recession and tightened monetary policy.

To understand why the money supply has to be tightened, note that the rise in the world interest rate puts depreciation pressure on the domestic currency (using the analysis of the small open-economy model with a floating exchange rate). Defending the fixed exchange rate, the monetary authority has to buy the domestic currency, which is to withdraw money from the economy.

The prediction from this virtual experiment rhymes with many episodes in the history of international finance. Whenever the US interest rate rises substantially, many small countries with rigid exchange rates would experience economic slowdown, monetary tightening, and even currency and banking crises. The Asian Financial Crisis in the late 1990s was a vivid example.

5.3.6 Case Study: The 1997–1998 Asian Financial Crisis

The Asian Financial Crisis started in 1997 and gripped much of East Asia and Southeast Asia. Among the most affected are Thailand, Indonesia, and South Korea. Malaysia and the Hong Kong Special Region of China were also badly hit.

The seed for the crisis was sown in the early 1990s when financial liberalization accelerated in Asia. Financial liberalization generally includes the removal of credit control, interest rate liberalization, privatization of banking, removal of capital account restrictions, and so on. Policy makers at the time generally believed in the free-market doctrine that financial liberalization was good for economic development. However, with inadequate financial regulation and macroprudential management, financial liberalization made the financial system fragile and prone to crisis.

The liberalization of capital account transactions (e.g., foreign lending, portfolio investment), in particular, played an important role in the Asian Financial Crisis. After the liberalization, international money poured into the region to search for higher yields. And local banks and firms ventured out to borrow the US dollar and Japanese yen, two popular funding currencies, at much lower interest rates than they can get domestically.

The Asian economies were booming even before financial liberalization. The massive capital inflow after the liberalization led to overheated economies and asset bubbles. Beneath the apparent prosperity, however, there was a combination of currency mismatch (liability in foreign currency, revenue in local currency) and term mismatch (short-term debt, long-term investment) in many Asian countries. A disaster was waiting to happen when banks and firms were in heavy foreign short-term debt and their investment was in the nontradable sector, real estate in particular.

Several factors contributed to the eruption of the crisis. An important one is the global monetary tightening. The average US federal funds rate in 1993 was 3.0%, which was very low by historical standard. From 1994, however, the US Federal Reserve started to raise interest rates. It rose to 4.2% and 5.8% in 1994 and 1995, respectively. And it stayed around 5.4% from 1996 to 1998. At the same time, the US dollar started to appreciate against other currencies. The dollar index climbed from around 83.4 in 1995 to around 93.9 in 1997 and 98.5 in 1998 (Table 2).

As the global monetary condition tightened, Asian currencies were under pressure. For example, the Thai Baht exchange rate was kept around 25 Baht/USD before July 1997. Although the government pledged to defend Baht, it lacked foreign reserves to make the pledge credible. Massive speculative attacks forced the Thai government to float the Baht on July 2, 1997, a date that marked the beginning of the Asian Financial Crisis. The devaluation of the Thai Baht started chain reactions that led to a region-wide crisis.

Table 2: The federal funds rate, dollar index, and the Asian exchange rates.

	Fed. Funds Rate (%)	Dollar Index	Exchange Rates (per Dollar)		
			Baht	Won	Rupiah
1993	3.02	89.9	25.3	805.8	2110
1994	4.20	88.4	25.2	806.9	2200
1995	5.84	83.4	24.9	772.7	2308
1996	5.30	87.2	25.4	805.0	2383
1997	5.46	93.9	31.1	953.2	4650
1998	5.35	98.5	41.3	1400.4	8025
1999	4.97	97.0	37.9	1189.8	7085
2000	6.24	101.8	40.2	1130.9	9595

†Data source: WIND.

After the floating, Baht devalued quickly and the Thai government had to ask for assistance from International Monetary Fund (IMF), which was supposed to act as the lender of last resort for member countries. Although the IMF did provide a rescue package, it demanded fiscal and monetary tightening on the Thai government and restructuring and reform on the financial sector. The fiscal and monetary tightening made the recession worse. And the chaotic financial restructuring sparked panic runs on many financial institutions in 1998. Later, IMF applied the same approach to other Asian countries and arguably deepened the crisis.[3]

Based on its model, the IMF projected a mild recession for Thailand in 1997 and a 3.5% GDP growth in 1998. However, the Thai economy went into a much deeper recession in 1998, with real GDP declining 7.6% in the local currency and 24.3% in dollar terms.

In Indonesia, the economy contracted 5.4% in dollar terms in 1997 thanks to the devaluation of the Indonesian Rupiah. The next year was much worse. The economy declined 13.1% in local currency and 32.9% in dollar terms. The mass unemployment, coupled with cutbacks in food and fuel subsidies for the poor (thanks to the fiscal tightening), prepared for violent social disruptions. During the May 1998 riots of Indonesia, mobs and organized gangs attacked the Chinese Indonesian communities, including vandalizing shops and sexually assaulting women. This further damaged confidence in the Indonesian economy.

In South Korea, the economy contracted 5.3% in dollar terms in 1997 thanks to the devaluation of the Korean Won. And the next year was more catastrophic. Although the real GDP contracted a mere 5.5% in the local currency, the dollar-denominated GDP declined by 57%.

Although there was speculation that China might follow other Asian countries and devalue its currency, China's Premier Zhu Rongji made it clear that China would not do it. By 1997, China had kept CNY pegged to the USD at 8.3 for over

Table 3: The impact of the Asian financial crisis.

	Real GDP growth (local currency)			GDP growth (in USD)		
	Thailand	S. Korea	Indonesia	Thailand	S. Korea	Indonesia
1994	8.0	9.2	7.5	13.8	13.3	17.9
1995	8.1	9.6	8.2	15.4	14.0	22.1
1996	5.7	7.6	7.8	8.1	12.8	7.5
1997	−2.8	5.9	4.7	−18.0	−5.3	−6.8
1998	−7.6	−5.5	−13.1	−24.3	−57.0	−32.9
1999	4.6	11.3	0.8	11.4	43.6	29.7
2000	4.5	8.9	4.9	−0.2	19.1	15.7

[†]Data source: WIND.

two years. The capital flow into China was mainly in the form of direct investments such as factories. Hence there was little capital flight when the crisis broke out. Furthermore, China had strict controls on capital flow, and it had enough foreign exchange reserves. When China pledged to keep the peg, therefore, it was credible.

There are important lessons to be learned from the Asian Financial Crisis. First, cross-border capital flow, especially in the form of short-term debt, is destabilizing for open economies. Capital control is helpful for developing countries to maintain macroeconomic stability.

Second, the crisis made it clear there was no lender of last resort for individual countries. Even for countries with relatively flexible exchange rates, if without capital control, they could not allow their currency to freefall since the devaluation might have bankrupted firms and banks in short-term foreign debt. To maintain a certain degree of exchange rate stability, countries have to insure themselves by accumulating foreign exchange reserves. And this is exactly what happened after the Asian Financial Crisis.

Many Asian countries, including China, built "war chests" of foreign reserves. Since foreign reserves were mostly in the US treasury bonds, building foreign reserves meant lending (or exporting savings) to the US. Consequently, former Federal Reserve chair Ben Bernanke complained in 2005 that the Asian reserve building contributed to a "saving glut" that kept the US long-term interest rate low despite the fact the Fed was raising the short-term interest rate.[4] The ultra low long-term interest rate arguably fueled the US housing bubble, the burst of which led to the 2008 Global Financial Crisis.

5.4 A Model of a Large Open Economy

The large open economy differs from the small one in that the capital outflow of the large open economy would have an impact on the world interest rate. We use

$F(r)$ to denote the net capital outflow, which is assumed to be a function of the real interest rate r. Since more capital outflow from the large economy depresses the world interest rate, we assume that $F(r)$ is a decreasing function, i.e., $F'(r) < 0$. And we use $X(e)$ to denote the net export, which is assumed to be a decreasing function of the nominal exchange rate e.

We consider the following model:

$$\begin{aligned} Y - C(Y-T) - G &= I(r) + F(r), & (12) \\ \frac{M}{P} &= L(r, Y), & (13) \\ F(r) &= X(e). & (14) \end{aligned}$$

The equation in (12) is the IS equation, describing the equilibrium in the market for loanable funds (or equivalently, the market for goods and services). The equation in (13) is the familiar LM equation, characterizing the equilibrium in the money market. The equation in (14) characterizes the equilibrium in the foreign exchange market, where the net export is the supply side and the net capital outflow is the demand side.

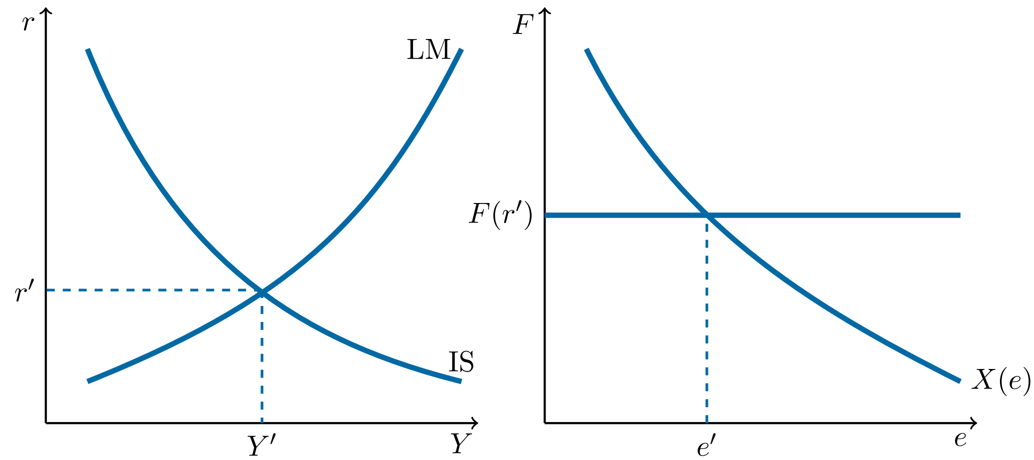
In this model of three equilibrium conditions, there are three endogenous variables: Y, r, and e. All other variables, including the sticky price P, are exogenously given. The model looks complicated, but it is easy to solve. Note that the exchange rate e does not appear in the first two equations. So Y and r can be determined as in the usual IS-LM model. With r determined, we can determine e by solving the equation in (14). Graphically, Figure 18 shows how the model can be solved. The analysis based on the model is straightforward.

Indeed, the directional prediction of policy impact on Y and r is the same as predicted by the usual IS-LM model. Quantitatively, however, there are some differences. For example, the effect of fiscal stimulus would be smaller than in the closed-economy IS-LM model, thanks to the *leakage* through trade. Recall that a fiscal stimulus would drive up the interest rate, which boosts the domestic currency. The appreciation of the domestic currency, in turn, discourages exports and stimulates imports. Thus the effect of fiscal stimulus leaks to foreign countries.

A monetary stimulus, however, may have a bigger impact on output and employment than in the closed-economy IS-LM model. The monetary stimulus not only drives down the interest rate, which stimulates domestic investment, but also devalues the domestic currency, stimulating the net export. That is to say, the large open-economy model characterizes a dual-channel monetary transmission:

$$M \uparrow \; \Rightarrow \left\{ \begin{array}{c} r \downarrow \; \Rightarrow I(r) \uparrow \\ e \downarrow \; \Rightarrow X(e) \uparrow \end{array} \right\} \; \Rightarrow Y \uparrow .$$

Figure 18: The large open-economy model.



5.5 A Summary

A closed economy can be regarded, conceptually, as an extremely large economy. Indeed, the economy of the earth must be closed since there is not interstellar trade so far. From the closed-economy model, we learn that a fiscal stimulus would raise the output and the real interest rate. And a monetary stimulus would also raise the output but lower the real interest rate.

On the other extreme is the small open-economy model, where the interest rate is exogenously given. If the exchange rate is fixed, a fiscal stimulus may raise output and employment. Monetary stimulus, however, is ruled out since monetary policy is no longer independent under the fixed-exchange-rate regime. If the small open economy has a floating exchange rate, it would enjoy an independent monetary policy. A monetary stimulus would lead to the depreciation of the domestic currency, which stimulates output by stimulating foreign demand. Fiscal stimulus, in contrast, is completely ineffective since the appreciation of the exchange rate would lower net export, offsetting the effect of increased government expenditure or a tax cut.

The large open-economy model analyzes an intermediate case between the closed economy and the small open economy with a floating exchange rate. The policy effect on the large economy is something between the effect on the closed economy and that on the small open economy. For example, the effect of a fiscal stimulus on the output would be positive since the effect of the same policy on the closed economy is positive and that on the small open economy with a floating exchange rate is zero. The effect on the large open economy would not be as large as on a closed economy, though, thanks to the leakage through trade.

6 The Keynesian AD-AS Models

Under the sticky-price assumption, we treat the general price level as exogenous in the IS-LM models. To analyze inflation, however, we must allow price to change, and thus treat price as endogenous in models. In this section, we endogenize price and introduce Keynesian AD-AS models that deal with both output and inflation.

Conceptually, we may regard the models in this section as *medium-run* models, in contrast to the short-run models that assume sticky prices and the classical long-run models that assume flexible prices. In the medium run, prices are subject to change but not as flexible as making monetary policy ineffective.

6.1 Aggregate Demand (AD)

6.1.1 From IS-LM to AD

Treating P as endogenous, we may reinterpret the IS-LM model as a special Keynesian AD-AS model:

$$\begin{aligned} Y &= C(Y-T)+I(r)+G && (15)\\ \frac{M}{P} &= L(r,Y) && (16)\\ P &= \bar{P}. && (17)\end{aligned}$$

The preceding model has three endogenous variables, Y, r, and P, and three equations. We may group the first two equations, (15) and (16), together, characterizing the aggregate demand (AD). Note that, given any price level P, the first two equations determine, as in the IS-LM analysis, an *effective demand*. The schedule of effective demand with respect to a whole range of P gives us the AD curve.

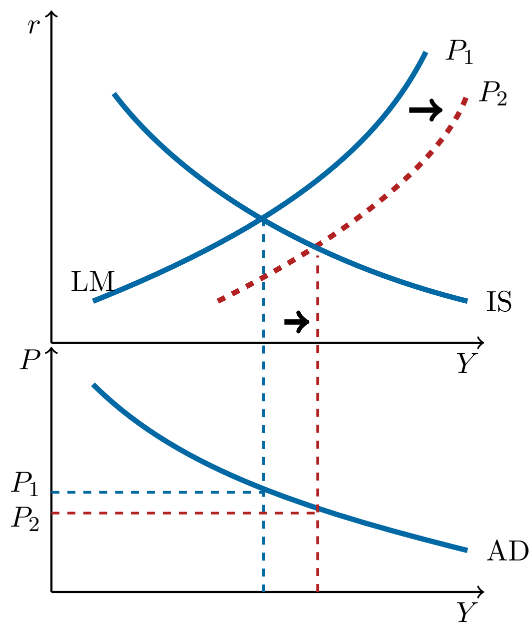
On the other hand, (17) characterizes the aggregate supply (AS) under the sticky-price assumption. The sticky-price assumption reduces a general AD-AS model to the IS-LM model, which is a simple AD-only model.

Mathematically, given a set of exogenous variables (G, T, M), we can solve the IS and LM equations for an equilibrium price level P (eliminating r, one of the endogenous variables). In other words, the IS-LM equations define an implicit function $P(Y, G, T, M)$, which we call the AD curve. Note that Y is an endogenous variable in $P(Y, G, T, M)$, and that G, T, and M are exogenous variables that may shift the AD curve.

To study the shape of the AD curve, we fix T, G, and M, and take total differentiation of both IS and LM equations:

$$\begin{aligned} dY &= C'(Y-T)dY + I'(r)dr,\\ 0 &= L(r,Y)dP + P\left(L_1(r,Y)dr + L_2(r,Y)dY\right). \end{aligned}$$

Figure 19: Derivation of the AD curve, $P_1 > P_2$.



We have thus transformed the IS-LM model into a *linear* model with endogenous variables being dY, dr, and dP. Eliminating dr from this system of equations, we obtain

$$\frac{dY}{dP} = -\frac{I'L}{P\left(L_1\left(1-C'\right)+I'L_2\right)} = -\frac{I'M}{P^2\left(L_1\left(1-C'\right)+I'L_2\right)}.$$

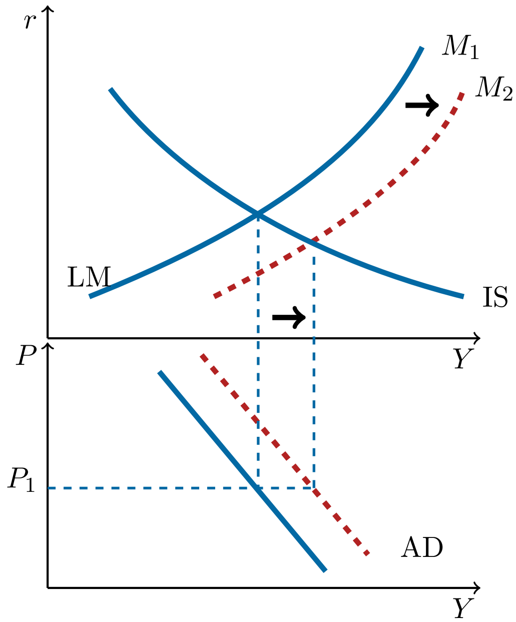
Since $I' < 0$, $L_1 < 0$, $0 < 1 - C' < 1$, we have $dY/dP < 0$, indicating a downward-sloping AD curve.

Figure 19 illustrates the graphical derivation of the AD curve. A decline in the price level P increases the real balance of the money supply. This shifts the LM curve to the right, yielding a higher effective demand Y. Hence the decline in P corresponds to a higher Y. In other words, the AD curve is downward-sloping.

To understand the mechanism behind the downward-sloping AD curve, note that the increase in the real money supply (brought by a decline in P) tends to lower the real interest rate, which further stimulates the investment demand, a major component of the aggregate demand.

And recall that the downward-sloping AD curve cannot be taken for granted in macroeconomics. As discussed in Chapter 3, the fact that almost every good and service has a downward-sloping demand curve does not prove a downward-sloping AD curve. The IS-LM model, as we have seen, illustrates the Keynesian argument for the downward-sloping AD curve.

Figure 20: The effect of a monetary expansion on the AD curve, $M_1 < M_2$.



6.1.2 The Effects of Fiscal and Monetary Policies

Fiscal and monetary policies can shift the AD curve. To see how the AD curve shifts horizontally, we can fix P and examine the effects of the changes in exogenous variables (G, T, M) on Y. But this is exactly the policy analyses based on the IS-LM model. We know the following from the analysis in Section 4.4:

$$\frac{dY}{dG} > 0, \quad \frac{dY}{dT} < 0, \quad \text{and} \quad \frac{dY}{dM} > 0.$$

This says that increased government expenditure, a tax cut, and monetary expansion can all shift the AD curve to the right, meaning that, at every price level P, the aggregate demand Y increases. Figure 20 illustrates the effect of a monetary expansion on the AD curve.

Using the AD curve, we can analyze the effect of policies on inflation, in the short run and in the long run. First, in the short run, the sticky-price assumption implies that the aggregate supply curve is horizontal ($P = \bar{P}$). The shift of the AD curve would thus fail to produce any inflation (Figure 21).

In the long run, however, the sticky-price assumption does not hold, and the economy would produce the output potential $\bar{Y}$ regardless of the price level. In this case, the aggregate supply curve would be vertical ($Y = \bar{Y}$). The shift of the AD curve to the right would thus only produce inflation and fail to raise output (Figure 22).

Figure 21: Short-run effect of a stimulus.

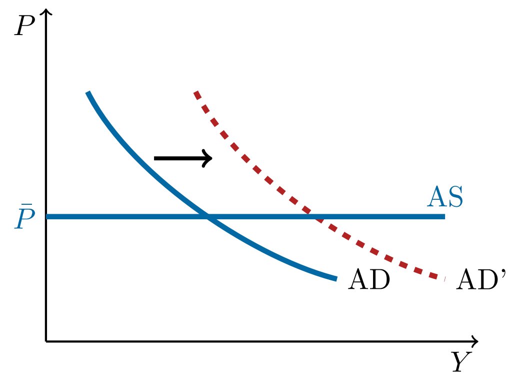


Figure 22: Long-run effect of a stimulus.

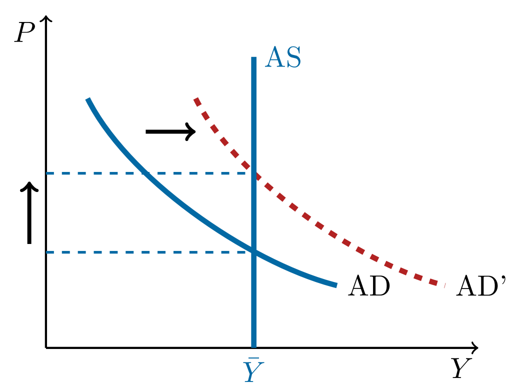
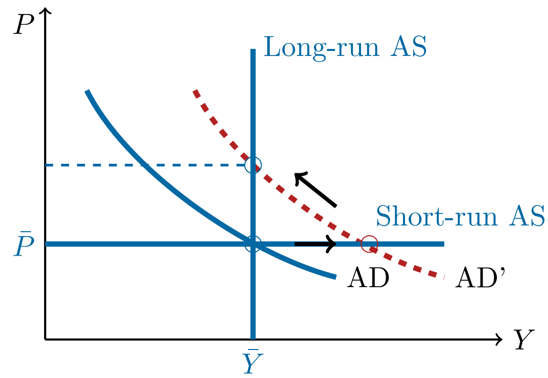


Figure 23: From short-run to long-run equilibrium.



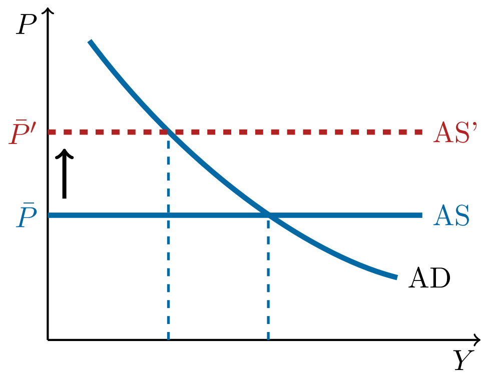
6.1.3 From Short-Run to Long-Run Equilibrium

We may combine short-run and long-run analyses to analyze an unnecessary stimulus. Suppose the economy is in its long-run equilibrium, producing the output potential $\bar{Y}$. But policy makers may want to stimulate the economy anyway, perhaps for a better prospect of staying in power. In the short run, the stimulus measures would raise the output above the natural level without producing much inflation, thanks to the sticky price. As time goes by, however, factor prices (labor, material, energy, etc.) would rise in the overheated economy, forcing firms to adjust production plans. Eventually, the economy would move from the short-run equilibrium back to the long-run equilibrium at $\bar{Y}$, but with a higher price level. See Figure 23 for an illustration.

The preceding narrative is a typical classical argument against government stimulus. However, if we start from an output level *below* the output potential, we can have an argument that is much friendlier to stimulus measures. An output level below the output potential is almost always associated with a high unemployment rate and low inflation (or even deflation). At this point, a fiscal or monetary stimulus would raise the output, reduce the unemployment rate, and start a virtuous cycle toward full recovery. This can be achieved without stirring much inflation since many resources are idle when the output is below the potential.

So knowing the state of the economy is important for making a policy judgment. Indeed, a sensible policy maker can be compared to a physician treating a sick man. He should make all efforts to know about the physiological condition of the sick. He should also try his best to know what has shocked the sick away from the healthy equilibrium. A fever caused by a flu should be treated differently from one caused by injury. For an economic policy maker, he should make all efforts not only to know how the economy is performing but also what kind of shock is affecting the economy. For example, inflation that is caused by demand shocks (e.g., asset bubble, investment boom) should be treated very differently from inflation that is caused by

Figure 24: Stagflation.



supply shocks (e.g., a severe drought, the formation of an international oil cartel, etc.).

6.1.4 Supply Shocks

If a shock reduces the supply of factor inputs or productivity, we call it a *negative* or *adverse* supply shock. A *positive* or *favorable* supply shock, in contrast, increases the supply of factor inputs or productivity. The Industrial Revolution, for example, is a favorable supply shock since it increases the productivity of the economy.

Policy makers do not have to respond to favorable supply shocks, since they generally increase social welfare without bringing inflation risks. The adverse supply shock, however, calls for policymakers' attention since it not only brings inflation but also reduces output and employment.

A famous example of adverse supply shock (or price shock) happened in 1973, when Organization of Petroleum Exporting Countries (OPEC) restricted oil supply. Consequently, the oil price soared, raising the cost of production all over the world. The rise in production costs pushed up the general price level. We often call inflation that is brought by increases in factor prices (labor, raw material, energy, etc.) cost-push inflation.

Given the aggregate demand curve, the exogenous increase in the general price level would shift the AS curve upward, resulting in a lower equilibrium output (Figure 24). The combination of inflation and economic stagnation is famously termed *stagflation*.

Is it sensible to use monetary tightening to bring down the inflation caused by the supply shock? The monetary tightening would shift the AD curve to the left, aggravating the recession and unemployment problem. This policy is not favorable, compared to doing nothing and waiting for inflation to come down.

Another policy option, which we may call the *accommodative* approach, is to conduct monetary or fiscal stimulation to raise aggregate demand and reduce unemployment. The accommodative approach would shift the AD curve to the right to restore the output to the natural level. However, the stimulus may stoke fears of even higher inflation.

The stagflation is, thus, a challenging situation for policy makers. Historically, the stagflation in the 1970s undermined support for Keynesian economics and policy making.

6.1.5 Case Study: The Oil Shocks in 1973 and 1979

In October 1973, OPEC proclaimed an oil embargo on Western countries that supported Israel during the Yom Kippur War. The embargo caused an oil crisis that raised oil and fuel prices dramatically in the Western world. The price of oil in the US, for example, had risen nearly 400% within half a year. This was historically called the first oil crisis.

The second oil crisis happened when the Iranian Revolution succeeded in 1979, replacing a pro-Western authoritarian monarchy with an anti-Western theocracy. The revolution created a widespread panic, sending oil prices doubling within the next year. As if to confirm the panic, the Iran–Iraq War broke out in September 1980. As a result, Iran's oil production almost came to a stop, and Iraq's oil output also dropped substantially.

Both oil crises led to stagflation in the Western world. In the US, for example, inflation rose to double digits after each of the oil crises. Meanwhile, real GDP growth declined to negative territory (Figure 25).

6.2 Aggregate Supply

The horizontal AS curve (under the sticky-price assumption) and the classical vertical AS curve are two special cases of a general theory of aggregate supply, which we now discuss.

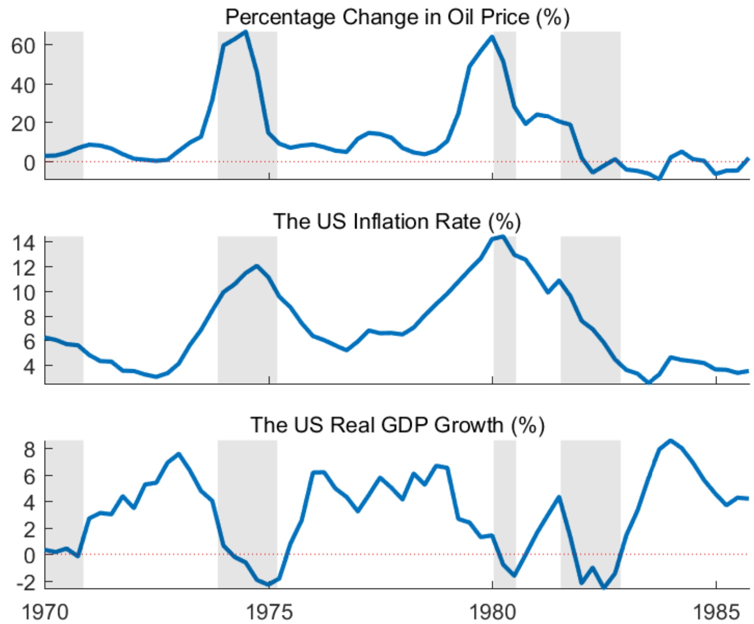
6.2.1 The Case for an Upward-Sloping AS Curve

Recall that the AS curve is the relationship between the general price level P and the total real output that the firms are willing to supply. We may characterize the AS curve by the following function:

$$P = P(Y, Z),$$

where Z is a vector of exogenous variables that may shift the AS curve, and $\partial P/\partial Y > 0$, that is, the AS curve is upward sloping. That the AS curve is upward sloping can be intuitively understood using the "imperfect information" argument. Suppose

Figure 25: The US stagflation in the 1970s and 1980s.



†Data source: fred.stlouisfed.org.

that each firm in the economy produces a single good and there are many goods in the market. If we assume that each firm has imperfect information on the overall price level, then, without perfect information, they may mistake an overall price rise as a rise in the relative price of the good that they produce. If the overall price level rises, many firms make more efforts to produce more, hence the upward-sloping AS curve.

We may further specify the AS equation in a linear form:

$$P = EP + \phi \left(Y - \bar{Y}\right), \tag{18}$$

where Y is output, $\bar{Y}$ is the output potential, P is the price level, EP is the expected price level, and ϕ is a parameter that measures the sensitivity of price to the deviation of output from the natural level. If the price level rises above the expectation, the output will exceed the output potential; if the price level declines, then the economy underperforms its potential. In this linear AS model, both $\bar{Y}$ and EP are exogenous variables. In particular, EP is the intercept of the AS curve, and changes in EP shift the AS curve along the vertical direction.

The linear AS equation in (18) can also be derived using the sticky-price argument. Assume that there are two types of firms in the economy:

(a) Flexible-price firms, which set prices according to

$$P_f = P + \gamma \left(Y - \bar{Y}\right), \quad \gamma > 0.$$

(b) Sticky-price firms, which set the price by

$$P_s = EP.$$

Let $0 < s < 1$ be the fraction of sticky-price firms. The overall price level would be

$$P = sP_s + (1-s)P_f = sEP + (1-s)\left(P + \gamma\left(Y - \bar{Y}\right)\right).$$

Re-arranging the terms, we obtain

$$P = EP + \frac{(1-s)\gamma}{s}\left(Y - \bar{Y}\right).$$

Obviously this is the same equation as in (18) with $\phi = (1-s)\gamma/s > 0$. If we draw the price level P on the Y-axis, which is conventional in economics, the AS curve has a positive slope of ϕ.

The slope of the AS curve is an important parameter for an economy. If the slope approaches infinity, then the AS curve would become vertical, i.e., the classical AS curve. The imperfect-information argument for the upward-sloping AS curve tells us that the AS curve would be steeper for those economies with a volatile aggregate price level, where people would learn to differentiate relative price changes from overall inflation. The sticky-price argument tells us that the AS curve would be steeper for those economies with a higher average level of inflation, where prices are relatively more flexible. Combining these two arguments, we may safely say that the AS curve would be steeper in economies with high and volatile inflation. It is thus no wonder that the classical approach to economics made a comeback in the 1970s, an era marked by exactly high and volatile inflation in developed countries.

6.2.2 Demand-Shock Recession Revisited

Suppose that the economy is initially at the full-employment level. If there is a negative shock to the aggregate demand (e.g., reduced investment sentiment), then the AD curve shifts to the left (Figure 26), resulting in an economy that produces less than the output potential. Meanwhile, the price level declines, resulting in deflation.

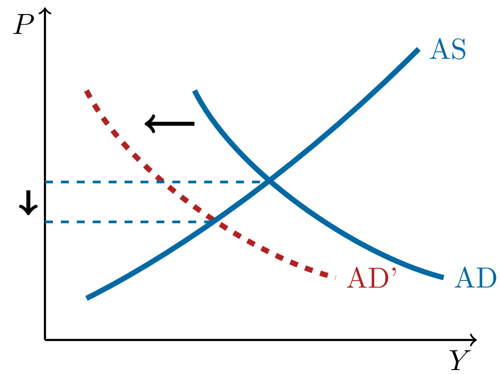
Deflation may further reduce aggregate demand. First, deflation makes the existing debt more difficult to service. It is equivalent to transfering wealth from debtors to creditors. Debtors are typically younger, having higher marginal propensity to consume, and are generally more entrepreneurial than creditors. Therefore, deflation may depress both consumption and investment.

Even an expectation of deflation may immediately dent aggregate demand. We may consider the following IS-LM model:

$$Y = C(Y-T) + I(i - E\pi) + G, \tag{19}$$

$$\frac{M}{P} = L(i, Y), \tag{20}$$

Figure 26: The effect of a negative AD shock.



where i is the nominal interest rate, and $E\pi$ is the expectation of inflation. In this model, the investment expenditure is a function of the ex ante real interest rate. And the endogenous variables are Y and i. Note that the expectation of deflation corresponds to a negative $E\pi$. When $E\pi$ turns from positive to negative, the IS curve will shift to the left, resulting in less output or income.

6.2.3 Responding to Demand Shocks

Relatively speaking, policy makers are more ready to respond to demand shocks than to supply shocks. As seen earlier, the government may conduct expansionary monetary or fiscal policies to stimulate aggregate demand when the economy slides into a recession.

And when there is a positive shock to the aggregate demand (e.g., a surge in investment sentiment) and inflation becomes a problem, then the government may raise interest rates (monetary tightening) or tax rates (fiscal tightening). The tightening would shift the AD curve to the left and cool down the economy.

6.2.4 Case Study: The Great Depression

Before the Great Depression, there was a spectacular bull market on Wall Street in the 1920s. People believed that the bull market was going to last forever, including the famous American economist Irving Fisher. And brokerage firms offered high-leverage financing for their customers to take advantage of the roaring market.

In late October 1929, the US stock market finally crashed. The initial crash led to forced selling by those who bought stocks with borrowed money and who had received "margin calls." A vicious cycle thus began. The stock crash turned the previous extreme optimism into extreme despair. New investment expenditure of the US economy almost came to a stop. The quick deterioration of the real economy further fed back to the stock market.

We may regard the Great Depression as an extreme version of the demand-shock recession. By mid-1932, the Dow Jones Industrial Average (DJIA) had lost almost 90% of its market value. The investment expenditure declined by a similar rate. The real output declined by about 30%. And the unemployment rate rose to 25% in 1933. Meanwhile, the economy went into deflation in 1930 and intensified in 1931. The deflation also contributed to the continued weakness.

The Great Depression did not hit China until 1932. The reasons why China escaped the initial shock waves were twofold. First, the Chinese economy in the 1920s and 1930s was agrarian, and it was not dependent on exporting to foreign markets. Second, the silver-based Chinese currency depreciated against major currencies that were on the gold standard. For example, a Chinese yuan was about 0.71 USD in 1927. In 1930 and 1931, the exchange rate dropped to 0.21 USD per yuan.[5] The sharp depreciation protected the agricultural sector and stimulated import-substituting industries. Furthermore, the overseas Chinese bought silver and sent their wealth home. The consequent monetary expansion led to a short-lived boom in China when other parts of the world suffered from deflation.

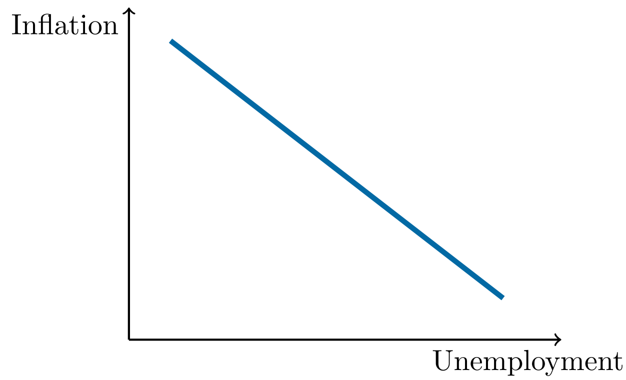
However, the silver stopped flowing into China in 1932 and started to flow out, although initially on a moderate scale. In 1934, the silver outflow became a serious problem when the US passed the Silver Purchase Act, which required the US government hold a quarter of the currency reserves in silver. In 1934 alone, silver worth 257 million yuan flowed out of China, and this did not include the silver smuggled out of the country. At the same time, the silver price doubled within a year. To the Chinese economy, this meant a dramatic appreciation of its currency and a radical monetary tightening due to the outflow of silver.

As a result, deflation finally hit China hard. From 1931 to 1934, farm prices fell 58%. Farmers were especially hard hit, since their taxes were fixed cash obligations. And the weather did not help. Drought and floods wreaked widespread devastation in 1934–1935. The rice harvest in 1934 was 34% below that of 1931.[6] Note that even in 1931, Chinese peasants were close to subsistence level. The Great Depression meant mass unemployment in the West, but famine and depopulation in China. It was under such circumstances that the Nationalist government gave up the silver-based currency in 1935.

6.2.5 The Phillips Curve

If the AS curve is upward sloping and fixed, shifts of the AD curve will produce changes in output and employment, along with inflation. Suppose that the economy is at the full-employment level and there is a negative shock (e.g., a financial crisis) to the aggregate demand. The shock would shift the AD to the left, resulting in less output, higher unemployment, and lower inflation, producing an inverse relationship between unemployment and inflation. The inverse relationship also remains when

Figure 27: The Phillips curve



the economy starts to expand as long as the boost to the economy comes mainly from the demand side. The government may conduct fiscal or monetary stimulus, shifting the AD curve to the right. As a result, output increases, unemployment declines, and inflation rises.

The celebrated Phillips curve is a hypothesis that there exists an inverse relationship between unemployment and inflation (Figure 27). As shown in the preceding analysis, the Phillips curve essentially suggests that short-term shocks to the economy mainly come from the demand side. When A. W. H. Phillips (1914–1975) first published his result in 1958, inflation was not a big concern. So policy makers thought that, based on the Phillips curve, they might reduce unemployment by stoking up inflation.

One of the modern versions of the Phillips curve takes the following form:

$$\pi = E\pi - \alpha(u - u^n) + v, \tag{21}$$

where π is the inflation rate, $E\pi$ is the expected inflation, u is the unemployment rate, u^n is the *natural rate of unemployment*, and v is a random shock to inflation.

Roughly speaking, inflation expectation is the average view of future inflation. Note that the expectation can be self-fulfilling: When everyone expects prices to go up, prices will go up as consumers hasten to purchase and sellers hesitate to sell.

We may call $(u - u^n)$ the *cyclical unemployment*, which is the deviation of the unemployment rate from the natural level. And α is a positive parameter that measures the sensitivity of inflation to cyclical unemployment. A drop in cyclical unemployment leads to the so-called *demand-pull inflation*. Note that when the unemployment rate drops, more workers have jobs, and they have more income to spend, hence there is higher demand in the economy.

We assume that natural rate of unemployment u^n is stable over time and independent of the cyclical unemployment. We call this assumption the *natural-rate hypothesis*. An alternative hypothesis is *hysteresis*, which states that the natural rate of unemployment may be time varying and that the level of the unemployment rate depends not only on the current state of the economy but also on the historical path. For example, a sustained period of unemployment may make the skill set of the unemployed obsolete, hence permanently raising the unemployment rate.

Finally, we may call the random shock v the *supply shock* and the inflation caused by v the *cost-push inflation*.

The Phillips-curve equation in (21) can be derived as follows. Let $p_t = \log(P_t)$. We modify the short-run AS equation in (18) as follows:

$$p_t = E_{t-1}p_t + \phi(Y_t - \bar{Y}) + v_t,$$

where we add a supply shock v_t and $E_{t-1}p_t$ represents the expectation of p_t given information available at time $t-1$. Subtracting p_{t-1} from both sides of the equation, we obtain

$$\pi_t = E_{t-1}\pi_t + \phi(Y_t - \bar{Y}) + v_t. \tag{22}$$

Note that $\pi_t = p_t - p_{t-1}$ and that $E_{t-1}p_{t-1} = p_{t-1}$. Plugging in Okun's law, $(Y_t - \bar{Y}) = -\gamma(u_t - u^n)$, we obtain the Phillips-curve equation:

$$\pi_t = E_{t-1}\pi_t - \alpha(u_t - u^n) + v_t,$$

where $\alpha = \phi\gamma$. The Phillips curve indicates a trade-off between inflation and unemployment: It may take the cost of inflation to reduce unemployment.

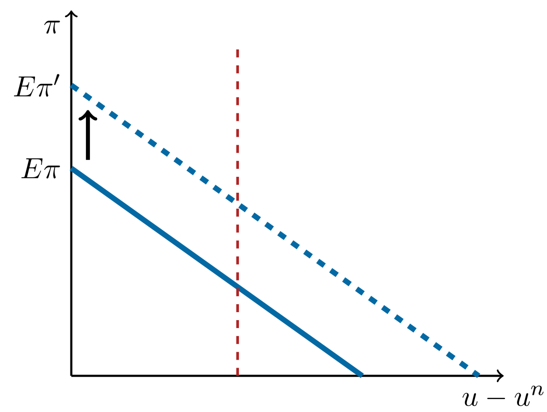
Thanks to (22), we may also say that the trade-off is between inflation and output gap. And we may call $1/\phi$ the *sacrifice ratio*, which measures the percentage of a year's real GDP growth that must be forgone to reduce inflation by one percentage point.

Examining the equation in (21), we see that the Phillips curve may be unstable due to the presence of v and $E\pi$, both of which can shift the Phillips vertically. When supply shocks (e.g., oil shock) were noteworthy in the 1970s, for example, the inverse relationship between unemployment and inflation all but disappeared.

Expectation also plays an important role in making the Phillips curve unstable. Under the *rational expectation* hypothesis, which states that people form expectations using all relevant information and that the expectation is correct on average, the expected inflation would increase as soon as policy makers consider a stimulus package. Consequently, the Phillips curve shifts upward immediately, foiling the attempt to trade more inflation for less unemployment (Figure 28). In the end, inflation may be stoked up with no reduction in unemployment.

In other words, there may be no trade-off between inflation and output (employment) when people can correctly evaluate policies and adjust their expectations of inflation. However, the nonexistence of a trade off between inflation and output (employment) is not necessarily bad news. The good news is that, if the rational expectation hypothesis holds, it may be painless to combat inflation. In other words,

Figure 28: The Phillips curve with inflation expectation.



a credible policymaker may bring down inflation without causing any increase in unemployment.

The rational expectation hypothesis is valuable for theoretical analysis. For empirical analysis, however, the inflation expectation requires more concrete forms. We may measure the inflation expectation by conducting a survey, asking a carefully chosen sample of people of their opinions on inflation. We may also employ econometric models to predict inflation and use the predicted value as expected inflation. The simplest econometric model for predicting inflation assumes that the one-period-ahead inflation equals the current inflation:

$$E_{t-1}\pi_t = \pi_{t-1}.$$

This assumption is called the *adaptive expectation of inflation*. If this assumption holds, then inflation tends to stay at the current level, implying persistence (or inertia) in the movement of inflation.

6.3 A Simple Dynamic AD-AS Model

A static model specifies a set of *simultaneous* relations among all variables. Given an external shock, the endogenous variables shift to another equilibrium immediately and stay there. To study the time-dependent transition from one equilibrium to another, we need dynamic models.

A dynamic model specifies a set of time-dependent relations, which necessarily involves lags of endogenous variables (predetermined variables). For example, if we assume adaptive expectation, the Phillips curve has the form $\pi_t = \pi_{t-1} - \alpha(u_t - u^n) + v_t$, which becomes a dynamic model. Note that the Phillips curve equation is a difference equation, in contrast to the differential equations we repeatedly use in Chapter 4. The differential equations describe smooth changes over time, while difference equations allow nonsmooth changes.

We first consider a static linear AD-AS model:

$$\begin{aligned} Y_t &= 100 + 0.4P_t + u_t, \\ Y_t &= 400 - 0.6P_t + v_t, \end{aligned}$$

where u_t and v_t denote shocks to the supply and the demand, respectively. Note that the model has two endogenous variables (Y_t and P_t), two exogenous variables (u_t and v_t), and two equations. We can *solve* the model and obtain

$$\begin{aligned} P_t &= 300 + (v_t - u_t), \\ Y_t &= 220 + \frac{1}{5}(3u_t + 2v_t). \end{aligned}$$

By "solve" we mean representing endogenous variables using exogenous variables as in the preceding example. Note that although the variables have time subscript, this is still a static model. In this model, a unit negative shock to u_t would result in a unit instantaneous price increase and a 0.6 decline in Y_t. The shock does not affect the future price or output.

If the AS curve does not shift, i.e., $u_t = 0$, then we have

$$\begin{aligned} P_t &= 300 + v_t, \\ Y_t &= 220 + 0.4v_t. \end{aligned}$$

This says that if shocks come only from the demand side, then price and output move together, which is consistent with a downward-sloping Phillips curve since the output is inversely related to the unemployment rate.

Now we assume that the aggregate supply reacts to the lagged price level. Then the model becomes

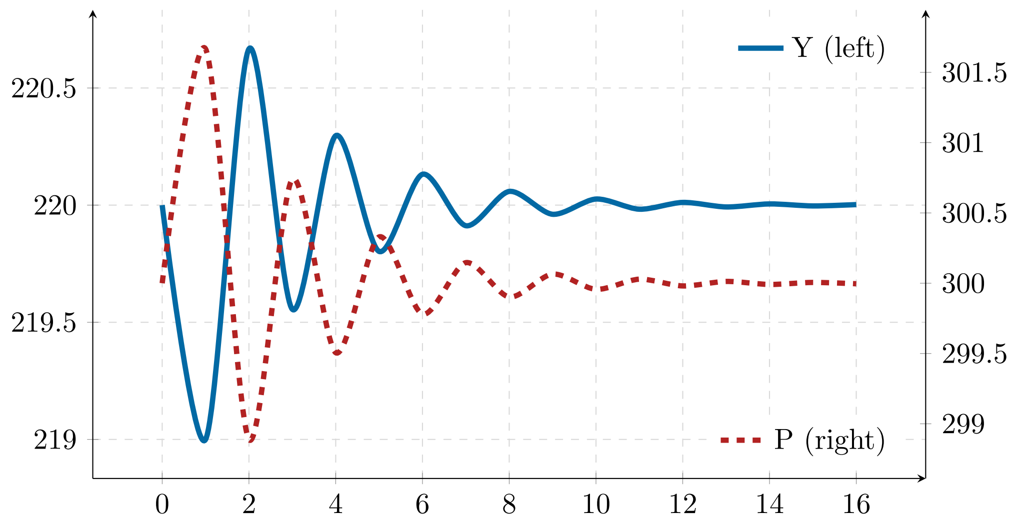
$$\begin{aligned} Y_t &= 100 + 0.4P_{t-1} + u_t, \\ Y_t &= 400 - 0.6P_t + v_t. \end{aligned}$$

The slight change makes the model dynamic, in which a past shock would affect the current and future economy. In this model, Y_t and P_t are endogenous variables, u_t and v_t are exogenous shocks, and P_{t-1} is a predetermined variable. In solving a dynamic model, we may treat predetermined variables as exogenous variables.

We may define a *steady state* of the dynamic model as the state in which the price is constant and there are no shocks ($u_t = v_t = 0$). It is easy to see that in the steady state, we have $P_t = 300$ and $Y_t = 220$.

In the dynamic model, an exogenous shock leads to a series of changes to endogenous variables. For example, suppose that the economy is initially at the steady state ($P_0 = 300$ and $Y_0 = 220$) and that there is a one-period negative shock to the AS at t=1 (say, $u_1 = -1$). Other than this, the supply and demand shocks remain zero, $u_t = 0$ for $t > 1$ and $v_t = 0$ for all t.

Figure 29: Dynamic effect of a supply shock.



The supply shock immediately reduces output and raises the price at $t = 1$. And the effect of the supply shock on the economy will not stop at $t = 1$. In the next period, Y_2 and P_2 depends on P_1, which further depends on u_1. Figure 29 shows a simulation study of the dynamic effect of the preceding supply shock on price and output in the simple AD-AS model. Note that if the demand shock is absent, the model produces an inverse relationship between price and output. We can do the simulation in Figure 29 using an Excel spreadsheet. Interested readers may want to play with the Excel file "simple_adas.xlsx" available at *http://jhqian.org/macrobook*.

6.4 A New Keynesian Dynamic AD-AS Model

We now present a more useful dynamic AD-AS model, which contains three equations: (i) an IS equation; (ii) a "Phillips curve" equation; (iii) a monetary policy rule. We first specify the three-equation model. Then we solve the model, that is, we represent each endogenous variable with exogenous and predetermined variables. Finally, we apply the model to macroeconomic analysis.

6.4.1 The Model and the Solution

We first specify an IS equation as follows:

$$y_t = \bar{y} - \alpha(r_t - r^*) + u_t,$$

where $y_t = \log Y_t$, $\bar{y} = \log \bar{Y}$, r_t is the real interest rate, u_t is the demand shock, α is a constant measuring how demand responds to changes in r_t, and r^* is a constant called the *natural rate of interest*. Based on the IS equation, we may understand r^* as the real interest rate that makes output equal to output potential in the absence of demand shocks.

We further specify r_t as the *ex ante* real interest rate:

$$r_t = i_t - E_t \pi_{t+1},$$

where i_t is the nominal interest rate and $E_t \pi_{t+1}$ is the expected next-period inflation at period t. Note that E_t denotes *conditional expectation* using information available at time t.

Then we use the Phillips-curve-like equation in (22) to model the dynamics of inflation:

$$\pi_t = E_{t-1}\pi_t + \phi(y_t - \bar{y}) + v_t,$$

where $(y_t - \bar{y})$ is the *output gap*, v_t is the *supply shock*, and ϕ is a constant measuring how inflation responds to the output gap.

Finally, to complete the model, we assume that the monetary authority determines the nominal interest rate by the following rule:

$$i_t = \pi_t + r^* + \theta_\pi(\pi_t - \pi^*) + \theta_y(y_t - \bar{y}),$$

where π^* is the inflation target; θ_π and θ_y are constants measuring how the monetary authority would respond to inflation and output gap, respectively; and r^* is the natural rate of interest. Based on this monetary policy rule, we may understand r^* as the real interest rate in a perfect state of the economy where the output equals the output potential and inflation equals the inflation target.

The preceding monetary policy rule implies that the monetary authority has two objectives: (1) Keep a moderate inflation rate; (2) Promote output and employment. The famous Taylor rule for the US Fed is given by

$$\text{FFR} = \text{inflation} + 2 + 0.5(\text{inflation} - 2) + 0.5(\text{GDP gap}),$$

where FFR represents the federal funds rate, which is the main target variable for the US Federal Reserve. The Taylor rule assumes that the natural rate of interest for the US economy is 2%. When inflation reaches the target 2 percent and the GDP gap is zero, then FFR should be 4%.

To make the conditional expectation $E_{t-1}\pi_t$ more concrete, we assume adaptive expectation of inflation (i.e., $E_{t-1}\pi_t = \pi_{t-1}$). Putting all the preceding equations together, we have

$$\begin{aligned}
y_t &= \bar{y} - \alpha(i_t - \pi_t - r^*) + u_t, && (23)\\
\pi_t &= \pi_{t-1} + \phi(y_t - \bar{y}) + v_t, && (24)\\
i_t &= \pi_t + r^* + \theta_\pi(\pi_t - \pi^*) + \theta_y(y_t - \bar{y}). && (25)
\end{aligned}$$

In this model, the endogenous variables are y_t, π_t, and i_t. The exogenous variables are $\bar{y}$, π^*, r^*, u_t, and v_t. And there is a predetermined variable: π_{t-1}.

It is useful first to define a steady state. Here we define the steady state as the state where inflation is constant and there are no shocks, $u_t = v_t = 0$ for all t. It is easy to see that in such a steady state, $y_t = \bar{y}$, $\pi_t = \pi^*$, $i_t = \pi^* + r^*$, and $r_t = r^*$. Thus, real variables (y_t, r_t) do not depend on monetary policy, and monetary policy only influences the nominal variables (π_t, i_t). Hence the steady state satisfies the classical dichotomy and monetary neutrality.

There are three equations and three unknowns (endogenous variables). We may solve the system of equations and obtain

$$\begin{aligned}\pi_t &= a_1(\pi_{t-1} + v_t) + a_2\pi^* + a_3 u_t, \\ y_t &= \bar{y} + a_4\left(\pi^* - \pi_{t-1} - v_t\right) + a_5 u_t,\end{aligned}$$

where $a_1 = \frac{1+\alpha\theta_y}{1+\alpha\theta_y+\phi\alpha\theta_\pi}$, $a_2 = \frac{\phi\alpha\theta_\pi}{1+\alpha\theta_y+\phi\alpha\theta_\pi}$, $a_3 = \frac{\phi}{1+\alpha\theta_y+\phi\alpha\theta_\pi}$, $a_4 = \frac{\alpha\theta_\pi}{1+\alpha\theta_y+\phi\alpha\theta_\pi}$, and $a_5 = \frac{1}{1+\alpha\theta_y+\phi\alpha\theta_\pi}$. Similarly, we may represent i_t as linear functions of exogenous and predetermined variables, simply by plugging the preceding solutions to π_t and y_t to the monetary policy rule in (25).

To apply the model, we must assign values to α, ϕ, θ_π, and θ_y. A natural approach would be to statistically estimate these parameters using data. But statistical estimation is most often infeasible. Instead, economists often rely on *calibration*, an approach that directly assigns values to the parameters in the model, drawing on empirical and experimental studies. Here, we simply set $\alpha = 1$, $\phi = 0.25$, $\theta_\pi = 0.5$, and $\theta_y = 0.5$. Furthermore, we assume $\bar{y} = 10$, $\pi^* = 2$, and $r^* = 2$. Next, we perform a couple of experiments on the model.

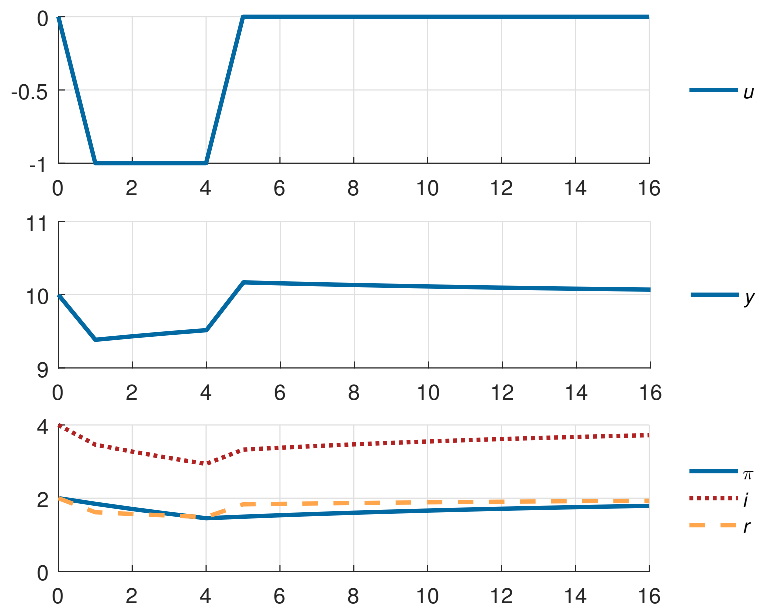
6.4.2 Demand-Shock Recession

Suppose that the economy is initially at the steady state and that there are four consecutive negative shocks to the IS equation, $u_1 = u_2 = u_3 = u_4 = -1$, each period representing a quarter. From the fifth quarter, the demand shock u_t is back to zero and $u_t = 0$ for $t > 4$. Figure 30 shows the dynamics following the shock.

The demand shock produces a recession that lasts only one quarter. However, the output gap remains negative for four quarters. From the second quarter, the economy starts to recover, thanks to the expansionary monetary policy (i.e., the real interest rate declines). In the fifth quarter, when the demand shock ends, the output is already above the output potential. At the same time, the monetary policy shifts to tightening, raising the real interest rate. The continued monetary tightening prevents the economy from further overheating. The output declines and converges to the output potential in the long run.

As the output declines in the recession, inflation also declines. The decline of both output and inflation leads to aggressive cuts in the nominal interest rate, from

Figure 30: Dynamic analysis of a demand-shock recession.



4% to below 3% in the fourth quarter. The aggressive monetary policy prevents inflation from steep declines. From the fifth quarter, inflation starts to pick up. But it stays below the target inflation and slowly converges to the target.

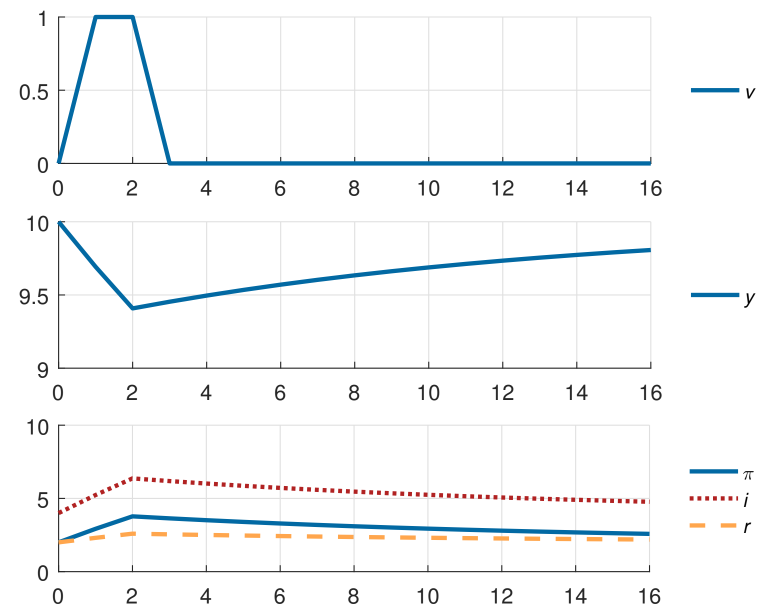
6.4.3 Supply-Shock Recession

Suppose that the economy is initially at the steady state and that there are two consecutive shocks to the Phillips-curve equation, $v_1 = v_2 = 1$. These shocks directly push up inflation. Hence we may interpret them as supply shocks or price shocks. From the third quarter, the supply shock v_t is back to zero, $v_t = 0$ for $t > 2$. Figure 31 shows the dynamics following the shock.

The supply shocks produce a recession that lasts two quarters. The economy starts to recover from the third quarter when the supply shock ends. The recovery, however, is slow. In the 15th quarter, the output is still below the potential.

The reason behind the slow recovery is that inflation rises substantially when the supply shocks occur. Inflation rises from 2% to nearly 4% in the second quarter. In contrast, output only declines by 0.6%. According to the monetary policy rule, the central bank should give higher priority to the inflation problem. Therefore, the central bank raises the nominal interest rate aggressively from 4% to 6.4% in the second quarter. The real interest rate rises during the recession. The hawkish monetary policy helps reduce inflation but slows down the recovery in output.

Figure 31: Dynamic analysis of a supply-shock recession.



Interested readers may use the Excel file "dynamic_adas.xlsx," available at *http://jhqian.org/macrobook*, for other experiments on the New Keynesian dynamic AD-AS model.

7 Keynes Theory of Employment and Investment

The IS-LM and AD-AS models are interpretations of Keynes's theory by the mainstream Keynesian or New Keynesian economists. These models are popular because they, using simple mathematics, have clear formulations with clear implications.

The mathematical models, however, lose some key insights of Keynes. In this section, we give an outline of Keynes's key ideas in his masterpiece, *The General Theory of Employment, Interest, and Money*.[7] The central questions are: (a) What determines the level of employment? (b) What determines the investment?

7.1 Theory of Employment

Keynes proposes an AD-AS model of employment that looks like the Keynesian Cross model. He postulates an *aggregate demand function* to characterize the demand side of the economy as a whole:

$$D = f(N),$$

where D is the proceeds (or revenue) that firms expect to receive from the employment of N workers. Note that the demand curve for a firm characterizes the changing quantity of a commodity that will be purchased at varying prices. For the economy as a whole, however, we cannot measure the total output in any simple unit, such as a ton. Instead, Keynes uses the number of workers employed to measure the total output. On the other hand, D corresponds to *price*, and it may be called the aggregate demand price. Keynes assumes that the aggregate demand function is increasing in N. That is, demand increases with employment.

On the supply side, Keynes postulates an *aggregate supply function*,

$$Z = \phi(N),$$

where Z is the *aggregate supply price* of the output from employing N workers. The term "aggregate supply price" needs a definition and some explanations. Keynes defines it as follows:

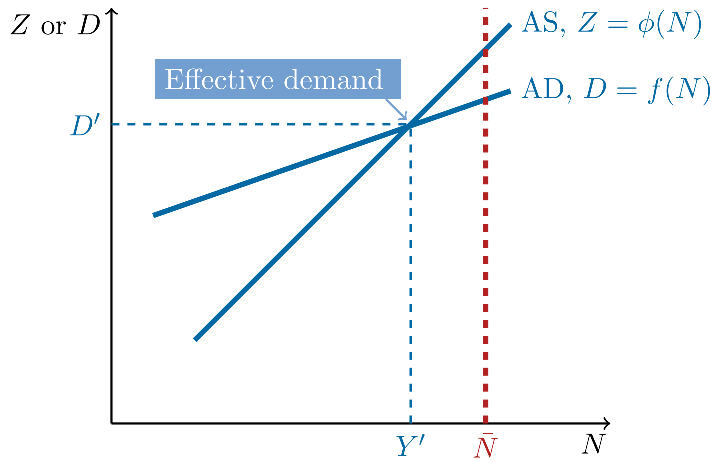
> *... the aggregate supply price of the output of a given amount of employment is the expectation of proceeds which will just make it worth the while of the entrepreneurs to give that employment.*

In a market economy, entrepreneurs make employment (thus, production) plans to maximize their profits. For a firm employing N_i workers, the manager may expect a minimum amount of proceeds, D_i. If he expects a revenue lower than D_i, he will reduce employment. Since both N_i and D_i can take simple summation, it is conceptually clear that, in aggregate, there is a minimum amount of proceeds that the whole entrepreneur class requires to employ a varying amount of total employment. The aggregate supply price is exactly the minimum amount of proceeds to induce a certain amount of employment.

The aggregate supply price should be higher than the expected factor costs, including wage costs in particular. We may assume that $Z = W \cdot N + R$, where W is the average wage, $W \cdot N$ is the total wage bill, and R represents the sum of economic profits and other factor costs. R may or may not increase as entrepreneurs consider more employment of labor. But it is unlikely that R decreases as entrepreneurs consider increasing labor inputs. To induce more employment, thus, the aggregate supply price must increase. As a result, the aggregate supply function is also increasing in N (Figure 32).

Furthermore, we may establish two facts about the AD and AS curves: (i) The AS curve has a steeper slope than the AD curve; and (ii) the AS curve is below the AD curve when the employment level N is very low. As a result, the AD and AS curves will have an intersection, which determines the equilibrium employment level N' and the effective demand D'. If the aggregate supply price is lower than the aggregate demand price for a given value of N, then firms will increase employment and production, driving up factor costs and hence the aggregate supply price, up to

Figure 32: Aggregate demand and aggregate supply.



the level of employment for which Z has become equal to D. If the aggregate supply price is higher than the aggregate demand price for a given value of N, then firms will find it unprofitable to offer so much employment. They will cut costs by slashing jobs and scaling down production to the level of employment for which Z has become equal to D. In other words, the entrepreneurs maximize their expectation of profits at the intersection of the AD and AS curves.

To see why the AS curve has a steeper slope than the AD curve, we note that wages are income to workers, and workers' consumption expenditure is a major part of the aggregate demand for entrepreneurs. Recall that the marginal propensity to consume is less than one, hence workers consume only part of their increased income. If we hold the average wage W constant, the increase of employment ΔN will bring an increase of $W \cdot \Delta N$ in the aggregate supply price. The increase in the aggregate demand price, however, will be only $MPC \cdot (W \cdot \Delta N) < W \cdot \Delta N$. That is to say, as more employment increases wage bills, the aggregate supply price rises faster than the aggregate demand price.

To see why the AS curve is below the AD curve when the employment level N is very low (say, far below the full-employment level), note first that the labor costs (and hence the aggregate supply price) should be very low if jobs are scarce. Second, note that people usually try hard to maintain a certain level of consumption, even when their incomes decline during recessions. When current income cannot cover current consumption, they may draw on their past savings. In the worst case, people have to maintain a minimal level of consumption just for survival. If there are no savings to draw from, they have to borrow from others, which is to draw on other's savings. Although the labor income goes down together with the labor costs during economic downturns, the aggregate consumption is more resilient thanks to the possibility of drawing on savings and borrowing. As a result, the aggregate demand price will be higher than the aggregate supply price when N is far below the full-employment level.

The intersection of the AD and AS curves determines, on the one hand, an equilibrium level of employment N' and, on the other hand, the *effective demand* D', which refers to the expected proceeds that just equal the necessary proceeds that make the employment of N' profitable to entrepreneurs.

There is no reason to believe that the intersection always falls on the full-employment level $\bar{N}$. Suppose that the economy is initially at the full-employment equilibrium, that is, $N' = \bar{N}$. If there is a negative shock to investment sentiment or propensity to consume, then the AD curve will shift downward. If the AS curve fails to shift downward by the same amount, then the new equilibrium employment will be less than $\bar{N}$. Note that sticky price, sticky wage in particular, may prevent the AS curve to adjust adequately for full-employment.

Keynes calls his theory of employment a *general* theory, in contrast to the classical theory, which he regards as a special case. The classical theory, as in Say's law, postulates that *supply creates its own demand*. As a result, the AD and AS curves must overlap, and competition between entrepreneurs must expand employment up to full employment. But Say's law holds only when money serves only as a medium of transactions. To assume that Say's law holds is equivalent to assuming that there are no obstacles to full employment. The repeated occurrence of prolonged mass unemployment, however, contradicts the classical doctrines.

7.2 Theory of Investment

7.2.1 The Importance of Investment

In Keynes's eyes, investment is important because the investment demand fills the gap between the effective demand and the consumption demand at the equilibrium employment level. As a result, the fluctuation of investment demand drives business cycles.

To see this, we may omit the effect of foreign trade and government expenditure and decompose the effective demand Y into two components:

$$Y = C(Y) + I,$$

where $C(Y)$ denotes an aggregate consumption function, which Keynes calls *propensity to consume*, and I denotes the demand derived from new investment. Note that we should understand Y within the consumption function as the aggregate income. We simplify analysis by assuming that the effective demand equals the aggregate income, which differs from the effective demand in Keynes's General Theory by a user cost of capital.

The propensity to consume and the new investment, together, determine the effective demand and the volume of employment. If there is insufficient new investment, the effective demand has to decline and, with it, the volume of employment. Indeed, a decline in the new investment, as analyzed in the Keynesian Cross model, may have a multiplier (negative) effect on total income and employment.

Since the marginal propensity to consume is less than one and it generally declines as income increases, the investment demand is especially important for wealthy societies, which have wider gaps between the aggregate income and the consumption expenditure. For a wealthy society, which typically enjoys a high productive power, investment means not only the possibility of future growth but also the necessity of maintaining the current living standards. But bad times hit. When a recession or depression occurs, the new investment demand may take a nosedive, forcing many to lose jobs, despite the fact that they are willing to work. Keynes calls this phenomenon the *paradox of poverty in the midst of plenty.*

7.2.2 Marginal Efficiency of Capital

It is of utter importance to study how investment fluctuates in the economy. Keynes in the General Theory offers a much richer account than the simple investment function in the IS-LM models.

Keynes introduces the concept of the *marginal efficiency of capital* to characterize the profitability of new investments. He defines the marginal efficiency of capital as

> *the rate of discount which would make the present value of the series of annuities given by the returns expected from the capital-asset during its life just equal to its supply price.*

Here, the series of annuities refers to the future cash flow of an investment or capital asset. We may also call it the *prospective yield* of the investment. The *supply price* of a capital asset refers to the price that would just induce the production of an additional unit of such assets. A more familiar name for the supply price is *replacement cost.*

In mathematical terms, we first denote the series of *expected* annuities by $Q_1, Q_2, \ldots, Q_n$. The marginal efficiency of capital is the discount rate r_m that solves the following equation:

$$\text{Supply price} = \frac{Q_1}{(1+r_m)} + \frac{Q_2}{(1+r_m)^2} + \cdots + \frac{Q_n}{(1+r_m)^n}.$$

The key word in this definition is *expected.* The current and past returns to a capital asset are irrelevant. Entrepreneurs, when deliberating new investment, have to form expectations of future yields.

For any given type of capital, new investment would depress the marginal efficiency of capital, mainly because the expected annuities per unit of capital would decline when the capital becomes more abundant. Equivalently, for each type of capital, there is a schedule of new investment corresponding to different marginal efficiencies of capital. We may aggregate all these schedules across different types of capital and obtain an economy wide schedule of new investment corresponding to different marginal efficiencies of capital. Keynes calls this the investment demand schedule or the schedule of the marginal efficiency of capital.

Entrepreneurs as a whole will add new investment until there is no longer any class of capital asset with a marginal efficiency of capital higher than the current interest rate. To help understand this, Keynes introduces the *demand price* of a capital asset as the present value of the expected annuities discounted by the interest rate r:

$$\text{Demand price} = \frac{Q_1}{(1+r)} + \frac{Q_2}{(1+r)^2} + \cdots + \frac{Q_n}{(1+r)^n}.$$

If the supply price of a capital asset is lower than the demand price or, equivalently, if $r_m > r$, then entrepreneurs will supply such capital assets, that is, add new investment.

7.2.3 The Inducement to Invest

The marginal efficiency of capital is defined in terms of the expectation of future yield and the current supply price of capital assets.

The supply price of capital assets may be mainly determined by the current technology and prevailing factor prices, such as wages and prices of energy and resources. A technological revolution, for example, may bring a lower supply price, which stimulates investment. In most historical periods, however, technology should be stable in the short term. Wages are notoriously sticky or stable in the short term. Therefore, in the absence of major supply shocks that may result in large fluctuations in energy and resource prices, the supply price of capital assets should be relatively stable.

The expectation of future yield, however, is often subject to major revisions, resulting in large fluctuations in the marginal efficiency of capital. The long-term expectation, in particular, is on extremely precarious ground. Often it is not that it is difficult to forecast the future yields, but that there is a lack of confidence in the forecasts. In modern terms, there are not only risks that can be dealt with using probability distributions but also uncertainties that make any long-term expectation cautious and subject to sudden revisions.

Keyes regards the fluctuation in the marginal efficiency of capital as the cause of business cycles. The optimistic evaluation of the marginal efficiency of capital leads to over-investment in the later phase of economic expansion. For a time, the feverish investment brings higher profits to corporations and confirms the optimism of the investors. Such a self-reinforcing activity, however, is also self-defeating. Sooner or later, the game has to stop since the feverish accumulation of capital would depress the marginal efficiency of capital. When this happens, optimism gives way to pessimism. As a result, the marginal efficiency of capital may collapse suddenly. The ensuing collapse in new investment leads to the collapse of the effective demand and, thus, an economic recession.

The recession continues as long as the marginal efficiency of capital is lower than the interest rate. The dismal state of the economy may change for the better when (i) the monetary authority slashes the interest rate to below the marginal efficiency of capital; (ii) the expectation of future yield improves as business confidence restores; and (iii) the supply price of capital declines by an amount enough for the marginal efficiency of capital to exceed the rate of interest.

7.2.4 The Role of the Stock Market

The stock market exerts a decisive influence on the rate of new investment. If the stock market valuation is low, then there is little incentive for entrepreneurs to make new investments since they can purchase capital assets in the stock market. If the valuation is high, then the entrepreneurs are more than willing to build new assets and try to float them, exploiting the high valuation.

The stock market provides for entrepreneurs and investors with a *conventional valuation* of almost every type of capital asset. The conventional valuation is imperfect but indispensable, because the expectation of future yields, especially the long-term expectation, rests on precarious grounds. The convention, however, has its weak points. That is, the convention itself is built on precarious grounds. Keynes lists five factors that contribute to the precariousness of stock market expectation of future yields.

First, as more and more shares are owned by retail investors or people who have no special knowledge of businesses, the stock market's overall ability to value assets has declined.

Second, short-term fluctuations in profits tend to have excessive influence on the market. Keynes gives two examples: (i) The share prices of American ice companies are seasonally higher in summer; and (ii) the recurrence of holidays may raise the market valuation of the British railway substantially.

Third, the conventional valuation of the stock market results from the mass psychology of a large number of ignorant individuals. The mass psychology, or opinion, may suddenly change due to factors unrelated to the prospective yield. In some abnormal times, the market may be subject to waves of optimistic and pessimistic sentiment. Although individual participants may be rational, the stock market as a whole may behave like a mad man. And the irrational behavior is, sadly, legitimate since there is really no solid ground for forecasting future yields.

Fouth, even professional investors and speculators are concerned not with making superior forecasts of future yields, but with foreseeing changes in the conventional valuation a little ahead of the others. In the ideal world, the professionals will correct the vagaries of the ignorant investors to their own benefits. For example, if the ignorant bid up the share price for no fundamental reasons, the professionals will sell the share to the ignorant (possibly buy a similar stock with similar fundamentals) for a profit. Those who succeed in such games would be contributing to society, in that, a more proper valuation of assets leads to more reasonable and stable investment.

In the real world, however, the *value investors*, who employ all available information to calculate the stream of future yields by the best methods available, may not dominate the market. Instead, the speculators whose only job is to guess crowd psychology may dominate the market valuation. There are several factors that may work against the predominance of the value investors.

For one, human nature desires quick results. It is exciting to make quick money, and remote gains are discounted at a high rate. Furthermore, to ignore short-term fluctuations, the value investors must have greater resources for safety and should avoid leverage. As Keyes says, "markets can remain irrational longer than you can remain solvent." Finally, when a committee is in charge of evaluating investment performance, the value investor will face the most criticism. The committee may not only pay excessive attention to short-term results but also penalize the *contrarian* character that is valuable to successful value investing. "Worldly wisdom teaches us that," Keynes says, "it is better for reputation to fail conventionally than to succeed unconventionally."

Fifth, once a collapse has occurred, the market recovery requires a revival of not only speculative confidence but also the confidence of banks to provide credit. The process of recovery can be lengthy.

In conclusion, the conventional valuation offered by the stock market has deep flaws. The violent fluctuations in the mass psychology of investors and speculators drive the fluctuations in the conventional valuation of prospective yields, which further drives the fluctuations in new investment, which further drives the business cycles. Keynes does not believe that investment should be solely determined by the stock market. He says,

> *When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done.*

On "Animal Spirit"

The instability of our economy comes not only from speculations but also shocks to the "delicate balance" between spontaneous optimism, which leads to actions, and cold mathematical expectation, which leads to inaction.

Keynes observes that a large proportion of our positive activities come as a result of the *animal spirit*, a spontaneous urge to do something. Undoubtedly, enterprises that are born out of hopes benefit society as a whole. If humans are all cold calculators of future benefits and costs, inaction would be the rule of human society, which then would not become as prosperous as achieved today.

The animal spirit, however, requires a supportive political and social atmosphere. Since the prosperity of our economy depends on the delicate balance of animal spirit and cold calculations, a shock to the political or social atmosphere may tip the balance and change the whole picture.

8 Minsky's Financial Cycle

The IS-LM and AD-AS models help us understand why, in recessions, an economy may employ less labor than the labor supply, producing less than the potential level of output. We have also studied how policies such as government expenditure, tax cuts, and monetary easing may get the economy back to the full-employment level.

However, we should not get the impression that the full-employment equilibrium is the norm and that fluctuations are exceptions. We should understand the concept of equilibrium as a dynamic process: Both the full-employment economy and the under-employment economy are in equilibrium. And, importantly, the full-employment equilibrium is as transitory as the underemployment equilibria.

When the economy gets back to the full-employment level, for example, the increased optimism will induce more investment. More investment will lead to higher corporate profit and labor income, reinforcing optimism. Meanwhile, asset prices keep rising and credit conditions keep loosening, as borrowers are more willing to borrow and lenders are more willing to lend. Full of inner dynamism, our economy cannot stay quietly anywhere. As soon as it recovers from a recession and reaches the full-employment state, it will move on to overheat.

The inner dynamism manifests itself most evidently in finance. While Keynes focuses on the role of the stock market, Hyman Minksky develops a theory that centers on the banking industry and the credit market. In his *Stabilizing an Unstable Economy*, Minsky characterizes a three-stage financial cycle. During the stage of *hedge finance*, the economy is dominated by *hedge units*, firms whose cash flows (from capital assets or other financial contracts) are sufficient to cover all cash-payment commitments (interests and principals) now and in the future. At this stage, the economy just starts to recover from a recession. The hedge units may be survivors from the previous financial crisis or new firms grown up from a conservative financial environment, which is also the legacy of the previous crash.

As the economy continues to recover, however, more firms become *speculative units* or even *Ponzi units*. For speculative units, the cash flow may be lower than cash-payment commitment (due to existing debt) in some short-term periods. Speculative units expect to borrow more to cover the gap. And, typically, they expect to roll over their debts when they mature. Compared to hedge units, the viability of speculative units depends not only on business conditions (i.e., cost and revenue) but also on conditions of the financial system.

For Ponzi units, the cash-payment commitment is not only higher than the cash flow but also higher than its income. Ponzi units expect to borrow more or sell assets to meet their future cash-payment commitments. As we can see, Ponzi units are even more vulnerable to worsening conditions in the financial system than speculative units. They need bankers who do not care about their worsening balance sheets. And they need the prices of their assets to go higher and the liquidity of these assets to be adequate.

At the later stage of an economic boom, more and more hedge units become speculative, and more and more speculative units become Ponzi. It happens not just because entrepreneurs and their bankers become more optimistic. It may also happen when revenues from new investments into some crowded industries are lower than expected, or when costs of business (interest rate, wage, etc.) rise substantially, forcing hedge units into speculative units, speculative units into Ponzi units.

When a substantial number of corporations engage in speculative or Ponzi finance, the financial markets and the economy become fragile. And the tendency for a robust economy dominated by hedge units to become a fragile one dominated by speculative or Ponzi units is engrained in our financial economy since engaging in speculative or Ponzi finance is highly profitable in a robust economy dominated by hedge units. The profitability of speculative and Ponzi finance will attract more converts, who will generate more credit and raise the demand for assets. The ensuing rise in asset prices will attract more converts to speculative or Ponzi units. The cycle repeats until the financial system is extremely fragile. At the final stage, only a small shock may be enough to tip the financial market to a crisis.

During the crisis, former speculative or Ponzi units may be liquidated, or they may go through debt restructuring, which is an effort to transform speculative or Ponzi units into hedge units. At least for a while, survivors of the business community and bankers will shy away from speculative and Ponzi finance. Thus, a robust economy dominated by hedge units re-emerges. The financial cycle, which propels the business cycle, goes on.

Note that the financial cycle does not simply repeat in the same style. Every cycle is different because financial markets are structurally changing over time. In the aftermath of each crisis, a responsible government has to identify what has gone wrong and tries to strengthen regulations to prevent that from happening again. However, when panic gives way to optimism on the part of market participants and when alertness gives way to complacency on the part of regulators, financial innovations will come back to life and make new forms of speculative and Ponzi finance possible.

Every cycle is, of course, also similar, especially in that there is always a bust following a boom. To modify Mark Twain's famous quote, the history of economic fluctuations doesn't repeat itself, but it often rhymes.

9 Concluding Remarks

In this chapter, we study a series of models that address short-run economic fluctuations. All these models share a common feature: The market is assumed to be imperfect in some way. The IS-LM models rely on the sticky-price assumption, which says that prices and wages are not flexible enough for markets to clear automatically. The AS model and the Phillips curve rely on the argument of sticky price or imperfect information. Keynes bases his theory of business cycles on an imperfect capital market. And, in the same spirit, Minsky describes a boom-and-bust financial cycle that drives expansions and recessions in the economy.

The economic history of the world speaks to the fact that the market economy is unstable. It was unstable before the government became big enough to influence the economy. It is still unstable after the government has become big and special institutions and policy instruments have been developed to smooth business cycles. Stabilizing an unstable economy is a very difficult job. But it is also a job that all responsible governments have to do. We leave this to the next chapter.

Notes

[1]For a detailed analysis of China's exchange rate regime, we refer to Su and Qian (2021), Structural Changes in the RMB Exchange Rate Mechanism, China & World Economy, 29 (2), 1–23.

[2]People's Bank of China, 2016, China Monetary Policy Report, Quarter One, Available from: www.pbc.gov.cn/en/3688229/3688353/3688356/3706393/index.html.

[3]For a forceful critique of IMF policies during the Asian financial crisis, read Joseph Stiglitz's *Globalization and Its Discontents*, New York: W. W. Norton.

[4]The Global Saving Glut and the U.S. Current Account Deficit, March 10, 2005, available at www.federalreserve.gov/boarddocs/speeches/2005/200503102/.

[5]Dietmar Rothermund, 1996, The Global Impact of the Great Depression 1929–1939, Routledge.

[6]The Cambridge History of China, Volume 13, Republican China 1912–1949, Part 2, edited by John K. Fairbank and Albert Feuerwerker.

[7]For a gentle introduction to *The General Theory*, I refer to Dudley Dillard's *The Economics of John Maynard Keynes: The Theory of Monetary Economy*, New Jersey: Prentice Hall.

References

[1] Bewley, Truman F., 1999, Why Wages Don't Fall in a Recession. Cambridge, MA: Harvard University Press.

[2] Blinder A.S., 1994, On Sticky Prices: Academic Theories Meet the Real World, in N.G. Mankiw, ed., Monetary Policy, University of Chicago Press, 117-154.

[3] Dillard, Dudley, 1948, The Economics of John Maynard Keynes: The Theory of a Monetary Economy.New Jersey: Prentice-Hall.

[4] Fleming, J. Marcus, 1962. Domestic Financial Policies under Fixed and Floating Exchange Rates". IMF Staff Papers, 9, 369–379.

[5] Mankiw, N.G., 2016, Macroeconomics (9th ed.), New York: Worth Publishers.

[6] Minsky, Hyman P., 1986, Stabilizing an Unstable Economy. Yale University Press.

[7] Mundell, Robert A., 1963, Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates. Canadian Journal of Economics and Political Science. 29 (4), 475–485.

[8] Rothermund, Dietmar, 1996, The Global Impact of the Great Depression 1929-1939. New York: Routledge.

[9] Samuelson, P.A., 1948, Economics, New York: McGraw-Hill.

[10] Stiglitz, J. E., 2002, Globalization and Its Discontents. New York: W.W. Norton.

[11] Su, Guanyu and Junhui Qian, 2021, Structural Changes in the RMB Exchange Rate Mechanism, China & World Economy, 29 (2), 1-23.