ECON 1550 Course Notes

Fernando Duarte

Spring 2026

Contents

Chapter I

Models in Economics

1 Models in Economics

"Economists do it with models"

Old Keynesian proverb

A model is a set of equations with economic meaning attached to them. It is important to realize that all models are wrong. Because models are not reality itself, any model we write down will always omit some aspect of reality. Once an element of reality is omitted, it is always possible to find predictions of the model that are at odds with the real world, which means that every model is wrong. Indeed, economic models omit *most* elements of reality, so they are often demonstrably at odds with many empirical observations.

Even though all models are wrong, some models are useful. Models have many uses, including differentiating causal from non-causal effects, allowing counterfactual analysis, rationalizing past economic outcomes, providing forecasts, evaluating the social welfare costs and benefits of policy, clarifying assumptions behind economic statements, and giving insight into specific mechanisms otherwise obscured by confounding effects. Many (perhaps most) economists evaluate models based on their usefulness rather than their correctness or accuracy. The choice of what elements from the real world to omit in a model is just as important as the elements that are included. A substantial source of disagreement among economists is, at its core, a disagreement about whether elements of the real world that were omitted by a model are important for the question under consideration and should instead be included.

Takeaway

All models are wrong, some models are useful.

1.1 Types of Equations

The models that we study in this course have three types of equations:

- **Identities** hold by definition or construction. For example, disposable income is defined by $Y_D \equiv Y - T$ where Y is income and T represents taxes. When an equation defines a variable, we use the symbol "$\equiv$" rather than a regular equal sign. Another kind of identity is an accounting identity, an equality that must be true regardless of the value of its variables by virtue of how the variables are defined or constructed. The national income identity

  $$\text{GDP} = \text{consumption} + \text{investment} + \text{government spending} + \text{exports} - \text{imports},$$

  is an accounting identity. Although none of the variables are defined by the accounting identity, the identity always holds (the equality is always true) because of the way in which all the variables are defined. When consumption increases by $1 and investment, government spending, exports, and imports remain the same, then GDP will always also be measured as having increased by $1. Another accounting identity is the balance sheet relation that

  $$\text{assets} = \text{liabilities} + \text{equity}.$$

- **Behavioral equations** are assumptions that capture some aspect of behavior that we include in a model. For example, a model can postulate that consumers behave so that aggregate consumption is one-half of disposable income. The behavioral equation that captures this behavior is $C = 0.5Y_D$. Behavioral equations hold by assumption and are the equations that distinguish one model from the next.

- **Equilibrium conditions** describe the relationship that must hold between variables in a model when economic forces are balanced. The main equilibrium condition that we will encounter is the condition that supply equals demand. For example, if money demand is assumed to be a function $L(\cdot)$ of the nominal interest rate i, and money supply is assumed to be equal to a constant level M^s, the equilibrium condition that money supply equals money demand is $L(i) = M^s$.

1.2 Types of Variables

All models in economics have two types of variables:

- **Endogenous variables** are variables explained within the model.
- **Exogenous variables** are variables that are taken as given. The behavior and value of exogenous variables are not explained by the model.

Exogenous variables are the inputs that we provide to the model, while endogenous variables are the outputs produced by the model.

The way we use models is by looking at how endogenous variables change when we decide to change the value of one or more exogenous variables. We are not allowed to decide on the value of endogenous variables, or to change their value "by hand" or "by assumption". Only the equations of the model can determine the value of endogenous variables.

Given a set of equations that comprise a model, it is not possible to tell just by looking at the equations which variables are endogenous and which variables are exogenous. A model must include, together with its equations, a list of which variables are endogenous and which variables are exogenous. Whether a variable is endogenous or exogenous is part of the "economic meaning we attach to equations" that was mentioned in the very first sentence of Section 1.

To **solve a model** is to write all endogenous variables only in terms of exogenous variables. To **solve for an endogenous variable** means writing that particular endogenous variable in terms of exogenous variables.

Parameters

We sometimes call certain exogenous variables **structural parameters**, or simply **parameters**. Parameters are exogenous variables that are **policy-independent** or **policy-invariant**, that is, do not change when government policy changes[1]. The exact meaning of policy-invariant can depend on the context or the model. For example, a variable that captures the overall productivity of the workforce is usually considered policy-independent in models of business cycles, but not in models of long-run growth.

Just as for endogenous and exogenous variables, there is no way to know just by looking at the equations of a model which of the exogenous variables are parameters. If a model has parameters, they must be identified separately. Specifying whether a variable is a parameter is not always relevant for the task at hand, in which case we don't need to waste time listing which of the variables are parameters.

Example 1.1. We build a simple model of consumption. The model has five variables: consumption denoted by C, taxes denoted by T, income denoted by Y, disposable income denoted by Y_D, and the marginal propensity to consume denoted by c_1. Taxes, income, and the marginal propensity to consume are exogenous. In addition to being exogenous, the marginal propensity to consume is assumed to be a parameter, whereas income and taxes are not. Consumption and disposable income are endogenous. We assume all variables are strictly positive and, in addition, that $c_1 < 1$. The model has two equations:

$$Y_D \equiv Y - T, \tag{1.1}$$

$$C = c_1 Y_D. \tag{1.2}$$

The first equation is an identity that defines Y_D. The second equation is a behavioral equation that describes consumer behavior. Consumers are assumed to consume a fixed fraction c_1 of their disposable income. This model has no equilibrium conditions.

[1]Economists – myself included! – sometimes consider parameters as a third type of variable (neither exogenous nor endogenous) but I am not sure I can explain why.

Let's now use the model to understand the effect of taxes on consumption. When taxes go up, disposable income goes down by definition. When consumers have a lower disposable income, they reduce consumption.

We cannot use this model to study the effect of consumption on taxes, since consumption is an endogenous variable. It does not make sense to start by saying "consider an increase in consumption" and then trace how the increase in consumption changes taxes. The reason is that taxes, being exogenous, cannot be explained within the model. Any changes in T must be postulated by us; T can only be changed for reasons outside the model. In addition, we cannot simply postulate that consumption increases as consumption is endogenous. If consumption increases, it must be because some change in exogenous variables induced the change.

Now we solve the model, that is, write the two endogenous variables, C and Y_D, only in terms of the exogenous variables Y, T, and c_1. Equation (1.1) already expresses Y_D only in terms of Y and T, so solving for Y_D is immediate. Plugging equation (1.1) into equation (1.2) gives C only in terms of Y, T, and c_1. Therefore, the model's solution is:

$$
\begin{aligned}
Y_D &= Y - T, \\
C &= c_1(Y - T).
\end{aligned}
$$

$\square$

Example 1.2. Consider now a different model that has the same equations as the model in the previous example (equations (1.1) and (1.2)), the same five variables (C, T, Y, Y_D, c_1), but a different set of exogenous and endogenous variables. The exogenous variables of this new model are consumption, income, and the marginal propensity to consume. The endogenous variables are taxes and disposable income.

In this model, it *is* possible to study the effect of consumption on taxes, since consumption is exogenous and taxes are endogenous. Consider then an increase in consumption. Equation (1.2) then implies that the higher consumption leads to higher disposable income. In turn, equation (1.1) shows that, by definition, the higher disposable income is associated with lower taxes.

Let's think about this for a second. This model claims that an increase in consumption leads to an increase in disposable income. Wouldn't that be nice! Of course, this model makes no sense, which is the point: Even though this model has the exact same equations and variables as the model in the previous example, the economic interpretation is completely different because the choice of exogenous variables was different.

One implication is that if in the real world we observe that C and T always go up and down together, two economists using two different models can have a very different interpretation of what is going on in the economy. And they can disagree even if their two models have the exact same equations! □

Shocks

Many times, economists use the word shock to describe changes in exogenous variables. Usually, the term shock is reserved for changes that are unforeseen, unforeseeable, or random, but this is not always the case.

A more formal definition of **shocks** is that they are economically interpretable primitive exogenous forces that satisfy three characteristics:[2]

- They are not explained by past and present endogenous variables.
- They are uncorrelated with other shocks.
- They represent either unanticipated movements in other exogenous variables or news about future movements in exogenous variables.

A technology shock may be the sudden discovery of the steam engine. A tax shock might represent the passing of legislation resulting from a change in political power. The outbreak of COVID-19 can be considered a health shock.

[2]The definition is adapted from: Ramey, Valerie A. "Macroeconomic shocks and their propagation." Handbook of macroeconomics 2 (2016): 71-162.

Takeaway

Models include three types of equations

- Identities hold by construction
- Behavioral equations capture assumptions
- Equilibrium conditions balance economic forces as in supply equals demand

and two types of variables

- Endogenous variables are explained within the model
- Exogenous variables are what the model takes as given
  - Parameters are policy-invariant exogenous variables
  - Shocks are economically interpretable, unanticipated, exogenous variables

1.3 Dynamic Models

A model is **static** when it only characterizes endogenous variables at a given point in time. When a model specifies the evolution of endogenous variables over time, we say that it is **dynamic**.

The list of values that a variable takes over time is called a **path**. We use subscripts to index time. When we use **discrete time**, periods are indexed by non-negative integers $(0, 1, 2, ...)$. For example, if Y_t is output in period t, then $\{Y_0, Y_1, Y_2\}$ is the path of Y in a three-period model. One period can correspond to a day, a month, a quarter, a year, etc. Suppose GDP in three consecutive years is $1 trillion, $1.2 trillion, and $1.4 trillion. The path of GDP is $\{1, 1.2, 1.4\}$.

We can also use **continuous time** and index time by a positive real number. For example, if time t is a real number in $[0, T]$, then a path for output is the collection of values of Y_t for all $t \in [0, T]$, sometimes also written as $\{Y_t\}_{t\in[0,T]}$ or $\{Y_t\}_{t=0}^{T}$. Just as we had to pick what one period means in discrete time, we have to pick units for continuous time. For example, if one unit of time is a year, our model with $t \in [0, T]$ and $T = 3$ encompasses three years. The value $Y_{0.5}$ is output after six months have passed.

When we specify the values of exogenous variables, we must now specify an entire path for each exogenous variable (one value for each time period) rather than a single value as we did before. Similarly, when we solve the model, we must solve for the entire path of each endogenous variable in terms of the paths of exogenous variables.

Takeaway

- Static models consider variables at a single point in time
- Dynamic models consider variables that evolve over time
  - The collection of values of a variable over time is called a path
  - Time can be discrete $t = 0, 1, 2, ...$ or continuous $t \in [0, T]$
  - When we solve a dynamic model, we write the paths of endogenous variables in terms of the paths of exogenous variables

Chapter II

Financial Assets

1 Financial Assets

A **financial asset** is a legal contract between two parties, the buyer and the seller. The contract works as follows:

i. The seller writes on a piece of paper a description of future promised payments. This piece of paper is the financial asset and whatever is written codifies the terms of the legal contract.

ii. The description can be as simple or as complex as desired. For example:

   - *"pay one dollar a day until the end of this year"*,
   - *"pay one dollar a day until the end of this year if, and only if, the stock market went up the previous day"*,
   - *"on the day Greece defaults on its debt, pay one dollar, and pay nothing otherwise"*.

iii. The buyer and the seller agree on a price P.

iv. The buyer gives the seller P dollars and the seller gives the piece of paper (the financial asset) to the buyer.

v. When one of the promised payments encoded on the paper comes due, the seller makes the payment to whomever happens to own the piece of paper at the time of the payment, be it the original buyer or a subsequent owner.

vi. If the seller does not make a payment as promised, the financial asset is said to be **in default**. Any violation of the promises, however small, triggers default. For example, default is triggered if payments are:

- late by a single day, or
- short by a single dollar.

In case of default, the owner of the financial asset can take the seller to court to recover the value of the violated promises.

A **financial liability** is a contractual obligation to deliver the payments associated with a financial asset. In the situation considered above, the seller has a financial liability corresponding to the legal obligation to make the payments promised by the financial asset it created.

Because financial contracts might specify payments from the seller to the buyer, from the buyer to the seller, or both, it is not always the case that the seller ends up with a liability and the buyer with an asset. Indeed, many financial contracts structure future payments so that no money is exchanged when the contract is signed. In this case, neither buyer nor seller can be said to have an asset or a liability. Instead, we refer to the contractual obligations of the buyer and the seller as the two **legs of the contract**.

The terms and logistics surrounding financial assets can be arbitrarily complex. A single contract can stipulate payments among more than two parties, spell out what court jurisdiction can be used in the event of default, prohibit re-selling the asset to third parties, specify what accounts must be used for payments, what happens in case of default, and so on.

Payoffs can be cash payments in dollars or in other foreign currencies, deliveries of commodities such as gold, oil, or soybean, and even transfers of other financial assets. Therefore, when talking about financial assets, we often use the word **payoffs** rather than **payments** to highlight this generality, though both terms are largely interchangeable. When all payments are in dollars, they are usually referred to as **cash flows**.

Takeaway

- Financial assets are legal contracts defined by their payoffs.

1.1 Bonds

A bond is a financial asset. Like all other financial assets, it is a piece of paper with promises of future payoffs written on it. When a bond is created and sold for the first time, we say that the bond is **issued**.

What distinguishes bonds from other types of financial assets—what makes a bond a *bond*—is that the amount and timing of the promised payoffs are fixed at issuance. Absent default, the sizes and dates of the payoffs are thus known with certainty from the moment the bond is issued, and they never change throughout the bond's lifetime.

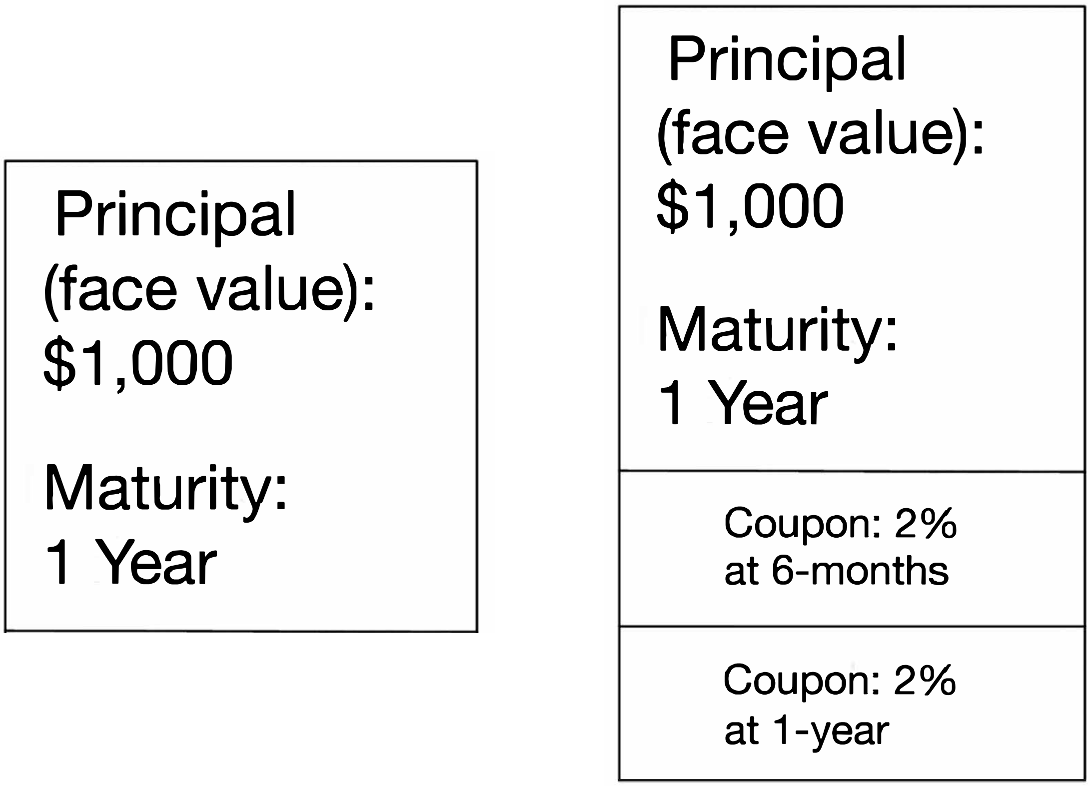
Zero-coupon bonds

The simplest kind of bond is one that promises a single payment at a specified date. Such a bond is called a **discount bond** or **zero-coupon bond**. Figure 1.1(a) provides an example. The promise is for $1,000, to be paid exactly one year after the bond is issued. The promised amount of $1,000 is called the **principal**, **face value**, or **par value** of the bond. The time left until the bond pays the principal is the **maturity** of the bond. After the bond is issued, the maturity decreases as time goes by. For example, if two months pass after a one-year-maturity bond is issued, the maturity of the bond is 10 months.

Coupon bonds

Figure 1.1(b) shows a different kind of bond, a **coupon bond**. It has the same face value and maturity as the bond in Figure 1.1(a), but includes two additional promises, known as **interest payments**, **coupon payments** or simply **coupons**. Coupons are promised payments that occur at fixed intervals. Coupons can be paid before or at maturity.

In the bond of Figure 1.1(b), the first coupon is paid six months after issuance, and the



(a) (b)

Figure 1.1: Panel (a) shows a one-year-maturity zero coupon bond with a principal of $1,000. This bond has a single fixed payment of $1,000 that occurs one year after the bond is issued.

Panel (b) shows a one-year-maturity coupon bond with a principal of $1,000 and a 2% coupon rate to be paid semiannually. This bond has two payoffs of fixed size at two fixed dates. The first payment occurs six months after the bond is issued and equals $10 (an annualized coupon rate of 2% is 1% over six months, and 1% of $1,000 is $10). The second payment occurs one year after the bond is issued and equals $1,010 (the principal of $1,000 plus the second coupon of $10).

second one is paid on year after issuance, concurrent with the principal. Coupons are expressed as a percentage of face value; the ratio of coupon payments to the face value is the **coupon rate**. The coupon rate is always **annualized** or **per annum**. For the bond in Figure 1.1(b), the coupon rate is 2% to be paid semiannually (two times a year, every six months). Because the 2% rate is annualized, each semi-annual payment corresponds to only 1% of face value. With a face value of $1,000, each coupon payment is $10. Together, the payments of this bond are as follows:

- a coupon of $10 six months after issuance,
- a coupon of $10 plus the $1,000 principal (i.e., $1,010 total) one year after issuance.

Takeaway

- Bonds are financial assets with fixed payments.
- Zero-coupon bonds only pay at maturity.
- Coupon bonds also have intermediate payments.

Bond Prices

The date and amount of a bond's payment are written on the bond and do not change when the bond changes hands. The price of the bond—the amount that the buyer pays the seller to become the owner of the bond—is *not* written on the bond.

Instead, the price is determined by supply and demand. The price can be different each time the bond is traded. Supply and demand can take different forms depending on how markets are organized. For example, there can be a single seller and a single buyer that negotiate over the price, many traders who buy and sell at a price that adjusts continuously to clear the market, or a single seller and many buyers with the price determined by an auction[1]

Bond Yields

The **interest rate** of a zero-coupon bond is the bond's rate of return until maturity. Denote the price of the bond at time t by B_t. Then the bond's interest rate at time t, denoted by R_t, is

$$R_t = \frac{\$1,000 - B_t}{B_t}. \tag{1.1}$$

[1]U.S. government bonds, for example, are typically sold at issuance through an auction. You can participate in these auctions yourself and buy U.S. government bonds directly from the U.S. Treasury at Treasury Direct.

The **yield to maturity** on a zero-coupon bond is the same as its interest rate; we use both terms interchangeably. Yield to maturity is often shortened to **yield.**[2]

Equation (1.1) can be solved for the bond's price B_t to get

$$B_t = \frac{\$1,000}{1 + R_t}. \tag{1.2}$$

Equations (1.1) and (1.2) allow us to convert between the bond's price B_t and yield R_t: prices and yields convey the same information. A higher price is associated with a lower yield, and a lower price with a higher yield. When we talk about bond yields, we are therefore just talking about the bond's price.

Bond yields are always expressed in annualized terms. To annualize a yield, we divide the raw (non-annualized) interest rate R_t by the maturity of the bond, with maturity expressed in years. Denote the maturity of the bond at time t expressed in years by m_t. Then the annualized yield is[3]

$$\text{Annualized } R_t = \frac{R_t}{m_t}.$$

For coupon bonds, the **yield to maturity** is the bond's return if held to maturity, assuming all coupon and principal payments are made as scheduled.

[2]A bond's **current yield** is the ratio of the coupon payment to the current price of the bond. The current yield is a different concept from yield to maturity. If you see the word yield by itself, it can mean current yield or yield to maturity depending on the context.

[3]This definition assumes interest compounds continuously. An alternative choice is to compound period by period, in which case the annualized yield $R_t^{\text{annualized}}$ is defined by

$$(1 + R_t^{\text{annualized}})^{m_t} = (1 + R_t).$$

It is always possible to translate interest rates that are annualized using continuous compounding to interest rates that are annualized by using period-by-period compounding. Both measure the same economic quantity in different units. Annualized yields using continuous and period-by-period compounding are numerically close to each other—and hence essentially interchangeable without converting units—when $R_t^{\text{annualized}}$ is not too big (say, 10%-15% or less). In these notes, we use continuous compounding because calculations are easier.

Takeaway

- Bond prices and yields convey the same information.
- An increase in the price of a bond is equivalent to a decrease in the yield of the bond (and vice-versa).

Risk-free bonds

Bonds that never default are called **risk-free** or **riskless**. Of course, no bond is truly default-free. However, U.S. government bonds are generally considered very close to default-free, so treating them as riskless is a reasonable approximation in many economic applications.

The **risk-free rate** is the yield of a riskless zero-coupon bond.

Price Risk

Let R_{mt} be the annualized yield at time t for a zero-coupon bond with maturity m. At time t, R_{mt} is known. After all, knowing R_{mt} is the same as knowing the bond's price and, if we can buy the bond, then its price must be known!

In contrast, the future yield $R_{m,t+1}$ is *not* known at t; it is first known at time $t+1$. However, at time t we can form some expectation—a best guess—regarding what $R_{m,t+1}$ will be. We denote this **expected yield** or **expected interest rate** by $R^e_{m,t+1}$. The expected yield $R^e_{m,t+1}$ *is* known at t.

Investing \$1,000 in a riskless two-year-maturity zero coupon bond at time t has payoffs of \$0 at $t+1$ and

$$\$1000 \times (1 + 2R_{2t}) \tag{1.3}$$

at $t+2$. The number 2 in front of R_{2t} is there because R_{2t} is expressed in annualized terms.

Investing \$1,000 in a riskless *one*-year-maturity zero coupon bond at time t has a payoff equal to

$$\text{payoff}_{t+1} = \$1000 \times (1 + R_{1t})$$

at $t+1$. Reinvesting payoff_{t+1} at $t+1$ on new riskless one-year-maturity zero coupon bond (that matures at time $t+2$) results in a payoff equal to

$$\begin{aligned} \text{payoff}_{t+2} &= \text{payoff}_{t+1} \times (1 + R_{1,t+1}) \\ &= \$1000 \times (1 + R_{1t}) \times (1 + R_{1,t+1}) \end{aligned} \tag{1.4}$$

at $t+2$. The *expected* payoff at $t+2$ is the same as the actual payoff in Equation (1.4) but replacing the actual or **realized yield** $R_{1,t+1}$ by the expected yield $R^e_{1,t+1}$.

Reinvesting the principal payoff of a bond into a new bond of the same maturity is called **rolling over** the bond. Equation (1.4) gives the payoff of rolling over a one-year-maturity bond for two years.

Buying the two-year bond at time t entails no risk: the payoff in Equation (1.3) is known at t. In contrast, rolling over a one-year bond for two years has a risky payoff because the future yield $R_{1,t+1}$ is not known at t. This kind of risk is called **price risk**. It is different from default risk, which is the risk that bond payments are not honored. Confusingly, risk-free bonds can have price risk.[4]

Takeaway

- There are two kinds of risks for bonds: default risk and price risk.
- Default risk is the risk that payments are not honored.
- Price risk is the risk that bond yields change in the future.

Bond Risk Premia

Equilibrium in Bond Markets Determines Risk Premia Investors do not like risk. If buying the two-year bond and rolling over the one-year bond twice had the same expected payoffs, investors would always prefer to buy the two-year bond. No one would want to hold the existing supply of one-year-maturity bonds, preventing the market for one-year bonds from being in equilibrium.

For the bond market to be in equilibrium, rolling over the one-year bond twice must have a

[4]Economics: Why make it easy when we can make it harder? Don't get me started about *heteroskedasticity*.

higher expected payoff than buying the two-year bond. The difference in expected payoffs between the two options is a **risk premium**. The risk premium is the compensation for taking on the price risk associated with rolling over the one-year bond. In the context of the yield curve, the risk premium is also referred to as the **term premium**.

Mathematically, equilibrium in the bond market requires that

$$\$1,000 \times (1 + 2R_{2t}) = \$1,000 \times (1 + R_{1t})(1 + R^e_{1,t+1}) + x_2. \tag{1.5}$$

The left-hand side of Equation (1.5) is the payoff from investing in the two-year bond, taken from Equation (1.3). The right-hand side is the expected payoff from rolling over the one-year bond twice (Equation (1.4) with $R_{m,t+1}$ replaced by $R^e_{m,t+1}$), plus the risk premium x_2.

Equation (1.5) can be simplified by ignoring a small cross-term. The product on the right-hand side can be expanded as:

$$(1 + R_{1t})(1 + R^e_{1,t+1}) = 1 + R_{1t} + R^e_{1,t+1} + R_{1t}R^e_{1,t+1}.$$

If R_{1t} and $R^e_{1,t+1}$ are not too big (say, 10% or less), then the term $R_{1t}R^e_{1,t+1}$ is much smaller than the other terms. After canceling the \$1,000 from both sides, Equation (1.5) can thus be approximated by ignoring the small term $R_{1t}R^e_{1,t+1}$, which results in

$$1 + 2R_{2t} = 1 + R_{1t} + R^e_{1,t+1} + x_2,$$

or

$$R_{2t} = \frac{1}{2}(R_{1t} + R^e_{1,t+1} + x_2). \tag{1.6}$$

In words, Equation (1.6) says that the two-year yield is the average of the future path of one-year yields, plus a risk premium.

Comparing an m-year bond to rolling over one-year-maturity bonds for m periods gives the general formula:

$$R_{mt} = \frac{1}{m}(R_{1t} + R^e_{1,t+1} + ... + R^e_{1,t+m-1} + x_m). \tag{1.7}$$

Once again, the m-year yield is the average of the future expected one-year yields, plus a risk premium.

Takeaway

- Equilibrium requires that the yield of a long-maturity bond equals the average of future expected one-year yields, plus a risk premium.

Nominal and Real Interest Rates

The interest rates and yields we have considered thus far are all **nominal** because the bond payments are denominated in dollars.

The **real interest rate**, also called the **real yield**, is defined as the nominal interest rate minus expected inflation:

$$r_{mt} = R_{mt} - \pi^e_{mt}, \tag{1.8}$$

where r_{mt} is the real interest rate (or real yield), R_{mt} is the nominal interest rate (or nominal yield), and π^e_{mt} is the inflation rate expected to prevail between time t and time $t+m$.

The real yield is the yield on a **real bond**, which has payments specified not in terms of dollars but in terms of the representative basket of goods and services of the economy. Investing an amount equal to one basket of goods in zero-coupon real bonds with maturity m at time t has a promised payoff of $(1+r_{mt})$ baskets of goods at $t+m$.

Combining the definition of the real yield in Equation (1.8) with Equation (1.7) gives:

$$\begin{aligned} R_{mt} &= \frac{1}{m}\Big[(r_{1t} + \pi^e_{1t}) + (r^e_{1,t+1} + \pi^e_{1,t+1}) + \ldots + (r^e_{1,t+m-1} + \pi^e_{1,t+m-1}) + x_m\Big] \\ &= \frac{1}{m}\Big[(r_{1t} + r^e_{1,t+1} + \ldots + r^e_{1,t+m-1}) + (\pi^e_{1t} + \pi^e_{1,t+1} + \ldots + \pi^e_{1,t+m-1}) + x_m\Big]. \end{aligned}$$

This last equation expresses the nominal m-year yield as the average of the expected one-year *real* yields, plus the average expected inflation, plus the risk premium. It is a useful decomposition of long-term nominal yields into three parts: one from real yields, another from inflation, and a third from risk premia.

1.2 Stocks

Stocks are financial assets. Like all other financial assets, they are defined by the future payoffs they promise.

Stocks are **issued**—created and first sold—by corporations, which we refer to as firms for short. A stock promises its holder payoffs equal to a fixed fraction of the lifetime profits of the firm. The size of the fraction is determined by the total number of stocks issued. For example, if a firm has issued 100 stocks, each stock promises to pay out 1% of the firm's lifetime profits. The payments that stock owners receive are called **dividends**.

Legally, each stock represents the ownership of a fraction of the firm. Because stocks represent shares of ownership, they are also called **shares** and the people who own them are **shareholders**.

The promise to pay a share of *lifetime* profits does not mean that all profits must be paid out to shareholders immediately after profits are made. Indeed, even though all profits will ultimately be paid as dividends, stocks make no promises about when dividends will be paid or how much will be paid each time. Dividend decisions are made by each firm's board of directors, who evaluate whether reinvesting profits or distributing them as dividends will generate greater value for shareholders.

Start-ups and fast-growing firms don't usually pay dividends. Many large, established companies pay dividends every quarter, typically a fixed percentage of the profits earned since the last dividend. Google, founded in 1998, did not pay dividends (through its parent company, Alphabet) until 2024.

References

Blanchard, Olivier. 2024. *Macroeconomics.* 9th ed. New York, NY: Pearson, July. ISBN: 978-0138119010. https://www.pearson.com/en-us/subject-catalog/p/macroeconomics/P200000010516/9780135343340.

Saros, Daniel E. 2019. *Principles of Political Economy: A Pluralistic Approach to Economic Theory.* Routledge. ISBN: 9781138039186.

Woodford, Michael. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press. ISBN: 9780691010496.

Chapter III

Intermediate Macro

1 A Closed Economy Model[1]

1.1 National Income Accounts

The measure of aggregate output in the national income accounts is **gross domestic product**, or GDP. We can measure GDP in three different but equivalent ways:

1. From the production side: GDP equals the value of the final goods and services produced in the economy during a given period. The important word here is final. We want to count only the production of final goods, not intermediate goods.

2. Also from the production side: GDP is the sum of value added in the economy during a given period. The value added by a firm is defined as the value of its production minus the value of the intermediate goods used in production.

3. From the income side: GDP is the sum of incomes in the economy during a given period.

The three ways to measure GDP give the same number.[2] We use Y to refer to GDP. Depending on the context, we refer to Y as output, production, value added, or income.

[1]This document based on Blanchard (2024).
[2]We ignore measurement error.

1.2 The Market for Goods and Services

Goods and Services

The economy produces many goods and services. We are only interested in aggregate production, so we construct a representative basket of goods and services and treat that basket as the single aggregate good in the economy. Output Y is the number of baskets of goods produced.

The Price Level

The price in dollars of one basket of goods is denoted by P. We refer to P as **the price level**.

Equilibrium in the Goods Market

Demand is given by:

$$Z \equiv C + I + G, \tag{1.1}$$

where Z is demand; C is aggregate consumption; I is investment in physical capital; and G is government spending.

Consumption C is an increasing function of **disposable income** $Y_D \equiv Y - T$:

$$C = C(\underset{(+)}{Y_D}), \tag{1.2}$$

where Y is income and T are taxes. The plus sign below Y_D signifies that the function $C(Y_D)$ is increasing in Y_D. We note that we have abused notation in equation (3.8): the symbol C on the left-hand side is a number while the symbol C on the right-hand side is a function. We will use the same abuse of notation throughout.

Investment I is a decreasing function of the **real interest rate,** r and an increasing function of income Y:

$$I = I(\underset{(-)}{r}, \underset{(+)}{Y}).$$

From now on, we ignore the dependence of investment on Y and simply write:

$$I = I(\underset{(-)}{r}).$$

We highlight that investment I is *not* investment in financial assets. It is investment in *physical capital* such as machines and buildings. Similarly, when we use the word capital in this context, we always mean *physical* capital, not *financial* capital (money invested in some financial assets).

Using $C = C(Y - T)$ and $I = I(r)$ in equation (1.1) gives:

$$Z = C(Y - T) + I(r) + G, \tag{1.3}$$

so Z is an increasing function of Y and G and a decreasing function of T and r.

The **supply of goods** is simply domestic production Y.

The equilibrium condition in the goods market is that the supply goods Y equals the demand for goods Z. We can write this as:

$$\underbrace{Y}_{\text{supply}} = \underbrace{C(Y - T) + I(r) + G}_{\text{demand}}. \tag{1.4}$$

We think of this equation as determining the equilibrium level of Y given T and r and G.

Table 1.1 lists the exogenous variables of the model in alphabetical order. Table 1.2 lists the endogenous variables in alphabetical order, together with the equations that determine them.

Goods market: exogenous variables

Variable	Description
G	government spending
r	real interest rate
T	taxes

Table 1.1

Goods market: endogenous variables and equations

Variable	Description	Equation	Type of equation
C	consumption	$C = C(\underset{(+)}{Y_D})$	Behavioral
I	investment	$I = I(\underset{(-)}{r})$	Behavioral
Y	income, production	$Y = C + I + G$	Equilibrium condition
Y_D	disposable income	$Y_D \equiv Y - T$	Identity
Z	demand for goods	$Z \equiv C + I + G$	Identity

Table 1.2

Example 1.1. We use the following functions:

$$C(Y - T) = 1 + \frac{1}{2}(Y - T), \tag{1.5}$$

$$I(r) = 2 - r. \tag{1.6}$$

Our goal is to solve the model, that is, express all endogenous variables in terms of exogenous variables only.

Plugging equations (1.5) and (1.6) into equation (3.3) gives:

$$Y = \underbrace{1 + \frac{1}{2}(Y - T)}_{C(Y-T)} + \underbrace{2 - r}_{I(r)} + G.$$

Solving for Y gives Y in terms of exogenous variables:

$$Y = \frac{1}{1 - \frac{1}{2}}\left(1 + 2 - \frac{1}{2}T - r + G\right) = 2\left(3 - \frac{1}{2}T - r + G\right). \tag{1.7}$$

The rest of the endogenous variables can be found by plugging in equation (1.7) into the equations from Table 1.2. □

Saving

Private saving, denoted by S^p is defined as the part of disposable income that is saved rather than consumed:

$$S^p \equiv Y_D - C.$$

Government saving, denoted S^g is defined similarly to private saving. The government's "income" is its tax revenue, T while its "consumption" is government spending:

$$S^g \equiv T - G.$$

National saving, sometimes called aggregate saving or just **saving**, is denoted by S and defined by:

$$S \equiv S^p + S^g.$$

The two types of saving we have defined, private and government, add up to national saving, denoted by S

$$S = S^p + S^g.$$

Using equation (3.3) and the equations for saving above, we get

$$I = S. \tag{1.8}$$

In a closed economy investment is always equal to saving. Equation (1.8) is where the name "IS" of the IS curve comes from.

The government's **budget deficit** is defined as $G - T$ which also equals the negative of government saving $G - T = -S^g$. Using the equations for saving given above, equation (1.8) can be written as:

$$S^p = I + (G - T). \tag{1.9}$$

Equation (1.9) states that private saving can take one of two forms: investment in physical capital I and purchases of government debt $G - T$. Investing in physical capital is saving because it involves foregoing some consumption today (equal to the amount invested) in order to have more consumption in the future (equal to the additional consumption that can be produced in the future with the additional physical capital).

When the budget deficit is positive the government is spending more than its income because $G > T$. Where does it get the difference? It gets it by borrowing from the private sector using government bonds. The private sector purchases the government bonds today, and the government repays bondholders in the future. Thus, budget deficits act as saving for the private sector, who consumes less today in order to have resources to buy government bonds, in exchange for higher future consumption when bonds pay out.

1.3 The Money Market

Real money demand is denoted by M^d/P. We assume the real money demand is given by:

$$\frac{M^d}{P} = \mathcal{L}(\underset{(-)}{i}, \underset{(+)}{Y}),$$

where $\mathcal{L}$ is a function that is decreasing in the **nominal interest rate** i (indicated by the minus sign below the i), and increasing in income Y (indicated by the plus sign under the Y). The **real money supply** is denoted by M^s/P.

The equilibrium level of real money, also called the **real money stock**, or the **real money balance**, is denoted by M/P. The real money stock is determined by the equilibrium condition in the money market that real money supply is equal to real money demand, $M^s/P = M^d/P$:

$$\frac{M^s}{P} = \mathcal{L}(i, Y). \tag{1.10}$$

Two models for the money market

We consider two different models for the money market. The money demand function $M^d/P = \mathcal{L}(i, Y)$ and the equilibrium condition $M^s/P = M^d/P$ are the same in both models. What differs is how the central bank behaves. Table 1.3 summarizes the two models, which are presented in more detail next.

Money market: two models

Model	Exogenous variable	Endogenous variable	Money market equilibrium
1	M^s	i	$M^s/P = \mathcal{L}(i, Y)$ determines i
2	i	M^s	$M^s/P = \mathcal{L}(i, Y)$ determines M^s/P

Table 1.3

Model 1: Exogenous money supply In this model, we assume that the real money supply M^s/P is exogenous while the nominal interest rate i is endogenous. We interpret this model as the central bank picking the money supply. The equilibrium condition $M^s/P = M^d/P$ determines the endogenous money demand and the behavioral equation $M^d/P = \mathcal{L}(i, Y)$ then determines the nominal interest rate.

This model is sometimes called the "old LM" or "classical LM" model. The idea is that the central bank's defining property is the ability to "print money". When the central bank prints money and uses it to buy bonds, money supply increases. In equilibrium, money demand must equal money supply, so money demand must increase. For money demand to increase, the nominal interest rate must go down so that people find bonds less attractive, prompting them to sell bonds in exchange for money. In the new equilibrium, the central bank holds more bonds, and the private sector holds more money.

Table 1.4 lists the exogenous variables for model 1. Table 1.5 lists the endogenous variables for model 1, together with the equations that determine them.

Money market Model 1: exogenous variables

Variable	Description
M^s/P	real money supply
Y	income

Table 1.4

Money market Model 1: endogenous variables and equations

Variable	Description	Equation	Type of equation
M^d/P	real money demand	$M^d/P = \mathcal{L}(\underset{(-)}{i}, \underset{(+)}{Y})$	Behavioral
i	nominal interest rate	$M^s/P = M^d/P$	Equilibrium condition

Table 1.5

Model 2: Exogenous nominal interest rate This model assumes that the nominal interest rate i is the exogenous variable, while the real money supply is determined endogenously. Now, instead of picking the real money supply, the central bank picks the nominal interest. Given the exogenous nominal interest rate i and the exogenous level of income Y; the real money demand is determined by $M^d/P = \mathcal{L}(i, Y)$. The real money supply is then determined by the equilibrium condition $M^s/P = M^d/P$.

This model is sometimes called the "new LM" model. Just as in Model 1, this model also takes into account the central bank's ability to print money. However, it also recognizes that the central bank has a monopoly on money creation, which allows it to pick not only the quantity of money (the money supply) but also the price of money (the nominal interest rate). Operationally, the way that the central bank "picks" the nominal interest rate is by allowing people to deposit money in central bank accounts—in the same way that you deposit money in your own checking account at your bank—that earn the interest rate chosen by the central bank. The central bank can pay any interest it chooses, since it can always print any amount of money needed to do so.

Table 1.6 lists the exogenous variables for model 2. Table 1.7 lists the endogenous variables for model 2, together with the equations that determine them.

Example 1.2. We use the following function for $\mathcal{L}$:

$$\mathcal{L}(i, Y) = 2 + Y - 0.2i. \tag{1.11}$$

The goal is to solve both money market models, that is, to express all the endogenous

Money market Model 2: exogenous variables

Variable	Description
i	nominal interest rate
P	price level
Y	income

Table 1.6

Money market Model 2: endogenous variables and equations

Variable	Description	Equation	Type of equation
M^d/P	real money demand	$M^d/P = \mathcal{L}(\underset{(-)}{i}, \underset{(+)}{Y})$	Behavioral
M^s/P	real money supply	$M^s/P = M^d/P$	Equilibrium condition

Table 1.7

variables in terms of exogenous variables only.

Model 1 The equilibrium condition for the money market is that real money demand M^d/P equals real money supply M^s/P. Since real money supply is exogenous, and $M^d/P = M^s/P$ already expresses money demand only as a function of exogenous variables.

Now we find i in terms of exogenous variables. Plugging equation (1.11) into equation (1.10) gives:

$$\frac{M^s}{P} = 2 + Y - 0.2i.$$

Solving for i gives:

$$i = 10 - 5\frac{M^s}{P} + 5Y. \tag{1.12}$$

Model 2 Now the nominal interest rate is exogenous. Real money demand in terms of exogenous variables is immediately given by $M^d/P = \mathcal{L}(i, Y) = 2 + Y - 0.2i$. Equating this real money demand to real money supply gives

$$\frac{M^s}{P} = 2 + Y - 0.2i,$$

which expresses the real money supply as a function of exogenous variables only. □

1.4 The IS-LM Model

The Expected Price Level

The **expected price level** denoted by P^e is the price level that we expect will prevail in some future period. You can think of it as our best guess of what the actual price level will be once we get to that future period. For example, let's use t to denote the current period (the present). In the current period, the price level is P_t. This current price level P_t is known at time t since we construct it by looking at currently available prices. The next period $t + 1$ is in the future. The actual or **realized** price level P_{t+1} that will prevail at $t + 1$ is not known at t. The future price level P_{t+1} will only be known at $t + 1$. However, in the present (at time t), we do have some expectations of what P_{t+1} will be, which is the expected price level P^e_{t+1}. We form these expectations at time t. Therefore, unlike the realized future price level P_{t+1}; the future expected price level is known in the present, at time t.

1.5 Inflation

Inflation denoted by π is the percentage change in the price level P over some period:

$$\pi_t \equiv \frac{P_t}{P_{t-1}} - 1. \tag{1.13}$$

For example, if the period t is the year 2025 and the period $t-1$ is the year 2024; π_t is the inflation rate in 2025, the percentage change in the price level between 2024 and 2025.

Expected inflation denoted by π^e is the expected percentage change in the price level over some period that is, at least in part, in the future. For example, if t is the current period, expected inflation for the period $t+1$ is:

$$\pi^e_{t+1} = \frac{P^e_{t+1}}{P_t} - 1, \tag{1.14}$$

and expected inflation for $t+2$ is:

$$\pi^e_{t+2} = \frac{P^e_{t+2}}{P^e_{t+1}} - 1.$$

We follow the convention that whenever we omit the time index subscript for a variable, then that variable corresponds to the current period. For example, if t is the current period, then π and π_t both denote inflation in the current period. We also follow the convention that whenever we omit the time index subscript for the expectation of a variable, the expectation is about the value of the variable in the next period. For example, if t is the current period, then π^e is the same as π^e_{t+1}. When necessary, we disambiguate by just stating in words what period is the current one, and what future period expectations refer to.

The Fisher Equation

The real interest rate r and the nominal interest rate i are related by:

$$r = i - \pi^e, \tag{1.15}$$

where π^e is expected inflation. Equation (1.15) is called the **Fisher equation**.

The IS Curve

Plugging the Fisher equation (1.15) into the equilibrium condition for the goods market in equation (3.3) gives the **IS relation**:

$$Y = C(Y - T) + I(i - \pi^e) + G. \tag{1.16}$$

The **IS curve** is the set of points (Y, i) that satisfy equation (1.16) when we plot them with values of Y in the horizontal axis and values of i in the vertical axis. Figure 1.1 shows a generic IS curve. The IS curve is always downward sloping. Despite its name, the IS curve is not necessarily curve and can certainly be a (downward sloping) straight line.

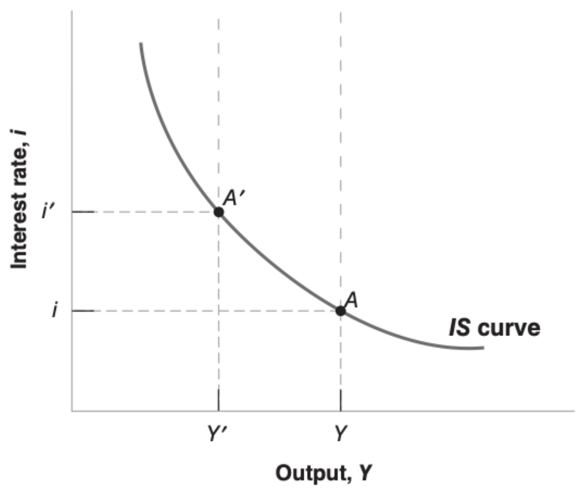


Figure 1.1: The IS curve

Example 1.3. The goal is to derive the IS curve. Using equation (1.7) and the Fisher equation, we get:

$$Y = 2\left(3 - \frac{1}{2}T - i + \pi^e + G\right).$$

Solving for i gives:

$$i = \underbrace{3 - \frac{1}{2}T + \pi^e + G}_{\text{intercept}} + \underbrace{\left(-\frac{1}{2}\right)}_{\text{slope}} Y, \tag{1.17}$$

which we recognize as the equation for a line with intercept $3 - \frac{1}{2}T + \pi^e + G$ and slope $-1/2$. The slope of this line is negative. We note that the only reason to solve for i explicitly is to make plotting the IS curve easier. We still think of the IS as determining Y for given values of i and of all other exogenous variables. When T or π^e or G change, the intercept of the IS changes. In this case, we say the IS shifts. If the intercept goes down, the IS shifts to the left; if the intercept goes up, the IS shifts to the right. We say the IS shifts left and right, rather than up and down, because we are thinking of how Y changes for given values of i (and a given value of i is a horizontal line, so higher or lower Y for that given value of i is the same as moving left or right along that horizontal line). □

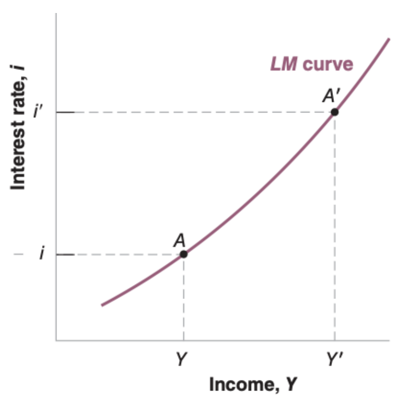
The LM Curve

Equation (1.10) is called the **LM relation** (where the L is for "liquidity" and the M is for "money").

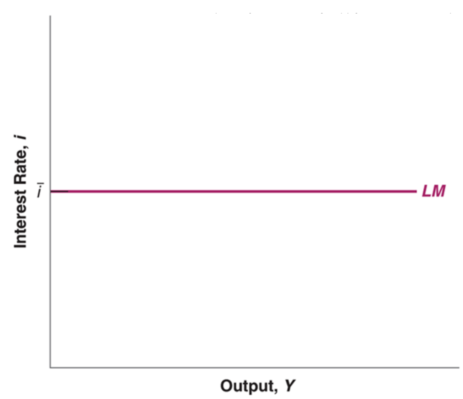
The **LM curve** is the set of points (Y, i) that satisfy equation (1.10) when we plot them with values of Y in the horizontal axis and values of i in the vertical axis. Despite its name, the LM curve is not necessarily curve and can certainly be a straight line.

In model 1, the LM curve is upward sloping because an increase in Y increases money demand and, with M^s exogenous, equilibrium requires a higher i. Figure 1.2(a) shows a generic upward sloping model 1 LM curve.

In model 2, the LM curve is a horizontal line because i is exogenous, and hence does not change with Y. Figure 1.2(b) shows a model 2 flat LM curve.



(a) Model 1: Upward sloping



(b) Model 2: Flat

Figure 1.2: The LM curve

Example 1.4. The goal is to derive the LM curve.

Model 1 If we use the function L from equation (1.11), we already found in the last section that:

$$i = \underbrace{10 - 5\frac{M^s}{P}}_{\text{intercept}} + \underbrace{5}_{\text{slope}} Y, \tag{1.18}$$

which we recognize as the equation for a line with intercept $(10 - 5M^s/P)$ and slope 5. When any of M^s or P change, the intercept of the LM changes. In this case, we say the LM shifts. If the intercept goes up, the LM shifts up; if the intercept goes down, the LM shifts down. We say the LM shifts up and down, rather than left and right, because we are thinking of how i changes for given values of Y.

Model 2 The equation for the LM curve is

$$i = \bar{\imath}, \tag{1.19}$$

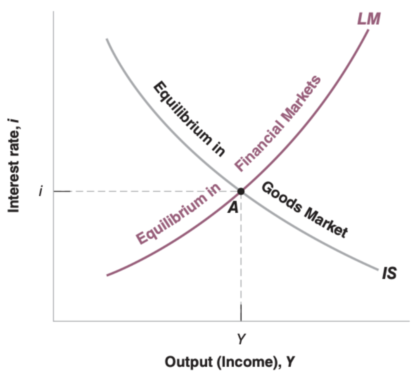
where $\bar{\imath}$ is the exogenous level of the nominal interest rate. □

Combining IS and LM

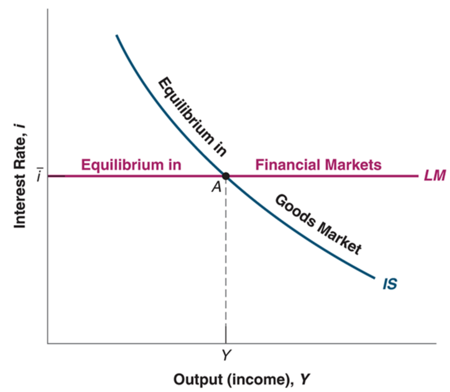
Points in the IS curve are combinations of Y and i that are consistent with equilibrium in the goods market. Points in the LM curve are combinations of Y and i that are consistent with equilibrium in the money market. The unique pair (Y, i) at the intersection of the IS and LM curves are therefore the values of Y and i that are consistent with equilibria in both the goods and money markets. Figure 1.3(a) shows the IS and LM curves together in one graph for model 1, and Figure 1.3(b) shows the same for model 2.

Just as we thought of the IS as determining Y given i and of the LM as determining i given Y. We now think of Y and i as being jointly determined by the IS and the LM—that is, by equilibrium in both the goods and money markets.

Table 1.8 lists the exogenous variables of the IS-LM model for model 1 in Table 1.3. Table 1.9 lists the endogenous variables for model 1, together with the equations that determine them.



(a) Model 1



(b) Model 2

Figure 1.3: IS-LM equilibrium

Tables 1.10 and 1.11 show the same but for model 2.

IS-LM Model 1: exogenous variables

Variable	Description
G	government spending
M^s/P	real money supply
T	taxes
π^e	expected inflation

Table 1.8

Example 1.5. **Model 1.** We start with the IS from equation (1.17) and the LM from equation (1.18):

$$\text{IS curve: } i = 3 - \frac{1}{2}T + \pi^e + G - \frac{1}{2}Y, \tag{1.20}$$

$$\text{LM curve: } i = 10 - 5\frac{M^s}{P} + 5Y. \tag{1.21}$$

We now find the equilibrium values of i and Y: that is, the values of i and Y at the intersection of the IS and the LM curves. First, equate the right-hand side of equation (1.20) to the right

IS-LM Model 1: endogenous variables and equations

Variable	Description	Equation	Type of equation
C	consumption	$C = C(\underset{(+)}{Y - T})$	Behavioral
i	nominal interest rate	$M^s/P = M^d/P$	Equilibrium condition
I	investment	$I = I(\underset{(-)}{i - \pi^e})$	Behavioral
M^d/P	real money demand	$M^d/P = \mathcal{L}(\underset{(-)}{i}, \underset{(+)}{Y})$	Behavioral
r	real interest rate	$r = i - \pi^e$	Identity
Y	income, production	$Y = C + I + G$	Equilibrium condition

Table 1.9

IS-LM Model 2: exogenous variables

Variable	Description
G	government spending
M^s/P	real money supply
T	taxes
π^e	expected inflation

Table 1.10

IS-LM Model 2: endogenous variables and equations

Variable	Description	Equation	Type of equation
C	consumption	$C = C(\underset{(+)}{Y - T})$	Behavioral
i	nominal interest rate	$M^s/P = M^d/P$	Equilibrium condition
I	investment	$I = I(\underset{(-)}{i - \pi^e})$	Behavioral
M^d/P	real money demand	$M^d/P = \mathcal{L}(\underset{(-)}{i}, \underset{(+)}{Y})$	Behavioral
r	real interest rate	$r = i - \pi^e$	Identity
Y	income, production	$Y = C + I + G$	Equilibrium condition

Table 1.11

hand side of equation (1.21):

$$3 - \frac{1}{2}T + \pi^e + G - \frac{1}{2}Y = 10 - 5\frac{M^s}{P} + 5Y.$$

Second, solve for Y to get:

$$Y = -\frac{14}{11} + \frac{10}{11}\frac{M^s}{P} - \frac{1}{11}T + \frac{2}{11}\pi^e + \frac{2}{11}G. \tag{1.22}$$

Last, plug Y into equation (1.20) or equation (1.21) and re-arrange to get the equilibrium value of i:

$$i = \frac{5}{11}\left(8 - \frac{M^s}{P} - T + 2\pi^e + 2G\right). \tag{1.23}$$

Model 2. The IS is the same as in model 1. The LM given in equation (1.19):

$$\text{IS curve: } i = 3 - \frac{1}{2}T + \pi^e + G - \frac{1}{2}Y, \tag{1.24}$$

$$\text{LM curve: } i = \bar{\imath}. \tag{1.25}$$

The equilibrium nominal interest rate is already expressed in terms of the exogenous variable $\bar{\imath}$. Plugging $i = \bar{\imath}$ into the IS curve and solving for Y gives Y in terms of exogenous variables only:

$$Y = 6 - T + 2\pi^e + 2G - 2i.$$

□

1.6 The Labor Market

Unemployment

Let L be the economy's **workforce** or **labor force** and N be the number of people in the workforce who are employed. Then the **unemployment rate** is defined by:

$$u \equiv \frac{L - N}{L}. \tag{1.26}$$

Production Function

The only factor of production is labor[3]. The aggregate **production function** is:

$$Y = N. \tag{1.27}$$

Price-setting

Since the only factor of production is labor, the nominal marginal cost for firms producing goods and services is the **nominal wage** W. If markets were perfectly competitive, firms would set the price P equal to marginal cost W and make zero profits. In the real world, there is market power and firms tend to set prices above marginal cost. We capture this idea by the following **price-setting relation**:

$$P = (1 + m)W, \tag{1.28}$$

where m is the **markup**. This price setting relation says that firms set the price of the goods that they sell to always get a profit rate of m. To see that the markup m is equal to the

[3]This is not as restrictive as it sounds: we are assuming that labor is the only factor of production that adjusts within the time horizons we are considering and ignoring the other factors because they remain unchanged.

profit rate, start by defining the profit rate by

$$\frac{\text{profit}}{\text{cost of production}}.$$

Each unit sold costs W to produce and sells for P. The profit per unit is $P - W$. Hence the profit rate for one unit is

$$\frac{P-W}{W} = \frac{(1+m)W - W}{W} = m.$$

Selling more units raises profits (the numerator) and costs of production (the denominator) proportionally, leaving the profit *rate* unchanged.

There are many factors that can affect the markup, including:

- The cost of production. If the cost of production increases, producers may pass some or all of the cost increase to consumers in the form of higher markups. An important example is the price of oil. When the price of oil increases, costs of production increase even keeping the nominal wage constant.
- Monopoly power. If a firm's monopoly power increases, it can increase the price even keeping labor and other production costs unchanged. Many different channels can influence the degree of monopoly power. For example, making it more difficult for new firms to enter markets increases the monopoly power of incumbent firms. More stringent anti-trust regulation, in contrast, should reduce monopoly power.

Wage-setting

Wages are set according to the following **wage-setting relation**:

$$W = P^e F(\underset{(-)}{u}, \underset{(+)}{z}). \tag{1.29}$$

In this wage setting relation the variable P^e is the expected price level (the price level we expect to hold in the future). The variable u is the unemployment rate. The variable z is a catch-all variable that captures all other elements that influence wages other than P^e and u. The function F is decreasing in u and increasing in z. That the function F is increasing in z is just a convention. There are therefore three elements that determine the wage are

P^e and u and z. We briefly explain why they matter for wages:

- Expected price level. The expected price level P^e is an important determinant of the nominal wage because what people and firms really care about is not the nominal wage but the real wage. Workers value money because of the goods and services they can buy with it, not for its own sake. If my nominal wage is \$1,000,000 but the price of one apple is \$10,000,000, I am not so happy with my gigantic nominal wage. We use P^e rather than P because after wages are set, they usually remain fixed (or close to fixed) for some time, with wage contracts renegotiated only infrequently. So the relevant price level is the one that will prevail between now and the future time when the wage is renegotiated rather than just the current price level. goods and services that one can afford after W is set are better captured by W/P^e.

- Unemployment. The wage setting equation says that when unemployment is high, wages are low. The reasoning is that when unemployment is high, employers have more bargaining power, since someone seeking a job must compete with a larger pool of unemployed workers. Conversely, when the unemployment rate is low, it is workers who have higher bargaining power, as many firms have to compete to hire among the smaller pool of unemployed people. A conceptual weakness of the unemployment rate as an indicator of the relative bargaining power of workers and firms is that it is based solely on information about the status of workers (employed or unemployed) and does not directly incorporate information on the hiring needs of employers. But the labor market involves costly search by employers for workers as well as by workers for jobs. In principle at least, a given unemployment rate could be consistent with either a strong labor market, with upward pressure on wages, or a weak labor market and low wage pressure, depending on whether job openings are plentiful or scarce. An alternative indicator of bargaining power that takes into account the state of workers and of firms is the ratio of the **vacancy rate** (job listings reported by employers divided by the labor force) denoted by v to the unemployment rate u. The ratio v/u is called **labor market tightness**. However, the notion of labor market tightness is more general. In periods when firms want to hire a lot of workers, unemployment is relatively low, there is upward pressure on wages and many unfilled job vacancies, we say the labor market is **tight**. Conversely, when there are a lot of people looking for jobs, firms are not looking to hire as much, vacancies are filled quickly, and there is a downward pressure on wages, we say that the labor market is **slack**. Thus, conceptually, what we would really like to include in the wage-setting equation (1.2) is a good measure of overall labor market tightness rather than simply unemployment. Using v/u is one good option.

- Other factors. The other factors are defined so that the convention that F is increasing in z applies. Examples of factors that enter z are:
  - Unemployment insurance, the payment of unemployment benefits to workers who lose their jobs. If u and P^e are given then higher unemployment benefits make unemployment less painful, so employers must offer a higher wage to attract workers (in this case higher unemployment benefits are associated with a higher z).
  - The minimum wage. An increase in the minimum wage may increase not only the minimum wage itself, but also wages just above the minimum wage, leading to an increase in the economy-wide average wage W at a given unemployment rate.
  - Employment protection laws, which makes it more expensive for firms to lay off workers. Such a change is likely to increase the bargaining power of workers covered by this protection (laying them off and hiring other workers is now more costly for firms), increasing the wage for a given unemployment rate.
  - Worker's productivity. If workers can produce more output in the same amount of time, their labor is more valuable, which allows them to bid for higher wages.

Table 1.12 lists the exogenous variables of the labor market model. Table 1.13 lists the endogenous variables together with the equations that determine them.

Labor market: exogenous variables

Variable	Description
L	labor force
m	markup
P^e	expected price level
Y	production
z	catch-all variable for factors that affect the nominal wage

Table 1.12

Labor market: endogenous variables and equations

Variable	Description	Equation	Type of equation
N	employment	$Y = N$	Behavioral
P	price level	$P = (1+m)W$	Behavioral
u	unemployment rate	$u = \frac{L-N}{L}$	Identity
W	nominal wage	$W = P^e F(\underset{(-)}{u}, \underset{(+)}{z})$	Behavioral

Table 1.13

Example 1.6. One example for the function F is:

$$F(u, z) = 1 - \frac{1}{4}u + z. \tag{1.30}$$

The goal is to solve the model.

The production function (1.27) immediately gives N as a function of the exogenous Y. Combining equations (1.1) and (1.27) gives u in terms of exogenous variables only:

$$u = \frac{L - N}{L} = 1 - \frac{N}{L} = 1 - \frac{Y}{L}. \tag{1.31}$$

Plugging equation (1.31) into (1.30) gives:

$$F(Y, z) = 1 - \frac{1}{4}\left(1 - \frac{Y}{L}\right) + z. \tag{1.32}$$

Plugging equation (1.32) into equation (1.2) gives:

$$\frac{W}{P^e} = \frac{3}{4} + \frac{1}{4}\frac{Y}{L} + z.$$

Solving for W we get W in terms of exogenous variables only:

$$W = \left(\frac{3}{4} + \frac{1}{4}\frac{Y}{L} + z\right)P^e.$$

Last, plugging into equation (1.1) gives P in terms of exogenous variables only:

$$P = (1 + m)\left(\frac{3}{4} + \frac{1}{4}\frac{Y}{L} + z\right)P^e.$$

□

1.7 The Phillips Curve

Aggregate Supply

Combining (1.1) and (1.2) gives:

$$P = (1 + m)P^e F(u, z). \tag{1.33}$$

We can re-write equation (1.3) as a relation between P and Y instead of P and u. As in the last example, combining equations (1.1) and (1.27) gives equation (1.31). Equation (1.31) says that lower unemployment is associated with higher output because it takes more employed people to produce more output. Using (1.31) in (1.3) gives:

$$P = (1+m)P^e F\left(1 - \frac{Y}{L}, z\right). \tag{1.34}$$

Equation (1.34) is called the **aggregate supply relation**, or AS relation for short. We think of the AS relation as determining the price level P for a given value of Y.

Figure 1.4 shows a generic AS curve. Just like all the other "curves", the AS curve can be a straight line. The AS curve is always upward sloping. Changes in the markup m and the expected price level P^e and the labor force L and the catch-all variable z shift the AS curve up and down.

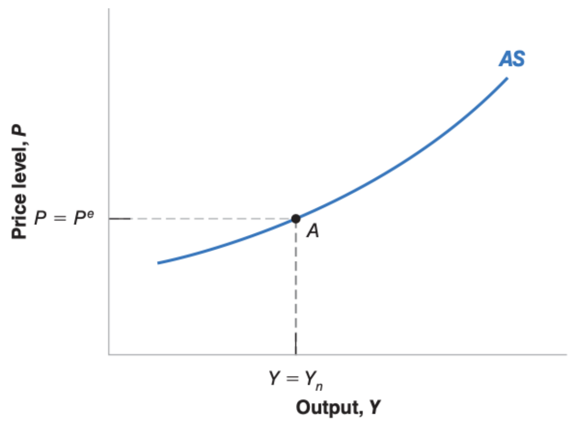


Figure 1.4: The AS curve

From now on, we assume a specific functional form for the function F:

$$F(u, z) = 1 - \alpha u + z, \tag{1.35}$$

where α is a positive parameter (a positive number taken as given). Plugging equation (1.31) into (1.35) gives:

$$F\left(1 - \frac{Y}{L}, z\right) = 1 - \alpha\left(1 - \frac{Y}{L}\right) + z. \tag{1.36}$$

Using equation (1.36) in equation (1.34) and re-arranging gives the AS curve:

$$P = \underbrace{(1+m)P^e(1-\alpha+z)}_{\text{intercept}} + \underbrace{\frac{(1+m)P^e\alpha}{L}}_{\text{slope}} Y, \tag{1.37}$$

which we recognize as the equation of a line with intercept $(1+m)P^e(1-\alpha+z)$ and slope $(1+m)P^e\alpha/L$. When any of m or P^e or α or z change, the intercept of the AS curve changes and the AS shifts up and down.

Linearization

The definitions of inflation and expected inflation in equations (1.13) and (1.14) imply that

$$P_t = (1+\pi_t)P_{t-1},$$
$$P_t^e = (1+\pi_t^e)P_{t-1}.$$

Adding time subscripts to equation (1.37) and plugging in the last two equations gives:

$$1+\pi_t = (1+m)(1+\pi_t^e)\left(1+z-\alpha+\frac{\alpha}{L}Y_t\right). \tag{1.38}$$

Distributing the product and subtracting 1 from both sides, we get

$$\pi_t = \pi_t^e + m + z - \alpha\left(1-\frac{Y_t}{L}\right) + \text{second order terms}, \tag{1.39}$$

where the second order terms are terms with cross-products of π_t^e and m and $u_t = 1-Y_t/L$. Since π_t^e and m and u_t are relatively small numbers (say, between 0 and 0.2), their cross products are small enough to be ignored and still get a good approximation. Ignoring these second order terms gives the **Phillips curve**:

$$\pi_t = \pi_t^e + m + z - \alpha + \frac{\alpha}{L}Y_t. \tag{1.40}$$

Figure 1.5 shows a generic Phillips curve (PC for short), where the vertical axis plots $\pi_t - \pi_t^e$ and the horizontal axis plots Y_t. Just like all the other "curves", the PC curve can be a straight line. The PC curve is always upward sloping. The variables π_t^e and m and z and α shift the PC up and down, while α and L change the slope of the PC.

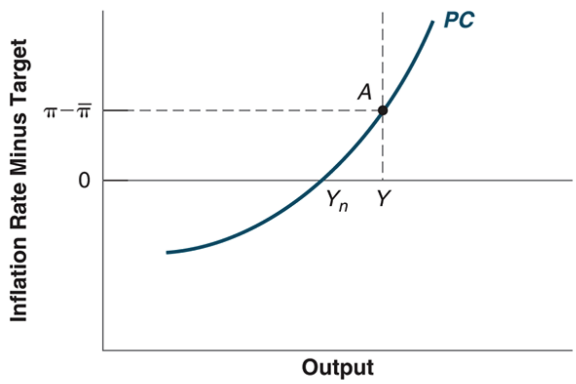


Figure 1.5: The Phillips curve

1.8 IS-LM-PC Model

The IS-LM-PC model combines equilibrium in the goods market (the IS relation), equilibrium in the money market (the LM relation), and equilibrium in the labor market (the Phillips curve, PC).

<u>The model</u>

Adding time subscripts, the IS relation in (1.16) becomes

$$Y_t = C(Y_t - T_t) + I(i_t - \pi_t^e) + G_t, \tag{1.41}$$

where we have used the Fisher equation (1.15) to write the real interest rate as $r_t = i_t - \pi_t^e$.

The LM relation in (1.10) with time subscripts becomes

$$\frac{M_t^s}{P_t} = \mathcal{L}(i_t, Y_t). \tag{1.42}$$

The PC in (1.40) remains unchanged.

Together with the definition of inflation in (1.13), equations (1.41), (1.42), and (1.40) determine the joint behavior of output Y_t and the nominal interest rate i_t and inflation π_t over time. Given an initial price level, the path of inflation can be used to infer the path of the price level P_t.

The natural rate of unemployment and potential output

The **natural rate of unemployment** denoted u^n is defined as the unemployment rate when inflation equals expected inflation, $\pi_t = \pi_t^e$. From (1.40) with $\pi_t = \pi_t^e$ we get

$$u^n \equiv \frac{m+z}{\alpha}. \tag{1.43}$$

The level of output when $\pi_t = \pi_t^e$ is called **potential output**, and is given by

$$Y^n \equiv L(1-u^n) = L\left(1-\frac{m+z}{\alpha}\right). \tag{1.44}$$

The difference $u_t - u^n$ is called the **unemployment gap** and $Y_t - Y^n$ is called the **output gap**.

Using equations (1.43) and (1.44), we can rewrite the Phillips curve in gap form:

$$\pi_t - \pi_t^e = -\alpha(u_t - u^n), \tag{1.45}$$

or, using the output gap,

$$\pi_t - \pi_t^e = \frac{\alpha}{L}(Y_t - Y^n). \tag{1.46}$$

Equations (1.45) and (1.46) say that unemployment below its natural level and output above potential are associated with inflation above expected inflation.

Expected inflation

Thus far, expected inflation has been taken as exogenous. A behavioral equation that makes expected inflation endogenous—a model of expected inflation—is

$$\pi_t^e = (1-\theta)\bar{\pi} + \theta\pi_{t-1}, \tag{1.47}$$

where θ is a parameter between 0 and 1, and $\bar{\pi}$ is another parameter that we refer to as the **inflation target**. We say that inflation expectations are **anchored** when $\theta = 0$ and **deanchored** when $\theta = 1$.

When $\theta = 1$ the Phillips curve (1.46) is called the "accelerationist" Phillips curve:

$$\pi_t - \pi_{t-1} = \frac{\alpha}{L}(Y_t - Y^n). \tag{1.48}$$

Equation (1.48) implies that whenever output is above its natural level, inflation increases over time, and whenever output is below its natural level, inflation decreases over time. In contrast, when $\theta = 0$ the Phillips curve is

$$\pi_t - \bar{\pi} = \frac{\alpha}{L}(Y_t - Y^n),$$

and inflation increases when the output gap $Y_t - Y^n$ increases, and decreases when the output gap decreases.

The difference between anchored and deanchored inflation expectations is of first-order importance for monetary policy. Consider a scenario in which the economy is in a boom, with unemployment below its natural rate and output above potential. The PC implies that, in this case, inflation is above expected inflation. Imagine monetary policy increases the interest rate over time and stops when unemployment is equal to its natural rate and output is equal to potential. With anchored expectations, the increasing interest rate makes inflation decrease over time as the output gap goes down. When output reaches potential, inflation equals the inflation target $\pi = \bar{\pi}$ so inflation ends up below its initial level. With deanchored expectations, inflation *increases* over time as output goes down, since the output gap is always positive along the path. When output reaches potential, inflation stops increasing. In contrast to the case of anchored expectations, inflation ends up above its initial level.

1.9 Sticky Prices

In the real world, **prices are sticky**: we do not see the price of individual goods changing instantaneously in response to evolving economic conditions, with prices of many individual goods unchanged for months at a time despite fluctuations in economic conditions. In addition, not all prices change at the same time. The stickiness of prices and the lack of synchronization in price changes across firms imply that the aggregate price level changes slowly over time.

Why are prices sticky? One reason is that changing prices is costly. The physical costs of changing prices are called **menu costs**, (think of a restaurant that has to print new menus when it changes prices). There are other reasons why prices are sticky: computing the new price and deciding to change prices can be time-consuming, especially in large companies; the necessary information about economic conditions may take time to get to firms and be digested; even when economic conditions change, if a firm's competitor does not change prices, it may be better to not change prices to remain competitive, etc.

1.10 The Short Run, the Medium Run, and the Long Run

We now want to think about how the economy evolves over time. It is useful to make the conceptual distinction between the short run, the medium run, and the long run.

- **The short run**: What happens to the economy from year to year. Movements in output are primarily driven by "movements in demand", that is, by equilibrium in the goods market. Firms are willing to supply any quantity at the given prevailing price.
- **The medium run**: What happens to the economy over a decade or so. In the medium run, the economy tends to return to the level of output determined by structural "supply factors", that is, by equilibrium in the labor market. Firms supply an amount of output equal to potential, and prices adjust to make the goods market be in equilibrium.
- **The long run**: What happens to the economy over a half century or longer. In the long run, what dominates is not fluctuations, but growth. Long-term growth is determined, among other things, by the growth rate of productivity, the education system, the saving rate, demographics, the rule of law, and the quality of institutions.

The short run, medium run, and long run can also be characterized by how prices behave:

- In the short run the price level P is fixed. In other words, P can be treated as exogenous in the short run. We can think of the short-run price level as being inherited from past decisions, and the short run as the time horizon over which changes in the sticky price level are small enough that they can be ignored.
- The medium run can be thought of as the amount of time it takes for all price changes to occur or, equivalently, for prices to behave as if they were not sticky at all. If prices

are not changing, then the price level must equal the expected price level and inflation must equal expected inflation. Therefore, the conditions $P = P^e$ and $\pi = \pi^e$ can be taken as the *definition* of the medium run.

- In the long run, we have **neutrality of money**: only real variables (rather than nominal ones) determine economic outcomes. The price level P and inflation π are irrelevant in the long run.

Medium-run adjustment

The Phillips curve (1.45) describes how inflation responds to the output gap. Suppose output is above its natural level $Y_t > Y^n$. Then (1.45) implies $\pi_t > \pi_t^e$ so the price level rises faster than expected. A higher price level reduces real money balances M_t^s/P_t; shifting the LM curve up in subsequent periods. This raises the nominal interest rate and lowers output, pushing output back toward Y^n.

Conversely, if output is below its natural level $Y_t < Y^n$ inflation is lower than expected. As the price level rises more slowly (or falls), real money balances increase, shifting the LM curve down and pushing output up toward Y^n.

In the medium run, once the economy reaches a situation in which output equals its natural level $Y_t = Y^n$; the Phillips curve implies that inflation equals expected inflation $\pi_t = \pi_t^e$. In that case, there is no systematic pressure for inflation to accelerate or decelerate.

References

Blanchard, Olivier. 2024. *Macroeconomics.* 9th ed. New York, NY: Pearson, July. ISBN: 978-0138119010. https://www.pearson.com/en-us/subject-catalog/p/macroeconomics/P200000010516/9780135343340.

Chapter IV

The II-XX Model

1 Internal and External Balance[1]

Section 19.1 of Chapter 19 in the textbook has a good discussion of internal vs external balance. If you have not yet read that section, you may want to do so before reading the rest of this note (even though this note is self-contained and should be understandable without having read Chapter 8).

Our initial understanding of internal vs external balance will require very few assumptions. In some sense, that is great: the fewer the assumptions, the more general we hope our model can be. On the other hand, we will be able to understand fewer things than if we had a more comprehensive model. After this first pass that requires few assumptions, we will return to the AA-DD and use it to reach sharper conclusions that have the same overall message.

Therefore, we now work under the following assumptions:

- There is a long-run level of output that coincides with the full-employment level Y^f. This means that, absent any changes, output Y will approach its full-employment level Y^f. Government policy can make the convergence of Y toward Y^f faster or slower,

[1]This chapter is based on: Roberto Rigobon, "15.014 Applied Macro- and International Economics II", Spring 2016. License: Creative Commons BY-NC-SA.

but it cannot change Y^f. In addition, given that Y^f is the "full-employment level" (whatever that means!), any deviations of Y from Y^f are undesirable for the people who live in this economy.

- There is a long-run level of the current account that is "sustainable", which we denote by XX^b. Sustainable means that, absent any changes, the current account CA:
  - will approach XX^b over time,
  - remains at XX^b once it gets there, and
  - has XX^b as a desirable value for reasons that we leave unspecified; perhaps sustainability itself is a worthy goal, or perhaps there is a good economic reason why people may prefer XX^b to other levels.

Given these three assumptions, an economy that has $Y = Y^f$ and $CA = XX^b$ is in its happy place. Our next task is to understand what combinations of output and exchange rates (Y, E) are compatible with $Y = Y^f$ and $CA = XX^b$ in the short run, that is, when the price level P is fixed.

1.1 The II Curve

In this section, we introduce the II curve. The II curve is the set of points (Y, E) that are consistent with full employment in the short run, i.e., the set of points that have $Y = Y^f$ and any E. As Figure 1.1 shows, the II curve is vertical.

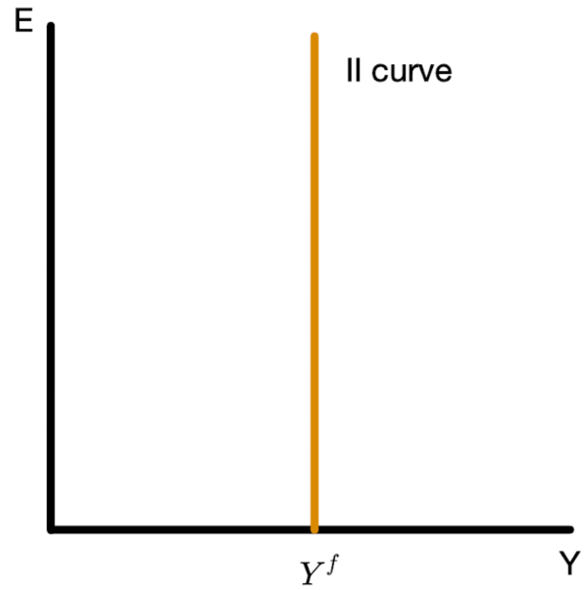


Figure 1.1:

1.2 Internal Balance

When the economy is at $Y = Y^f$, we say it is in "internal balance". At $Y = Y^f$, there is no reason for a policymaker to try to stimulate or cool down the economy in the short run. We have been referring to Y^f as the full-employment level, but there are closely related terms that describe the same idea: the "natural level of output", the "non-accelerating-inflation level of output", the "potential level of output". At $Y = Y^f$, the economy is operating at an efficient and non-inflationary level. Associated with this full-employment level of output is a "natural rate of unemployment", u^n, defined as the rate of unemployment that is consistent with $Y = Y^f$ and a constant price level P^f. Let L be the economy's workforce and N be the number of people in the workforce that are employed. Then the unemployment rate is defined by

$$u = \frac{L - N}{L}. \tag{1.1}$$

If we assume that the only factor of production is labor[2] and that the economy's aggregate production function is

$$Y = N,$$

then the natural rate of unemployment and the "full-employment" level of output are related by

$$u^n = 1 - \frac{Y^f}{L}.$$

Figure 1.2 shows a representation of the adjustment dynamics for the economy as time goes by. The arrows point toward the II line to signify that there is an economic force (as of now unspecified) that tends to move the economy toward Y^f.

When $Y > Y^f$, we describe the situation as one of "overheating". Compared to the more desirable state in which $Y = Y^f$, people are working overtime, inflation is too high, and other resources are being used "above capacity", which can result in higher overall output but with lower productivity and faster capital depreciation.

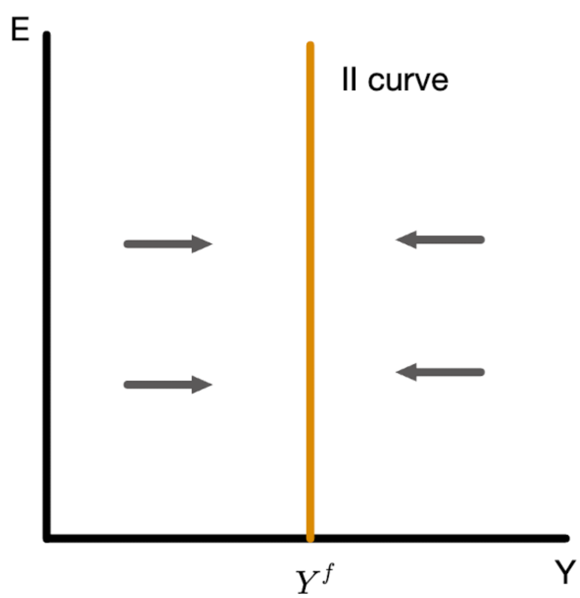


Figure 1.2:

[2]This is not as restrictive as it sounds: we are assuming that labor is the only factor of production that adjusts within the time horizons we are considering and ignoring the other factors because they remain unchanged.

When $Y < Y^f$, we describe the situation as "unemployment". People would like to work more, there is idle capacity or underutilization of resources, and inflation is too low.

1.3 The XX Curve

The XX curve is the set of points (Y, E) that keep the current account constant and equal to a given value XX in the short run (with P fixed):

$$XX = CA\left(\frac{EP^*}{P}, Y, Y^*\right) \tag{1.2}$$

We already encountered the XX curve in the context of the AA-DD model. When plotted in a diagram with Y in the horizontal axis and E on the vertical axis, it is increasing. We know it is increasing because CA is increasing in E and decreasing in Y (keeping P, P^*, Y^* constant).

We now want to have deeper economic understanding of why the XX is increasing. Figure 1.3 shows an arbitrary starting point (Y_0, E_0).

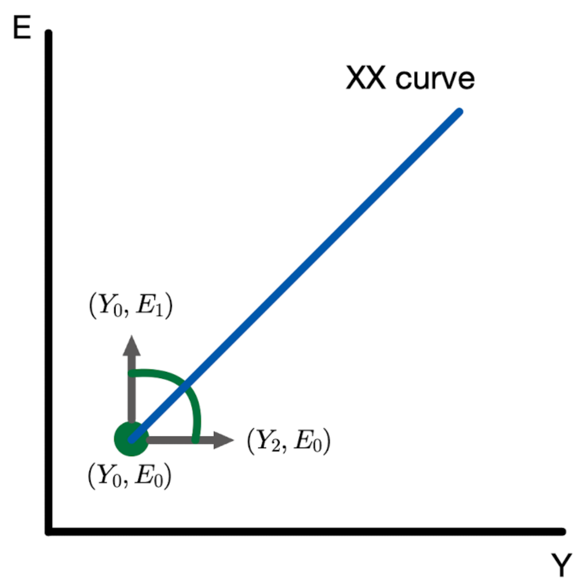


Figure 1.3:

The value of CA at that point is XX_0, so we have that:

$$XX_0 = CA\left(\frac{E_0P^*}{P}, Y_0, Y^*\right)$$

Imagine that there is a nominal depreciation that keeps output fixed: from (Y_0, E_0) we move to (Y_0, E_1) with $E_1 > E_0$. In the short run, P and P^* are fixed, so the nominal depreciation is also a real depreciation. Domestic goods have become cheaper relative to foreign goods. Imports go down, exports go up, and the current account improves (goes up). So when we start at (Y_0, E_0) and move north in the Figure 1.3, CA goes up.

Now imagine a different change that keeps the exchange rate fixed but increases output: from (Y_0, E_0) we move to (Y_2, E_0) with $Y_2 > Y_0$. Higher Y implies that there is a higher demand for foreign goods. Imports go up and the current account deteriorates (goes down). So when we start at (Y_0, E_0) and move east in Figure 1.3, CA goes down.

Going north we have a situation of higher CA, while going east leads to a lower CA. If we move from the point (Y_0, E_1) to the point (Y_2, E_0), we must cross a point of zero change in CA. Said differently, you can always find a combination of movements to the north and to the east that lead to no change in CA. This point of no change in CA to the northeast of (Y_0, E_0) must therefore be on the same XX curve as (Y_0, E_0), which in this case is the one that has $CA = XX_0$.

If we connect all the points that give us a zero change in CA, we trace the entire XX curve. We have found that movements along the XX curve keep CA unchanged because the improvement in CA due to a real depreciation (the movement to the north) is exactly offset by a deterioration in CA due to increased imports that were caused by higher income[3].

1.4 External Balance

We define external balance as the situation in which the current account is at its sustainable level XX^b. Figure 1.4 depicts the dynamics of convergence toward XX^b. First, the figure shows the XX curve with an associated value of the current account equal to XX^b. For any combination of points (Y, E) above this XX^b, the current account is higher than XX^b. In this case, the figure shows arrows that point down and toward the curve to signify dynamics that, over time, make the current account decrease toward XX^b. Similarly, for points below

[3]Remember that Y is both aggregate output and aggregate income. That aggregate output equals aggregate income is an accounting identity. If you don't recall this relation, go back to your notes from intermediate macro or our quick review of this issue in the lecture slides on National Accounts.

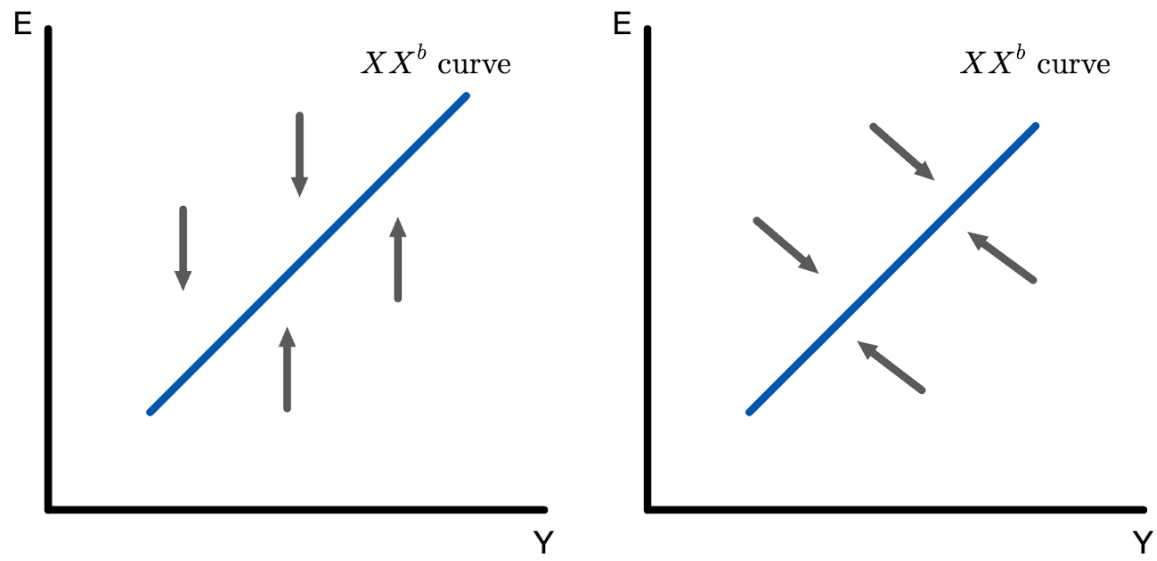


Figure 1.4:

the XX^b curve that have a current account lower than XX^b, the left panel of the figure shows arrows pointing up and toward the curve, to signify dynamics that make the current account increase toward XX^b. Arrows that point straight up and straight down are not the only possibility. Indeed, the adjustment toward XX^b does not have to occur only through the exchange rate and can occur through combinations of E and Y. The right panel of Figure 1.4 shows an example in which E and Y are both changing as the adjustment takes place. The arrows point toward the XX^b curve at some angle rather than vertically up and down, as in the right panel of the figure.

The lack of specificity about how the adjustment toward XX^b is one of the drawbacks of this simple model vis-a-vis a model with more structure such as the AA-DD.

When the current account is above XX^b, we say that the economy is in a situation of "surplus", while when the current account is lower than XX^b, we say that the economy is in a situation of "deficit". Here the word deficit refers to a *current account* deficit (relative to the level XX^b) and not a *fiscal* deficit. A fiscal deficit is characterized by government purchases larger than taxes, $G > T$, while the current account deficit is characterized by $CA < XX^b$. While current account and fiscal deficits are not the same thing, we do see in the real world that persistent domestic fiscal deficits are often sustained, at least in part, by borrowing from the rest of the world, which brings the current account down.[4]

1.5 The Four Zones of Economic Discomfort

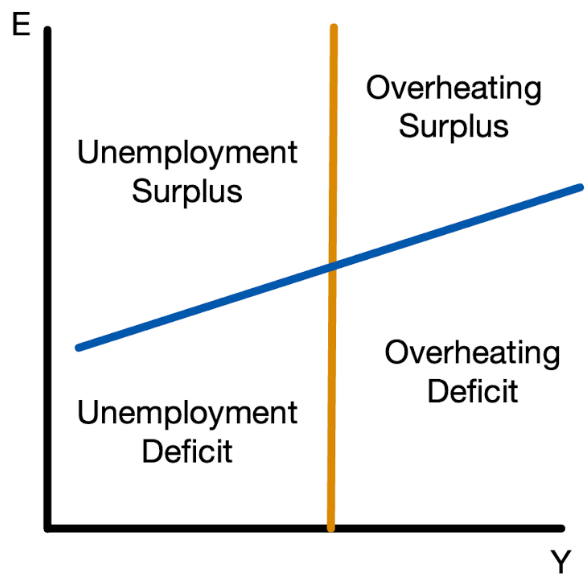


Figure 1.5:

Figure 1.5 shows the result of putting Figures 1.2 and 1.4 together. We can use this figure to diagnose the state of the economy belonging to one of the four "zones of economic discomfort". In each of these zones, we have an absence of internal of external balance of a different kind. For example, the south-east zone represents an overheating economy with current account deficits, while the north-east zone has overheating and surplus.

We can also combine the dynamics from Figures 1.2 and 1.4, as is done in Figure 1.6. An important question is how long it takes for an economy to trace the whole path in the real world.

Without any policy intervention, the answer is that it can take many years. Policy intervention, if appropriate, can speed up the convergence toward the long-run equilibrium. However, the effects of policy also take time to be felt in the economy. In the United States, monetary policy actions usually start to have an effect on output and the current account

[4] Recall that if S^p is private saving, $S^g \equiv T - G$ is public saving, and I is investment, then $S^p + S^g = I + CA$. For given S^p and I, higher public saving (equivalently, lower public deficit) is associated with a lower current account.

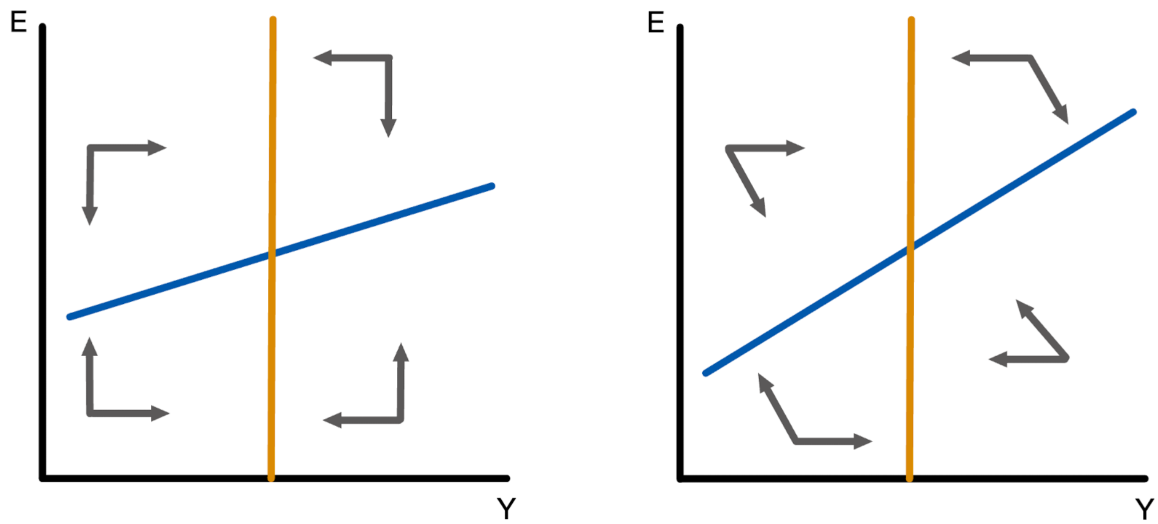


Figure 1.6:

after a few months, and have their largest effect between one and two years after the actions are taken. Fiscal policy can be faster or slower, depending on the nature of the policy. For example, one would expect that mailing stimulus checks to households would have a faster effect than increasing spending on education. The *nominal* exchange rate usually adjusts more rapidly. In the context of the AA-DD model, the convergence to the long-run was assumed to happen "monotonically", that is, with Y and E always moving in one direction between the short run and the long run. With a slower adjustment along the Y or the E directions, or with different ways to converge to XX^b, the resulting dynamics can be very different. Figure 2.1 shows three possible cases. On the left panel, we have the type of monotonic adjustment that was commonplace in the AA-DD framework. The middle panel shows a transition from overheating to unemployment, and the right panel shows a transition from deficit to surplus.

2 The Latin Triangle

We can extend our framework to include social constraints. We will model these social constraints by assuming that there is a level of the real wage above which there is "social peace" and below which there is "social unrest". When there is social peace, there is little social pressure exerted on the government to enact reform or significantly change course.

Social unrest, on the other hand, is a situation in which people's standard of living is not acceptable, prompting society to make demands on the government. There are many levels of social unrest: from just complaining, to demonstrations, to riots, to revolution. Our way to capture social constraints is, of course, an extreme simplification of much more multidimensional and complicated issues. But even with its simplicity, we will gain what I think is valuable insight.

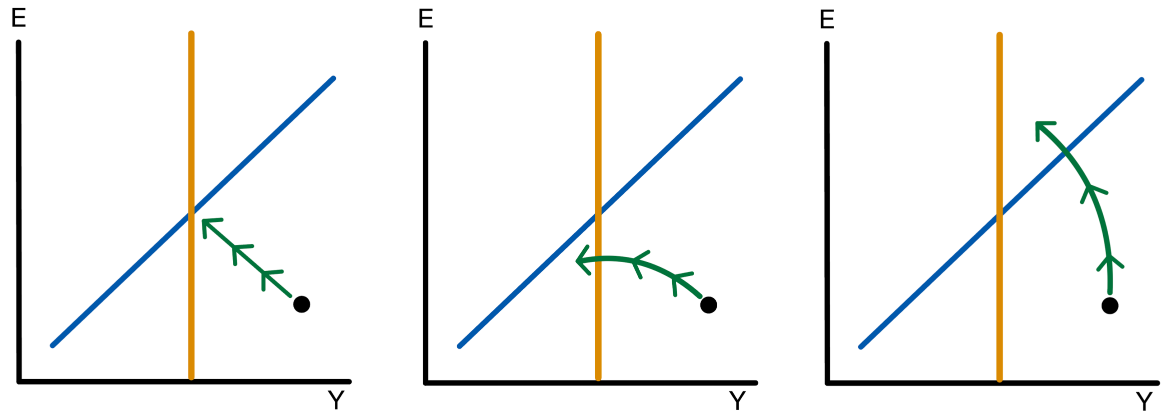


Figure 2.1:

2.1 Social Peace Line

Figure 2.2 plots the social peace curve SP as a function of output Y and the real exchange rate q. To interpret this SP curve, we first note that the real wage and the real exchange rate are inversely related[5], so moving north in the figure implies lower real wages. Social peace occurs when the economy is below the SP curve, and social unrest when the economy is above the SP curve. If social peace is achieved at a high enough real wage independent of the level of Y, the SP curve is flat.

In general, however, we can easily imagine an upward sloping SP curve. It is reasonable to assume that people may be equally happy with a lower real wage if Y is higher. For example, people may be equally happy if they have a lower real wage but the government provides better healthcare (which makes Y go up through higher G). Of course, there is a question about how to finance such change, but the social peace line is not about financing; that

[5]Technically, the *expected* real wage is inversely proportional to the real exchange rate. See Appendix 1 for details.

is what the XX is for. Here we are mostly interested in happiness. Another reason for an increasing SP curve is that people dislike to live in an economy with higher unemployment (which in our framework means low Y) for a given level of the real wage. Nevertheless, a flat SP curve suffices to make the points we want to make, and makes everything simple, so we assume the SP is flat. In addition, any aspects that we have explicitly omitted can be incorporated as "shifters" of the SP curve. Therefore, even though the schedule only depends on real wages, changes in inequality, quality of democracy, legitimacy of the government, corruption, crime, freedom of speech, etc., can be analyzed by shifting the SP curve.

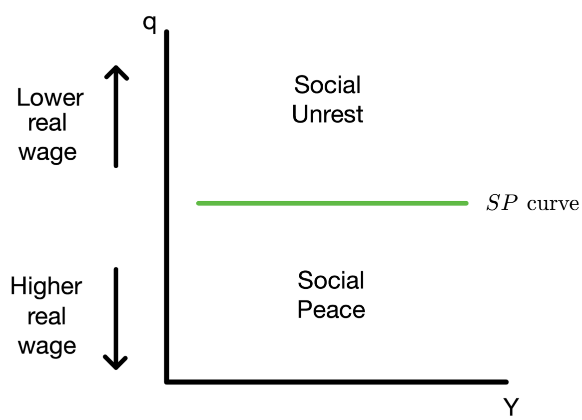


Figure 2.2:

The next step is to put the SP curve together with the II-XX framework. Where does the SP curve lie? Above, below, or through the intersection of the II and XX curves? Nothing forces the SP curve to cross at exactly the same point at which II and XX intersect.

The most interesting case is when the SP curve is below the long-run equilibrium. In Figure 2.3, we show the three curves together. Although we now have q rather than E on the vertical axis, the shapes of the II and XX curves remain unchanged since we keep P fixed (and P^* is exogenous). There are three points of intersection. At point A, the economy has internal and external balance, but real wages are too low and citizens are rioting. At point B, wages are high enough for social peace, external balance is achieved, but the economy is not in internal balance and has unemployment. Finally, at point C, wages are high enough to achieve social peace, the economy achieves internal balance, but there is external imbalance with a current account deficit.

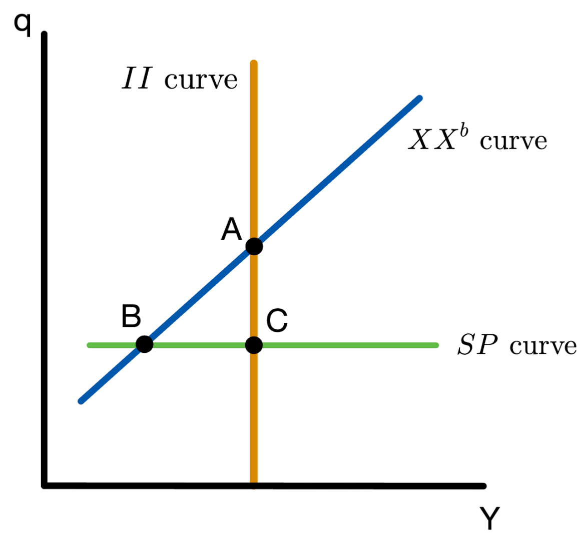


Figure 2.3:

These three points highlight the conflict between social and economic objectives. We can give some more evocative labels to the three points: point A is the *IMF equilibrium*, point B is *Europe*, and point C is *Populism*. Each of the three options has a problem. In the IMF equilibrium, the economic equilibrium has been achieved, but socially and politically, the situation is unsustainable. In Europe, even though salaries are high, standards of living are high, and the current account is in balance, the labor market exhibits significant unemployment. So, wages are high, but only for the those who actually have a job. Finally, in Populism, wages are high and unemployment is low, but this is at the expense of a significant current account deficit. Borrowing from the rest of the world is needed to keep the economy at this level of demand and exchange rates, and borrowing that much forever is not sustainable.

Because none of these options are sustainable in the long run, the economy cannot stay on any one of the three options forever. Figure 2.4 shows that, instead, the economy cycles over the options.

Assume the economy starts with a relatively conservative government at point A (IMF). Of course wages are low, but everybody is working and the current account is at a sustainable level. The low standards of living slowly lead citizens to some complaints. Later the complaints increase enough to reach levels where public demonstrations and even riots occur. The complaining takes a toll on the political landscape. The government responds by increasing wages. Because the government is relatively conservative, it does not want to increase foreign debt. Or perhaps it cannot borrow more from abroad due to prior agreements with the IMF, or other constraints arising from foreign creditors. To keep the current account at its sustainable level while increasing wages, the economy moves along the XX^b schedule from point A to point B.

As the economy is moving from A to B, unemployment starts to increase. However, at the beginning, the unemployed are few and disorganized. Politically speaking, it is relatively easy to increase wages and make the small yet increasing number of unemployed people pay the cost of the adjustment. However, when the economy reaches point B, the unemployment rate is large enough that workers get organized, and they start applying political pressure. Elections are approaching, so the government decides it is time to respond to people's demands. The government decides to keep salaries high, but implements expansionary fiscal and monetary policy to stimulate the economy and reduce unemployment. Interest rates drop, taxes are cut, highways are built.

The incumbent government moves to C and makes all citizens happy. Of course, the government is financing all the expansion through a massive current account deficit, borrowing internationally. After the elections, the government finds itself with large fiscal and current account deficits. The "country risk" increases as it becomes increasingly obvious that the creditworthiness of the country is becoming questionable. The government cuts spending somewhat, but the reduction in government spending is small relative to the already large interest payments on all the accumulated debt. Refinancing the government's debt is becoming more expensive, as creditors demand higher and higher interest rates in compensation for the increased country risk. Soon, the government starts having difficulty borrowing from abroad at ever-increasing interest rates. Eventually, the country defaults, faces a balance of payments crisis, or becomes unable to borrow a sufficient amount to keep the economy at full employment with high wages. There is a crisis. The country calls the IMF for help. Of course the IMF is willing to help, but the assessment of the situation reveals that the economy is not competitive, that wages are too high, and that the economy needs and adjustment program that reduces fiscal and current account deficits. The economy needs to return to conservativeness. The government cuts expenses drastically, increases taxes, and devalues the currency to stimulate exports and improve the current account. And the cycle starts again.

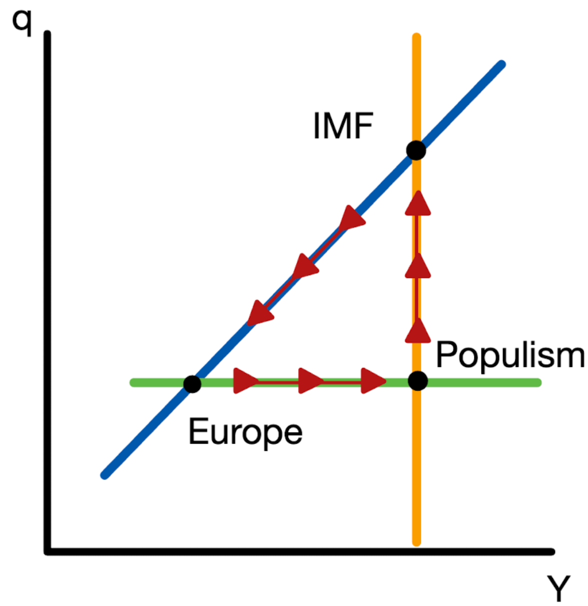


Figure 2.4:

This cycle is remarkably and scarily realistic, especially in many Latin American countries – hence the name “Latin Triangle”. We have made some very minimal assumptions thus far, yet captured a deep-rooted cycle that comes from the fact that economic objectives sometimes collide with political and social aspirations. The cycle acknowledges that there are no simple answers and that, over time, there will inevitably be periods of painful adjustment – social, political, economic. It is easy to blame incompetent politicians, greedy foreign institutions, or our corrupt political rivals for these episodes. The Latin Triangle gives us a way to more deeply understand the ultimate forces behind some of the more immediate causes of bad outcomes.

The model can help us conceptualize the social narratives that shape social sentiment in different economic environments. Let’s now revisit the cycle from from the perspective of how society sees itself. A country in the populist phase is financing its standard of living with borrowing. Unemployment is at a record low, salaries are high, aggregate demand is booming, and the economy has been doing "well" for quite a while. The only bad indicator in the economy is that it is running a current account deficit. So, yes, the economy is not competitive (since wages are "too high"), but if the future is as bright as the past, this is transitory. This false sense of invincibility, mostly driven by a sequence of previous "good economic times", masks the underlying lack of economic sustainability. This is a time of excess optimism. The future and present are perceived to be much better than what they truly are. Elections take place and the incumbent government wins.

Of course, just after the elections, the continuing borrowing puts pressure on the exchange rate, taking a toll on the central bank balance sheet. Reserves plummet to dangerous levels and the country's currency suffers a balance of payments crisis. How is that possible? How is it possible that such a successful country with only some minor weaknesses is being so unfairly treated? Financial speculation is blamed for the unwarranted attack on the currency and on the central bank. The country has no option but to devalue, cut aggregate demand, and call international institutions to repair the damaged financial situation. In fact, in some cases, the country needs to restructure or outright default on its debt. When this happens, the country is excluded from international financial markets, the interest rate increases to unthinkable levels, and a massive credit drought takes place. The government claims that the unfair adjustment is being imposed by these evil international financial institutions, amplified by speculators and vulture funds. The country has "no other option"; it is almost as if the country has lost part of its sovereignty.

At the IMF equilibrium, economic balance is reestablished, but at the cost of a significant drop in standards of living. On the economic side, the government runs a tight fiscal ship. There is neither a fiscal nor a current account deficit. There is no excess unemployment. On the social side, however, real wages are significantly lower. With depreciation comes domestic inflation. Slowly but surely, the purchasing power of consumers declines. Even though the economics is sound, the social contract is not. Poverty increases, inequality deteriorates, consumers start expressing their discontent privately and, later on, publicly. This is a time in which consumption is depressed. The bright future is now gone. The unrest prompts the government to deal with the low standards of living by compensating workers with salary increases, subsidies, a better social safety net, etc. The outcome is that the cost of the workers goes up, but because the country has been excluded from borrowing they cannot run a current account deficit. Hence, the only option is increase wages by cutting aggregate demand. The expenses by the government are increasing but, at the same time, taxes are increasing as well. The economy is now on its way to the Europe equilibrium.

Wages are high, but unemployment starts to increase. At the IMF equilibrium, everybody was employed but upset at the low wages. As wages start improving, the majority becomes happier. Only a few unemployed are disgruntled but by and large people see an improvement in their standard of living. Politically, this is palatable situation. The problem of rising unemployment can be addressed by a more generous unemployment insurance and a strengthening the social safety net. In the meantime, because the country has been running relatively good economic policy, it can, and does, re-enter international financial markets. After some time, the number of unemployed has increased so much that it is clear that the situation is no longer sustainable.

The more open financial markets and the high unemployment makes political rhetoric shift toward a more populist stance. Impositions by foreigners cannot determine the fate of millions of citizens; it is unfair that a country that was once so successful on its own is now at the mercy of heartless financial speculators. In response, government spending increases, taxes are cut, interest rates are cut, and all is financed with borrowing (mostly from abroad). At the beginning, the borrowing is manageable. The low initial fiscal debt levels allow the government to borrow at a low rate. The situation improves, unemployment declines, and wages start to go up. The government believes it has found the right policy. Why not do more of the same? And they do. The economy is now moving toward populism. The only way to sustain the economy and fulfill campaign promises is to keep borrowing more, increasing aggregate demand, and keeping wages high. Citizens feel better, and their contentment is reflected in the polls. Elections are nearby and there is no incentive to go back to orthodoxy. This is the time to concentrate on what is important: the welfare of our citizens – and their votes. And just like that, we are back to populism, and the cycle repeats itself.

2.2 Dynamics

Figure 2.4 shows a counter-clockwise spiral. If we compare these counter-clockwise dynamics with those in Figure 1.6, we see that the movements from the Europe point to Populism, and from Populism to IMF, are consistent with the "natural" adjustment dynamics of the economy. In other words, even if the government kept policy unchanged, the tendency of the economy to approach internal and external balance is compatible with the Europe-to-Populism-to-IMF dynamics. Of course, our narrative of the Latin Triangle makes clear that governments do change policies along the way, which can speed up, slow down, or change the "natural" dynamics. The point is that on this part of the Triangle, the government will find it generally easier to steer the economy since it is not going against the built-in adjustment mechanisms of the economy.

In contrast, the transition from IMF to Europe does imply an active government policy that goes against (and overcomes) the tendency of the economy to converge toward internal balance.

3 The II-XX with Domestic Demand and Real Exchange Rates

We now put more structure on the model by making stronger assumptions on how the economy adjusts over time toward internal and external balance. As anticipated, the stronger assumptions will allow us to have a sharper characterization of the dynamics that the economy follows.

To incorporate these two adjustment mechanisms into our model, it will be convenient to look at diagrams with domestic demand A and the real exchange rate q in the axes rather than output Y and the nominal exchange rate E as we had done before (we already took some steps in this direction during the Latin Triangle discussion, where we plotted q rather than E on the vertical axis). With A on the horizontal axis and q on the vertical axis, the adjustments toward internal and external balance will be very simple. Internal balance will be achieved through vertical movements (changes in q) while external balance will be achieved through horizontal movements (changes in A). The switch from Y and E to A and q also has the conceptual advantage of separating domestic policy (changes in A) from external policy (changes in q).

Therefore, our first job is to express all the equations we will need as functions of A and q. After that, we study the adjustment dynamics.

3.1 Domestic Demand

Let's recall that demand for domestic goods is not the same as domestic demand, since domestic demand includes demand for all goods, both domestic and foreign. Demand for domestic goods is

$$Z = C + I + G + CA,$$

while domestic demand is

$$A = C + I + G. \tag{3.1}$$

In the present context, domestic demand is sometimes referred to as "domestic *a*bsorption", which is why I have denoted it with an A. The term absorption is supposed to evoke the idea that when demand for domestic goods goes up, some of that demand is "absorbed" by A, while the rest is "lost" to imports. Domestic demand is also referred to as "aggregate demand", since it is the total demand for all goods, foreign and domestic.

As usual, we maintain the assumption that consumption is an increasing function of disposable income: $C = C(Y - T)$. Using $C = C(Y - T)$ in equation (3.1) gives:

$$A = C(Y - T) + I + G \tag{3.2}$$

so A is an increasing function of Y. In addition, since marginal propensity to consume out of disposable income is less than one (consumers on aggregate save at least some part of their income), an increase in Y is associated to a less than one-for-one increase in A.

The current account is also modeled just as before:

$$CA = CA(EP^*/P, Y, Y^*),$$

where CA is a function that is increasing in the real exchange rate EP^*/P, decreasing in domestic income Y, and increasing in foreign demand Y^*.

With a supply of goods given by Y, the equilibrium condition in the market for domestic goods is

$$Y = C + I + G + CA. \tag{3.3}$$

To write our model in terms of A, we combine equations (3.1) and (3.3) to get

$$A = Y - CA(EP^*/P, Y, Y^*)$$

Now use that $q = EP^*/P$ to write

$$A = Y - CA(q, Y, Y^*) \tag{3.4}$$

3.2 The *II* Curve in Terms of *A* and *q*

In this context, we define the II curve as the set of points (A, q) compatible with a goods market equilibrium that has $Y = Y^f$ in the short run (i.e., with P fixed). Using $Y = Y^f$ in equation (3.4) gives the II curve:

$$A = Y^f - CA(q, Y^f, Y^*). \tag{3.5}$$

Figure 3.1 plots the II curve from equation (3.5) with A on the horizontal axis and q on the vertical axis. The II curve gives the combination of domestic demand levels and real exchange rates that, keeping prices fixed, gives a short-run equilibrium in the goods market that has output equal to its full-employment level.

In this model, A and q are both endogenous. We can take q as given and see what happens to A, or we can take A as given and see what happens to q. In the end, the model determines the joint behavior of A and q. Taking q as given, the II curve provides the answer to the following question: For a given value of q, what is the level of domestic demand (or

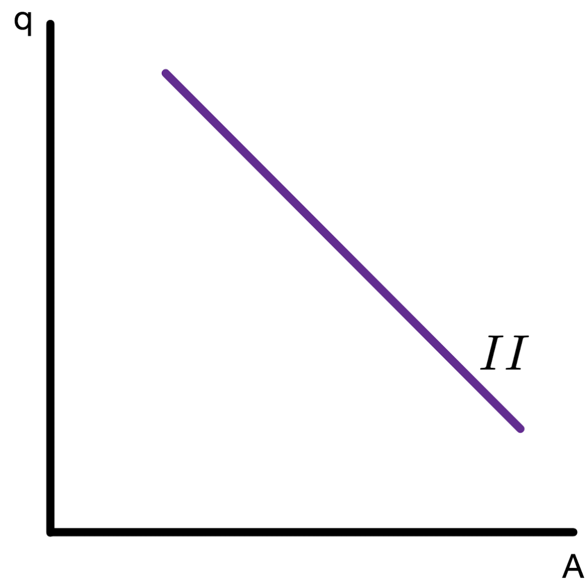


Figure 3.1:

absorption A) required to make output equal to the full-employment level of output? Now let's assume that A is given. First, consider a given A that is much smaller than Y^f. In order for output to equal Y^f, there has to be a large current account surplus. The only way to achieve a large current account surplus is to have a real exchange rate that is sufficiently high (a large enough depreciation). Consider now a given A that is above Y^f. If $CA = 0$, then we must have $Y > Y^f$. In order to have $Y = Y^f$, we need a current account deficit. The only way to achieve a deficit in the short run is if q is low enough (a large enough apreciation).

We note that this II curve is conceptually similar to the II curve that we introduced in Section 1, but not exactly the same. The earlier version of the II curve was defined as the set of points (Y, E) that are consistent with $Y = Y^f$ in the short run. The new II curve in equation (3.5) gives the set of points (A, q) that are consistent with a short-run goods market equilibrium that has $Y = Y^f$. In addition to considering (A, q) instead of (Y, E), the new II introduces the requirement that (A, q) can *support an equilibrium* with $Y = Y^f$. The earlier II curve, in contrast, just required $Y = Y^f$ without any reference to equilibria.

3.3 The *XX* Curve in Terms of *A* and *q*

Let's now write the XX^b curve as a function of A and q. Equation (1.2) with $XX = XX^b$ and $q = EP^*/P$ gives the XX^b curve:

$$XX^b = CA(q, Y, Y^*). \tag{3.6}$$

To write CA as a function of A instead of Y, we solve equation (3.2) for Y to get:

$$Y = C^{-1}(A - I - G) + T, \tag{3.7}$$

where $C^{-1}(\cdot)$ is the inverse function of the increasing function $C(\cdot)$. For example, if

$$C(x) = 1 + \frac{3}{4}x \tag{3.8}$$

then

$$C^{-1}(x) = \frac{4}{3}(x - 1). \tag{3.9}$$

Plugging equation (3.7) into equation (3.6) gives:

$$XX^b = CA\left(q, C^{-1}(A - I - G) + T, Y^*\right).$$

We can write this equation as

$$XX^b = \widetilde{CA}(q, A, Y^*) \tag{3.10}$$

where the function $\widetilde{CA}$ is defined by

$$\widetilde{CA}(q, A, Y^*) \equiv CA\left(q, C^{-1}(A - I - G) + T, Y^*\right).$$

Since $C(\cdot)$ is increasing and $CA(\cdot, \cdot, \cdot)$ is decreasing in its second argument, $\widetilde{CA}$ is also decreasing in its second argument. It is good to keep in mind that sometimes, with some abuse of notation, we write $XX^b = CA(q, A, Y^*)$ instead of the more correct but cumbersome $\widetilde{CA}(q, A, Y^*)$.

Example 3.1. We keep using equation (3.8). We additionally assume that

$$CA(q, Y, Y^*) = 2q - 0.5Y + 0.1Y^*.$$

Using (3.9), we have

$$\begin{aligned}
CA(q, Y, Y^*) &= 2q - 0.5Y + 0.1Y^*, \\
&= 2q - 0.5[C^{-1}(A - I - G) + T] + 0.1Y^*, \\
&= 2q - 0.5\left[\frac{4}{3}(A - I - G - 1) + T\right] + 0.1Y^*, \\
&= 2q - \frac{2}{3}A + \frac{2}{3}(I + G + 1) - 0.5T + 0.1Y^*.
\end{aligned}$$

So, in this case

$$\widetilde{CA}(q, A, Y^*) = 2q - \frac{2}{3}A + \frac{2}{3}(I + G + 1) - 0.5T + 0.1Y^*.$$

$\square$

Figure 3.2 plots the XX^b curve from equation (3.10) with A on the horizontal axis and q on the vertical axis. This is the same curve as the upward sloping curve from Figures 1.3, 1.5 and 1.6, but plotted in a different axis. The XX^b curve is the set of points (A, q) that are consistent with a current account equal to XX^b in the short run. It answers the following question: For a given value of q, what is the level of domestic demand (or absorption A) required to make the current account equal to its sustainable level XX^b? Just as for the II curve, we can also take A as given and ask what happens to q. The XX^b curve therefore also answers the question: For a given value of A, what is the level of q required to make the current account equal to its sustainable level XX^b?

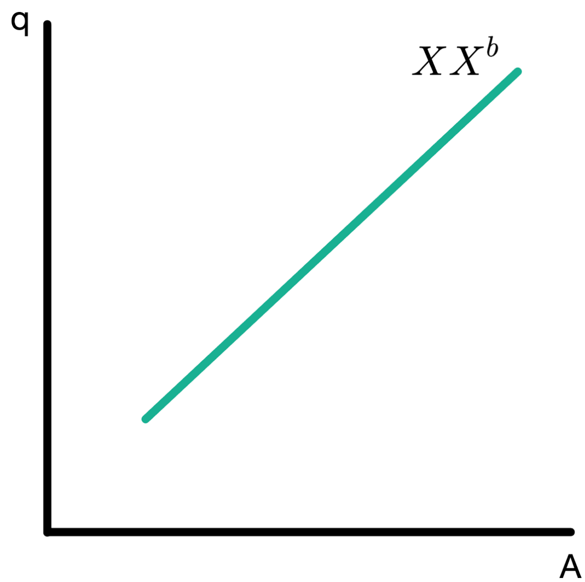


Figure 3.2:

3.4 The Four Zones of Economic Discomfort

The left panel of Figure 3.3 shows the II and XX^b curves in the same plot, with A and q in the axes. Similar to Figure 1.5, we can characterize the four regions of economic discomfort based on the combinations of internal and external imbalances that prevail.

The points above the II curve have $Y > Y^f$ and, conversely, the points below the II curve have $Y < Y^f$.[6] For the XX^b curve, we have that points above it have $CA > XX^b$ while points below it have $CA < XX^b$. Points exactly on II have $Y = Y^f$ and correspond to points in which there is internal balance. Points exactly on XX^b have $CA = XX^b$ and correspond to points in which there is external balance.

[6]To see that $Y > Y^f$ is the same as being above the II curve, use equations (3.5), (3.4), and that domestic demand A is increasing in Y.

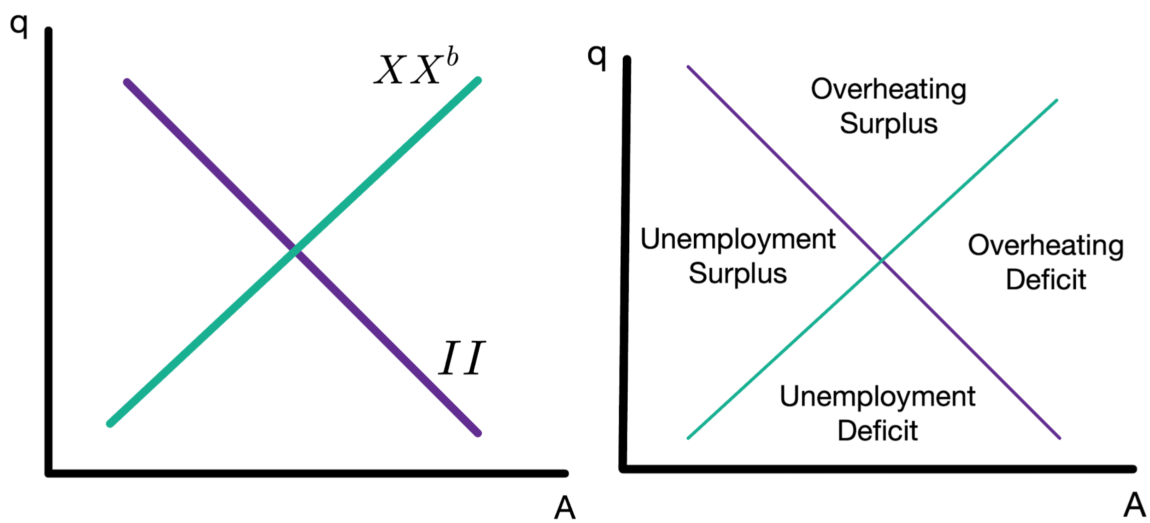


Figure 3.3:

3.5 Adjustment Mechanisms

We are finally ready to make the stronger assumptions regarding how the economy adjusts over time toward internal and external balance.

Adjustment Mechanism Toward Internal Balance

- The economy tends toward internal balance through adjustments in the price level P. When $Y > Y^f$, P is increasing. When $Y < Y^f$, P is decreasing. As time goes by, absent any changes in exogenous variables or parameters, the changes in P will make Y approach Y^f.

What does this adjustment mechanism imply for the dynamics of A and q? The left panel of Figure 3.4 shows the dynamics that arise from adjustments in P. All the points above the II curve have $Y > Y^f$. For these points, P is increasing over time. When P is increasing over time, q is decreasing over time (since $q = EP^*/P$). Conversely, for points below the II curve we have $Y < Y^f$, P decreasing and q increasing. Thus, internal balance is achieved by movements in the real exchange rate q, which are movements in the north-south direction.

In the AA-DD model, we had already encountered an adjustment mechanism driven by changes in P. In fact, for the AA-DD model, it was the *only* endogenous adjustment mechanism. When P changed, it induced shifts in both the AA and the DD. The shifting of the curves stopped when their intersection gave an equilibrium with $Y = Y^f$. Thus, starting from any short-run equilibrium, the adjustments in P pushed the economy toward a long-run equilibrium. One aspect that is different in the current model compared to the AA-DD model is that we now assume that adjustments in P serve as an adjustment mechanism for internal balance only. In the AA-DD, changes in P induced changes in Y, q and E, moving all of them toward their long-run equilibrium levels. As a consequence, changes in P also induced changes in A in the AA-DD model (although we never looked at that directly when we studied the AA-DD). In contrast, in the present II-XX model, changes in P induce changes in q only; neither Y nor E change with P.

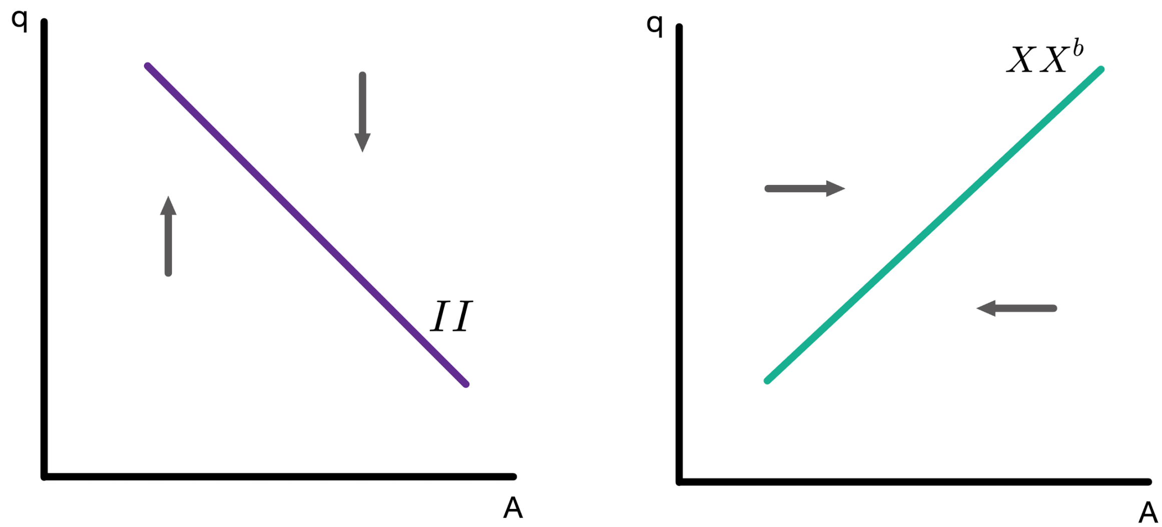


Figure 3.4:

Adjustment Mechanism Toward External Balance

The adjustment mechanism we consider now is a new mechanism that is absent in the AA-DD framework and that we have not encountered (in this course) until now.

- The economy tends toward external balance through adjustments in aggregate consumption C. When the current account is below its sustainable level XX^b, consumption is decreasing over time. Lower consumption leads to lower aggregate demand A, lower output and lower income. Lower income makes imports go down, which makes the current account increase and get closer to XX^b. Conversely, when the current account is above its sustainable level XX^b, consumption, output and income are increasing over time. The current account is deteriorating and moving toward XX^b.

The right panel of Figure 3.4 shows the dynamics that are brought about by this adjustment mechanism.

Let's now understand the economic logic behind it. At the deepest level, the adjustment mechanism of the current account through consumption and aggregate demand comes from two sources. The first is the existence of an aggregate budget constraint for the entire economy that says that we cannot finance too high a level of consumption with ever-increasing borrowing. The second is that people value consumption and will not choose to consume less in every period if it can consume more in every period with the same resources.

The existence of a budget constraint provides the adjustment mechanism of CA toward XX^b when $CA < XX^b$. The preference for higher consumption provides the adjustment mechanism of CA toward XX^b when $CA > XX^b$.

Imagine the economy has an unsustainable current account deficit, which in this context means $CA < XX^b$. In the right panel of Figure 3.4, points with $CA < XX^b$ correspond to points below the XX^b line. When $CA < XX^b$, the country is consuming more than what is sustainable in the long run and financing this consumption by borrowing from the rest of the world. As time goes by, the country's debt to foreigners keeps increasing. The borrowing being unsustainable means that, eventually, the borrowing must go down, and the debt levels must stabilize (or else the budget constraint is violated). When the borrowing goes down, consumption levels cannot be sustained and must go down as well. Lower consumption leads to lower domestic demand. In a goods market equilibrium, lower demand implies lower output. Since aggregate output equals aggregate income[7], aggregate income goes down. Lower aggregate income implies that demand for imports is lower, so equilibrium imports go down. Lower imports make CA go up, which is an adjustment toward XX^b.

[7] See footnote 3.

In the chain of events described in the last paragraph, the unsustainably high borrowing must *eventually* go down. There is in principle no reason why we can't keep an unsustainable deficit, or even increase the deficit, for some time before having to reduce consumption. After all, whether a current account level is sustainable is a long-run issue. Short-run movements away from sustainability may be consistent with long-run sustainability if they are reversed quickly enough. However, in our model we do assume that the adjustment toward long-run sustainability and toward external balance begins immediately. This does not rule out all temporary movements away from sustainability. There can be movements away from sustainability brought about by changes in exogenous variables or parameters. However, it is good to keep in mind that, in the real world, endogenous movements away from sustainability can occur, do occur, and are sometimes quite persistent.

Imagine now that the economy has an unsustainable current account surplus, $CA > XX^b$. When $CA > XX^b$, the country's income is higher than its aggregate demand. The difference between income and demand is lent out to the rest of the world. As time goes by, the domestic economy accumulates wealth in the form of foreign assets – the IOUs that entitle the domestic economy to future repayments from the rest of the world. This accumulation of wealth is a form of saving, as it foregoes consumption today (by lending to foreigners) in exchange for consumption in the future (when we use the IOU repayments to increase consumption). The accumulation of saving becomes unsustainable when it is not desirable for consumers to keep lending forever at the current level, as that would imply that, for all future periods, wealth is being accumulated faster than it is consumed. And if wealth is forever being accumulated faster than it is consumed, it is possible to consume some of the wealth today and also keep all future consumption unchanged. Therefore, when the current account surplus becomes unsustainable, it is optimal for consumers to consume more and lend less. The reduced lending to the rest of the world lowers the current account toward its sustainable level, while the higher consumption increases domestic demand.

In sum, countries cannot borrow at an unsustainable level forever, and do not want to save at an unsustainable level forever. In either case, the external imbalances are corrected over time by changes in aggregate demand A. In the right panel of Figure 3.4, this corresponds to movements in the east-west direction. When the external imbalance is in the form of a surplus, the economy is above the XX^b line and the adjustment toward external balance happens by increasing A over time. When the external imbalance is in the form of deficits, the economy is below the XX^b line and the adjustment toward external balance happens by lowering A over time.

3.6 Adjustment Dynamics

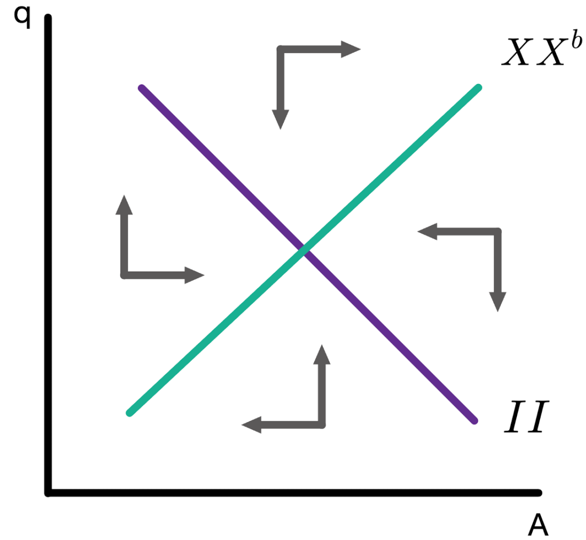


Figure 3.5:

Combining the dynamics from the two panels of Figure 3.4 gives Figure 3.5. Each of the four zones of economic discomfort have their own dynamics. For example, the top quadrant has overheating and surplus. Because it has overheating, the price level is increasing, which pushes q down. Because it has surpluses, consumption is increasing, which pushes A to the right (toward higher values).

Figure 3.6 places the economy at a starting point in the top quadrant and traces the path that results from following the adjustment dynamics from Figure 3.5. The result is a clockwise spiral.

We can get some further insight by looking at the strength of each of the two adjustment mechanisms for different combinations of A and q. Figure 3.7 illustrates the strength of the adjustment mechanisms with arrows attached to the path traversed by the economy

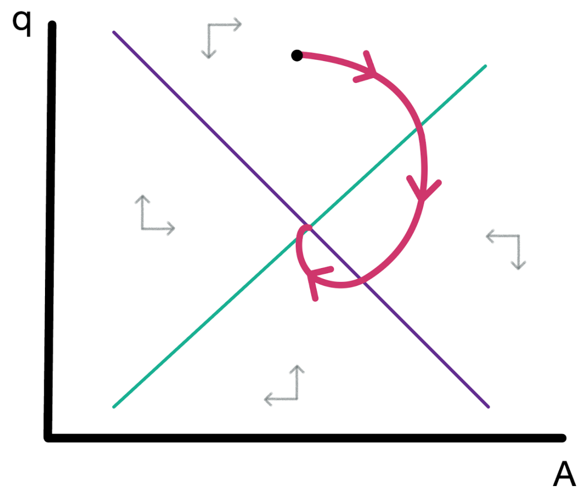


Figure 3.6:

over time. Longer arrows denote a "stronger" adjustment. We equate stronger adjustments with faster adjustments; we can think of the length of an arrow as indicating the speed at which the economy is moving in the direction of the arrow. The arrows pointing up or down correspond to adjustments in q that move the economy toward internal balance while the arrows pointing east or west correspond to adjustments in A that move the economy toward external balance. A natural assumption, which we make from now on, is to have adjustments be faster when the economy is far away from balance, and slower when the economy is closer to balance.

In Figure 3.7 the red path starts with two arrows that are relatively large and of similar size. The arrows are large because the economy is far from the lines. The arrows are of similar size because the distance to the purple II curve is similar to the distance to the green XX^b curve. We expect the economy to move down and to the right at approximately similar speeds. As the economy moves down and to the right, it gets closer to the green XX^b curve of external balance. As it gets closer to XX^b, the arrow pointing to the right becomes shorter. For example, the arrow pointing to the right at the midpoint of the red path is shorter than the arrow pointing to the right at the beginning of the red path.

At the midpoint of the red path, we also see that the arrow pointing to the right is shorter than the arrow pointing down. The arrow pointing right is shorter because the economy is closer to the green external balance XX^b curve than to the purple internal balance II

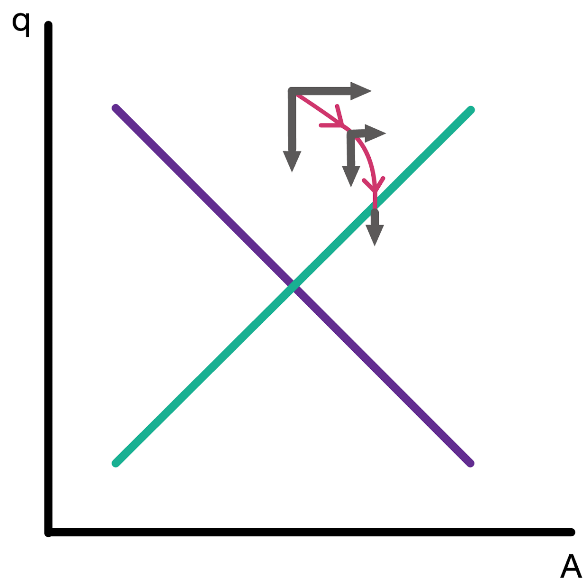


Figure 3.7:

curve. Adjustment toward external balance (moving to the right) is happening at a slower pace than the adjustment toward internal balance (moving down).

When the economy is at the end of the red path, it is on the green XX^b line, so it is in external balance. When the economy is in external balance, there is no force pushing the economy toward higher or lower A, which is why there is no horizontal arrow at that point (equivalently, the arrow has zero length). On the other hand, there still is internal imbalance in the form of overheating, which pushes q down. The economy is therefore moving straight down. The length of the arrow pointing down at the end of the red path is smaller than the two other arrows pointing down in the red path because the economy is closer to internal balance at the end of the red path than at any other point in the path.

The same reasoning applies when the economy finds itself on the purple II line. In this case, there would be no north-south adjustment and the economy would be moving only in the east-west direction toward the green XX^b line. You can try starting the economy in different places and trace the path it follows. The result will always be adjustment dynamics that trace a clockwise spiral path that converges toward the intersection of the II and XX^b curves.

4 Environmental Sustainability

We now look at how to incorporate environmental concerns to the framework. As usual, we simplify the problem, not with the intention of trivializing it, but to have a manageable model that, despite its simplicity, is still useful.

With environmental concerns, we do not see cycles like the ones that arise in the Latin Triangle. The reason is that the environment moves much slower than economic and social events. The slow-moving nature of the environment makes some of the tensions between environmental sustainability and all other goals more dangerous. If the environment were already deteriorating at a fast enough pace, we probably would have taken corrective actions already. Instead, social choices are more focused on the economic and social cycles.

At the end of this section, we discuss the type of technological progress that can help resolve the tension between economic and environmental goals.

4.1 Environment Schedule

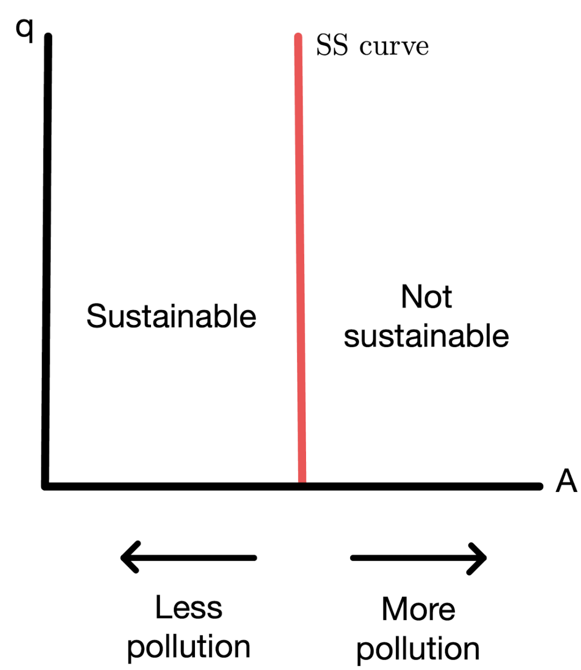


Figure 4.1:

We assume that there exists a level of aggregate demand A below which the impact on the environment is sustainable. This assumption is meant to capture that some of the environmental problems of excessive consumption of natural resources (renewables or not), excessive production and pollution, congestion, and many others, require a large enough level of consumption and investment. We do not assume that sustainability depends on exchange rates – which sounds like a reasonable assumption, especially in the long run. Figure 4.1 shows a diagram that represents the idea of an aggregate level of A below which we have environmental sustainability. We refer to the vertical line that determines this cut-off level of A as the SS curve. To the right of SS, the environment suffers. The more to the right, the more it suffers. To the left of the SS curve, the environment is protected and consumption is sustainable. We have once again simplified and labeled the complex effects that lead to a less sustainable environment as "more pollution", and those that improve the environmental sustainability as "less pollution".

4.2 Kyoto's Triangle

Given the slow-moving nature of the environment, the uncertainty in the speed and type of future technological progress, the scientific disagreements in long-run environmental predictions, and the likelihood and variety of potential policies that may or may not be implemented, it is difficult to know where to place the SS relative to the other curves we have been studying. We will look at the case in which the level of consumption implied by having internal and external balance (the IMF equilibrium) is already too high to be consistent with a sustainable environment. Even if other placements of the SS curve are possible and interesting, the case we look at presents a number of challenges that, to me, seem relevant in today's world and will become even more relevant as time goes by.

Figure 4.2 adds the SS curve to Figure 2.3, with the SS to the left of the IMF equilibrium. The SS goes exactly through the Europe point, but this is not so important. If we move the SS slightly to the left or to the right, other points of intersection would arise among the curves. However, not much would change in terms of the insights we are interested in.

In Figure 4.2, the intersection of the red environmental sustainability SS curve with the purple internal balance II curve is labeled "Kyoto". Now we have four locations to think

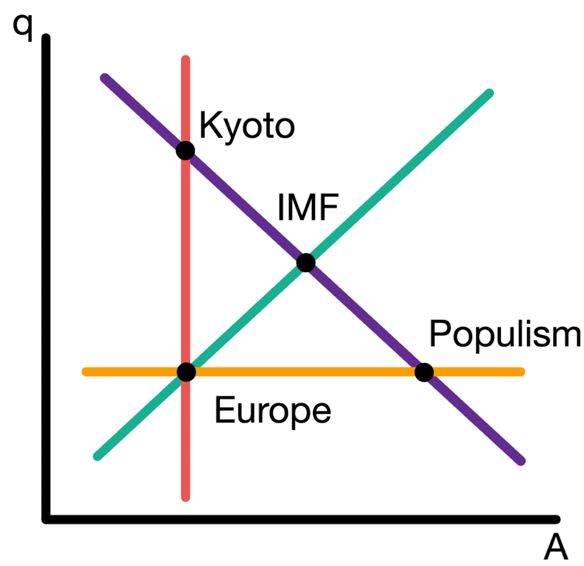


Figure 4.2:

about, each of which has its own challenges. Table 4.1 summarizes them.

At the IMF equilibrium, the economy is in internal and external balance, but it exhibits social tensions and the demand implies the environmental sustainability – polluting, for simplicity – is not attained. In other words, we have economic balance, but neither social nor environmental balance. Even though there is full employment and there are no current account pressures, the economy is unsustainable.

In the point labeled Europe, the country has external balance, the standard of living is relatively high – given the high real wage – and the demand is low enough that the impact in the environment is low enough for sustainability. The only problem is that the country has a relatively large level of unemployment. So, even though the economy is sustainable along the social, environmental, and external balance dimensions, labor markets are in significant disarray. This is sustainable for a while, but the social safety net required to keep the unemployed content is a large burden for the economy.

At the point labeled Populism, political and social concerns are minimal. Real wages are high and employment is at the natural rate. People are very happy regarding economic matters. However, the country is running an unsustainable current account deficit that requires borrowing from foreigners – the equivalent of running down savings. In addition, aggregate demand is so high that the environmental impact is the worse out of the five options we are considering. This is clearly an unsustainable situation. In general, however, the escape valve is the current account. In most cases, a sharp reduction in (or outright shutdown of) foreign lending is what forces the economy to adjust – not the environmental impact.

Finally, the point labeled Kyoto is in many ways similar to Europe. Aggregate demand is low enough that the environmental impact is sustainable. However, instead of workers having to suffer from unemployment, it is real wages that are too low. At the Kyoto point, because wages are depressed, the costs of production are low by international standards, the export sector is very competitive and the current account has a large surplus. The label "Kyoto" is evocative not only of Japan, which is arguably in a low-demand, depressed-wage situation (or used to be until recently), but also of the Kyoto Protocol, one of the first international treaties to reduce greenhouse gas emissions. Saying that wages in Japan are depressed does not mean that wages in Japan are lower than their counterpart in other less developed countries. Wages in Japan are depressed relative to the wages that would prevail at the IMF equilibrium. In other words, wages are below the level required to achieve external balance. Because wages are low, the economy has a current account surplus. This is also an unsustainable situation. First, there is tension in the political system. Second, not all countries in the world can run surpluses – by definition, every surplus must be balanced by a deficit.

	Internal Balance	External Balance	Social Peace	Environment
IMF	✓	✓	Conflict	Pollution
Europe	Unemployment	✓	✓	✓
Populism	✓	Deficit	✓	High Pollution
Kyoto	✓	Surplus	Conflict	✓

Table 4.1:

4.3 Productivity Improvements

Technological improvements can increase production without increasing the use of resources. This is one of the "solutions" to environmental sustainability issues that has received a lot of attention in the public discussion. However, technological improvements have two implications. First, if production increases without using additional resources, the SS shifts to the right, which allows the economy to sustain a larger demand with a smaller impact on the environment. Second, if technological improvements come with an increase in productivity, the full-employment level of output Y^f also increases – the same labor force can now produce more output. Changes in Y^f shift the II curve.

The overall effect of technological improvements depends crucially on how much the SS and II curves shift. Figure 4.3 shows three possibilities. On the left panel, the technological improvement has a relatively large impact on production and a relatively small impact on environmental sustainability. In this case, the IMF equilibrium has higher demand and a lower wage than before, making social unrest more severe. If the change in SS is small enough, even the original IMF equilibrium (at the intersection of the dashed purple initial II curve and the green XX^b) remains unsustainable.

The center panel of Figure 4.3 shows a productivity increase that makes both the SS and the II shift more than in the left panel. If the II did not shift, then the IMF equilibrium would be sustainable. However, the technological improvements have made much more production feasible. The new IMF equilibrium can support a level of demand that is too high to be environmentally sustainable, even after the large shift in the SS.

Last, the right panel of Figure 4.3 shows the kind of technological progress that solves the environmental problems. This kind of "green" technology increases the level of production that is environmentally sustainable without simultaneously increasing demand beyond that new environmentally sustainable point.

The bad news is that social unrest remains in all panels of Figure 4.3. To address social peace and environmental sustainability together, we must consider shifts in the XX^b curve.

Figure 4.4 shows three possibilities when the XX^b curve shifts. These shifts can come from

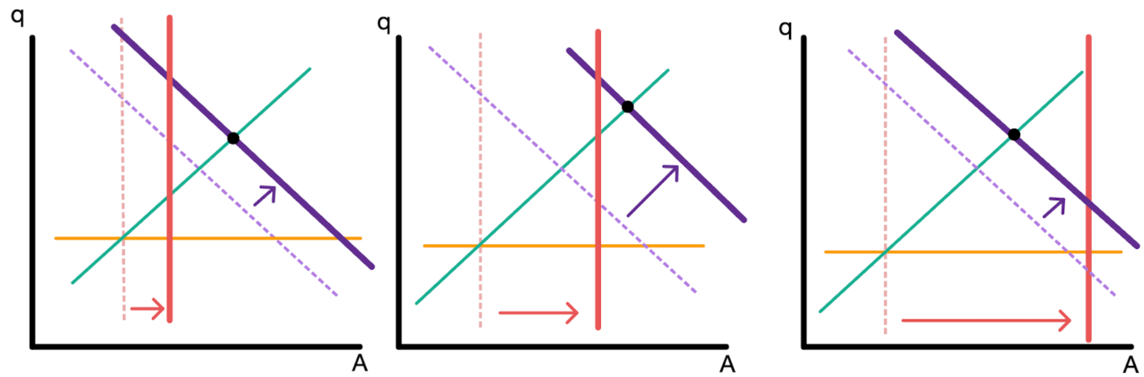


Figure 4.3:

many sources. Here we still consider technological changes, which can induce changes in the level XX^b that is sustainable for the current account, or in foreign demand Y^* if technology diffuses globally.

On the left panel, the technological improvement has a large impact on the XX^b curve but a small impact on environmental sustainability. In this case, the IMF equilibrium implies a much higher standard of living, a higher demand, and, if the improvements are large enough, social peace. On the other hand, the environmental impact, measured by the horizontal distance between the SS curve and the IMF equilibrium, is now larger.

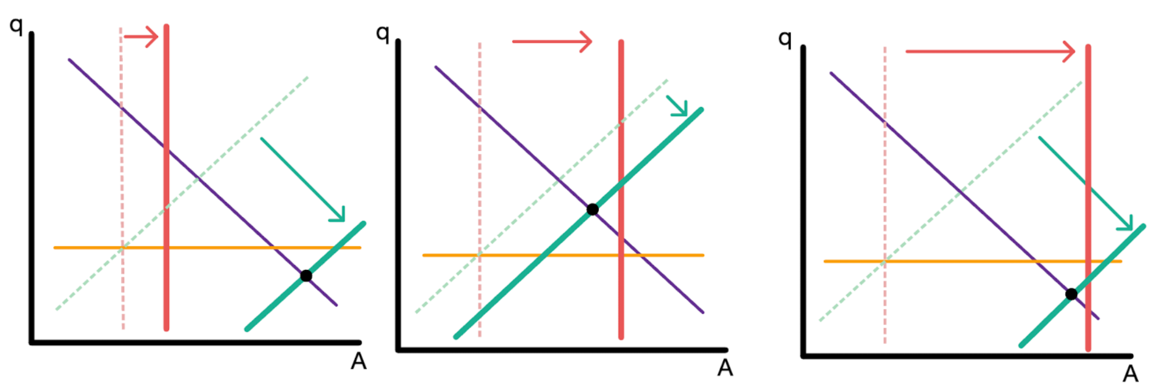


Figure 4.4:

The center panel of Figure 4.4 shows a technological change that makes the SS shift significantly, and the XX^b shift little. The improvement in environmental sustainability is so large that the IMF equilibrium is to the left of the SS. There is no negative environmental outcome associated with this type of technological progress. On the other hand, the Latin

Triangle and its implied cycles are a concern. Even though the environment is sustainable, the economy and the politics are not.

Last, the right panel of Figure 4.4 shows the kind of technological progress that solves both the Latin Triangle and the environmental problems.

1 Appendix

1.1 Labor Market

In this Appendix, we develop a small model of the labor market, taken straight from intermediate macro. You do not need to know this model of the labor market, I will not ask you about it in exams.

In our labor market model, labor is the only factor of production and $Y = N$, as we had assumed above. Since the only factor of production is labor, the marginal cost for firms producing goods and services is the wage W. If markets were perfectly competitive, firms would set the price P equal to marginal cost W and make zero profits. In the real world, there is market power and firms tend to set prices above marginal cost. We capture this idea by the following price-setting relation:

$$P = (1 + m)W. \tag{1.1}$$

where m is the markup. This price setting relation says that firms set the price of their goods based with a markup above their marginal cost W.

Wages are set according to the following wage setting relation:

$$W = P^e F(\underset{(-)}{u}, \underset{(+)}{z}). \tag{1.2}$$

In this wage setting relation, P^e is the expected price level (the price level we expect to hold in the future), u is the unemployment rate, z is a catch-all variable that captures all other elements that influence wages, and F is a function that is decreasing in u and increasing in z (as represented by the plus and minus signs in the formula). There are therefore three elements that determine the wage: P^e, u and z. We briefly explain why they matter for wages:

- Expected price level. The expected price level P^e is an important determinant of the nominal wage because what people and firms really care about is not the nominal wage but the real wage. Workers value money because of the goods and services they can buy with it, not for its own sake. If my nominal wage is \$1,000,000 but the price of one apple is \$10,000,000, I am not so happy with my gigantic nominal wage. We use P^e rather than P because after wages are set, they usually remain fixed (or closed to fixed) for some time, with wage contracts renegotiated only infrequently. So the relevant price level is the one that will prevail between now and the future time when the wage is renegotiated rather than just the current price level. goods and services that one can afford after W is set are better captured by W/P^e.

- Unemployment. The wage setting equation says that when unemployment is high, wages are low. The reasoning is that when unemployment is high, employers have more bargaining power, since someone seeking a job must compete with a larger pool of unemployed workers. Conversely, when the unemployment rate is low, it is workers who have higher bargaining power, as many firms have to compete to hire among the smaller pool of unemployed people.

- Other factors. The other factors are defined so that when z goes up, wages go up. That F is increasing in z is arbitrary, we could have just as well assumed the opposite and change the meaning of z. One example of a factor that enters z are unemployment benefits. For given u and P^e, higher unemployment benefits make unemployment less painful, so employers must offer a higher wage to attract workers (in this case higher unemployment benefits were associated with a higher z).

Combining (1.1) and (1.2) gives

$$P = (1+m)P^e F(u,z). \tag{1.3}$$

In the long-run, $P = P^e$. Using $P = P^e$ in (1.3) defines the natural rate of unemployment

u^n:

$$1 = (1+m)F(u^n, z).$$

Given u^n, we can find the full-employment level of output Y^f using equation (1.1).

Equation (1.1) implies that the expected real wage is:

$$\frac{W}{P^e} = \frac{1}{1+m}\frac{P}{P^e} \tag{1.4}$$
$$= \frac{1}{(1+m)(1+\pi^e)} \tag{1.5}$$

where I have used the definition of expected inflation $\pi^e \equiv P^e/P - 1$.

1.2 Real uncovered interest parity

Using the definitions of the real exchange rate

$$q \equiv EP^*/P \tag{1.6}$$

and of expected inflation

$$\pi^e \equiv P^e/P - 1, \tag{1.7}$$

we have that

$$\begin{aligned}
\frac{E^e}{E} &= \frac{q^e \frac{P^e}{P^{*e}}}{q \frac{P}{P^*}} \\
&= \frac{q^e}{q}\frac{1}{P^{*e}/P^*}\frac{P^e}{P} \\
&= \frac{q^e}{q}\frac{(1+\pi^e)}{(1+\pi^{e*})} \\
&= \left(1 + \left[\frac{q^e}{q} - 1\right]\right)\frac{(1+\pi^e)}{(1+\pi^{e*})}
\end{aligned}$$

A linear approximation gives

$$\frac{E^e}{E} \approx 1 + \left(\frac{q^e}{q} - 1\right) + \pi^e - \pi^{e*}$$

Using this linear approximation in the uncovered interest parity condition

$$R - R^* = \frac{E^e}{E} - 1$$

gives

$$R - R^* = \left(\frac{q^e}{q} - 1\right) + \pi^e - \pi^{e*}$$

Using the definition of the real interest rate

$$r^e = R - \pi^e, \tag{1.8}$$

we get the real uncovered interest parity condition

$$r^e - r^{e*} = \left(\frac{q^e}{q} - 1\right) \tag{1.9}$$

1.3 Real Wages and Real Exchange Rates

Using (1.8) and (1.9) in (1.5) gives

$$\begin{aligned}\frac{W}{P^e} &= \frac{1}{(1+m)(1+\pi^e)} \\ &= \frac{1}{(1+m)(1+R-r^e)} \\ &= \frac{1}{(1+m)\left(1+R-r^{e*}-\left[\frac{q^e}{q}-1\right]\right)}\end{aligned}$$

A linear approximation gives

$$\begin{aligned}\frac{W}{P^e} &\approx 1 - m - \left(R - r^{e*} - \left[\frac{q^e}{q} - 1\right]\right) \\ &= 1 - m + r^{e*} - R + \left(\frac{q^e}{q} - 1\right)\end{aligned}$$

which shows that, for a given domestic interest rate R, the expected real wage W/P^e is proportional to real expected depreciation $q^e/q - 1$. In turn, for a given expected real exchange rate q^e, real expected depreciation is inversely related to the current real exchange rate q. It follows that, for given R and q^e, the expected real wage W/P^e and the real exchange rate q are inversely related.