Macroeconomic Theory and Policy (2nd Edition)

Andolfatto, David

Simon Fraser University

1 January 2008

MACROECONOMIC THEORY & POLICY

2nd Edition

David Andolfatto

Simon Fraser University

2008

Chapters

Chapter Contents

1. **The Gross Domestic Product**. National Income and Product Accounting. Measurement Issues. Nominal versus Real GDP. Trends and Cycles. Schools of Thought.

2. **Output and Employment.** The Market for Goods and Labor. General Equilibrium in a Static Income-Leisure Model. A Neoclassical Interpretation of the Business Cycle. Welfare-Reducing Stabilization Policy.

3. **Uncertainty and Expectations**. Decision-Making Under Uncertainty. Ex ante versus ex post outcomes. Rational and Irrational Expectations. Multiple Equilibria and Animal Spirits. Welfare-Improving Stabilization Policy.

4. **Unemployment**. Labor Market Flows. A Simple Indivisible-Labor Model of Employment Choice. Reservation Wages. Modeling Unemployment. Government Insurance Policies.

5. **Consumption and Saving**. A 2-Period Endowment Model. International Borrowing and Lending. The Real Rate of Interest. The Intertemporal Budget Constraint. Consumer Demand and Desired Domestic Saving. The Current Account and the Trade Balance. Transitory, Anticipated, and Persistent Productivity Shocks. Reverse Causality.

6. **Fiscal Policy**. Accounting and Data. Government Spending and Taxation in a Static Model. Lump-Sum versus Distortionary Taxation. Government Transfer Policies. Government Deficits in a 2-Period Model. The Intertemporal Government Budget Constraint. The Ricardian Equivalence Theorem. Transitory Government Expendenture Shocks and Tax-Smoothing Policy. Tax versus Bond-Financed Increases in Government Purchases.

7. **Capital and Investment**. Domestic Investment Demand in a 2-Period Model. A Small Open Economy with Saving and Investment. Interpreting the Cyclical Behavior of the Trade Balance. Closed Economy Analysis and the Determination of the Real Rate of Interest. The IS Curve and the Keynesian Cross. Interpreting Policy Implications Based on Conventional Wisdom.

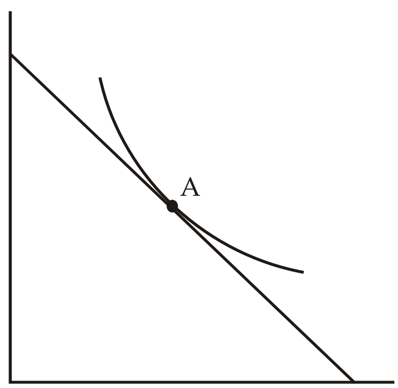
8. **Money and Inflation**. A Simple Overlapping Generations Model. The Demand for Real Money Balances. Monetary Equilibrium. The Welfare-Enhancing Role of Monetary Exchange. Money Neutrality and Non-Neutrality. Inflation, Seigniorage, and War Finance. The Laffer Curve. A Simple Model of Money and Banking. Government versus Private Supply of Money. International Monetary Systems. Nominal Exchange Rate Indeterminacy. Multilateral and Unilateral Fixed Exchange Rate Regimes. Speculative Attacks. Monetary Union.

9. **Economic Growth and Development**. Long-Run Evidence. A Malthusian Model with Fertility Choice. Children as an Investment Good. Private Property Rights and Asset Markets. Accounting for the Great Transition. Recent Development Patterns. Special Interests.

Preface

There is today a large discrepancy in the way macroeconomic theory is practiced by researchers and in the way it is taught to undergraduates. I am not exactly sure why this is the case. Perhaps some instructors–trained in older methods–feel that modern macroeconomic theory, with its insistence on microeconomic rigor, obscures the forest for the trees. Perhaps some instructors–trained in modern methods–feel that the modern approach is simply too complicated (mathematical) for the average undergraduate to absorb. Whatever the reasons, I believe that they are all wrong.

While modern methods can indeed be complicated, the basic economic forces embedded within a (say) dynamic, stochastic, general equilibrium model can be expressed in a very simple and intuitive manner. Moreover, this can be done largely with the aid of diagrams of the form:



This is not a complicated diagram. In mathematical terms, it requires nothing more than highschool alegebra to analyze formally (if the simple calculus is first provided–something I would recommend for introductory or intermediate level courses). For better or worse (and I would argue the former), this simple diagram summarizes the essential ingredients of any economic theory. And yet, its presence is strangely absent in the way macroeconomics is commonly taught. There is no reason for why this should be the case.

The diagram above represents ideas that are familiar to any economist. The first is that people have **preferences** defined over different commodities; and they they are generally **willing to substitute** across commodities. The second is that people face **constraints**; and that these constraints dictate the **ability to substitute** across commodities. Individual behavior is presumed to reflect the interaction between preferences and constraints. Aggregate behavior constitutes a collection of individual behaviors that are in some sense consistent with each other. That is, we construct Leviathan (cover) by adding up his individual pieces.

This simple way of organizing thought is at the same time intuitive and remarkably flexible. It is intuitive because almost anyone can relate to the idea that incentives influence behavior. It is flexible because "commodities" can be defined in any number of ways; so that the diagram can be brought to bear on almost any economic phenomenon. For example, I use this diagram to develop theories of labor supply, consumption and saving, money demand, and fertility choice.

There is one other great benefit associated with the diagram. In particular, because it makes explicit reference to peoples' preferences, it can be used to evaluate the **welfare consequences** of government policy. Exactly how one is to do this without reference to individual preferences escapes me entirely. Moreover, by being explicit about individual preferences, one quickly learns to be cautious in terms of relating conventional measures of macroeconomic activity to any sensible notion of social welfare. A higher GDP or a higher current account surplus, for example, do not necessarily imply a higher level of welfare; and Leviathan may find it optimal to allow for some positive level of unemployment.

I should like to offer a response to a criticism that one sometimes hears in relation to the modern macroeconomic analysis. Throughout the text, I make liberal (but not exclusive) use of to so-called "representative agent" assumption. This I do primarily for pedagogical purposes. In fact, the modern approach in no way depends on the representative agent formulation (there is nothing that prevents one from introducing as much heterogeneity as desired). While some conclusions are no doubt sensitive to the assumption, there are many others that are not (and it is these latter conclusions that deserve emphasis). In any case, I find it ironic that the criticism of this abstraction is most often leveled by those who prefer the older methodology; which, by its very nature, is typically cast in terms of a representative agent (e.g., in the form of aggregate behavioral relationships that make no reference to individual differences).

There is something else that I should point out. I make virtually no attempt in the text to describe what has become a highly influential branch of modern macroeconomic theory; the so-called New-Keynesian approach. This is regretable primarily because so much of what we read every day in the newspaper pertaining to monetary policy appears to be couched in this language. Nevertheless, I choose to omit it for the following reasons. First, I am not a big fan of the approach (there are better ways, in my view, to investigate the key questions in monetary policy). Sticky prices and wages may be important at some level, but probably do not factor into the "big" economic questions. And I still do not know what an "inflation shock" is supposed to represent in reality. Second, there are already many textbooks out there that do an adequate job of explaining the approach, so that there is no need to repeat things here.

On the other hand, I do take the time to take seriously some key Keynesian insights (Keynes was not a New-Keynesian). Contrary to what some may believe, the modern approach can be used to make precise Keynes' notion of coordiation failure and animal spirits. In short, the modern approach does not preclude the possibility that macroeconomic stabilization policies are in some way desirable. Being explicit about the circumstances in which this may be the case, however, makes clear the assumptions required to generate the result (and the real world limitations that are likely to impinge on policymakers in designing their policies).

At this stage, I would like to thank all my past students who had to suffer through preliminary versions of these notes. Their sharp comments have contributed to a much improved text. I would especially like to thank Sultan Orazbayez and Dana Delorme, both of whom have spent long hours documenting and correcting the typographical errors in earlier drafts. Thoughtful comments were also received from Bob Delorme, Janet Hua Jiang, and many others that are too numerous to mention here. Thank you all very much.

David Andolfatto

Burnaby, British Columbia

December, 2007

CHAPTER 1

The Gross Domestic Product

1. Introduction

The Gross Domestic Product (GDP) is an economic statistic that one hears quoted frequently in the news and elsewhere. But what exactly is this GDP thing supposed to measure? And why should anyone care about whether it is measured at all?

Most people have at least a vague idea that the GDP represents some measure of 'economic performance.' One often hears, for example, that a country with a higher GDP is performing better than one with a lower GDP; or that a rapidly growing GDP is better than a stable, or declining GDP. This idea of the GDP as a measure of economic performance is held so widely and (at times) accepted so uncritically, that on these grounds alone, it probably deserves closer scrutiny.

Before we can talk sensibly about GDP and why it might matter, we should have a clear understanding of how the term is defined and measured. Most countries in the world have a government agency (or agencies) responsible for collecting and aggregating measures of economic activity. You can find a list of these agencies at the following website (the United Nations Statistics Division):

http://unstats.un.org/unsd/methods/inter-natlinks/sd_natstat.htm

In Canada, our national statistical agency is called *Statistics Canada*.[1] Among other things, Statistics Canada maintains a system of national **Income and Expenditure Accounts** (IEA). The following quote, taken from the Statistics Canada webpage, sums up their own (somewhat naive) view of the world:

> The Income and Expenditure Accounts are the centre of macroeconomic analysis and policy-making in Canada. They are a means by which Canadians can view and assess the performance of the national economy. The accounts provide both a planning framework for governments and a report card on the results of the plans that governments carry out. At the core of the Income and Expenditure Accounts (IEA) is the concept of Gross domestic product (GDP) and its components.

[1] See: http://www.statcan.ca/

The statement above makes clear that GDP is considered a core concept. So let's take some time to investigate its measurement and potential usefulness.

2. Definition of GDP

Here is a standard definition of GDP:

GDP: The total value of final goods and services produced in the domestic economy over some given period of time.

From this definition, we gather that the GDP represents some measure of the level of production in an economy. For this reason, the GDP is commonly referred to as **output**. Keep in mind that output constitutes a **flow** of goods and services. That is, it represents the value of what is produced over some given interval of time (e.g., a month, a quarter, or a year). Food, clothing, and shelter services produced over the course of a year all contribute to an economy's annual GDP.

Let us now examine the definition of GDP more carefully. Note first of all that the GDP measures the 'value' of output. We will discuss the concept of 'value' in some detail later on; but for now, assume that value is measured in units of 'dollars' (feel free to substitute your favorite currency). When output is measured in units of money, it is referred to as the **nominal GDP** (at current prices).

Output takes the form of **goods and services**. What is the difference between a good and a service? A good is an object that can be held as inventory; while a service is an object that cannot be stored. Think of the difference between an orange and a haircut. Note that any good is likely valued only to the extent that it yields (or is expected to yield) a service flow; as when I consume that orange, for example.

Next, note that the definition above makes reference to **final** goods and services. A final good is to be distinguished from an **intermediate** good. An intermediate good is an object that is produced and utilized as a input toward the production of some other good or service *within the time period of consideration.*

An example may help clarify. Imagine that last year, an economy produced \$200 of vegetables, \$150 of fertilizer, \$100 of bread, and \$50 of flour. Imagine further that all of the fertilizer was used in the production of vegetables and all of the flour was used in the production of bread. One might be tempted to conclude that the annual GDP for this economy is \$500, but this would be wrong. In fact, the GDP is equal to \$200 + \$100 = \$300; that is, the total value of the final goods produced (bread and vegetables). The value of the intermediate goods is excluded from this calculation because their value is already embedded in the value of the final goods. That is, when you pay \$2.00 for a bundle of carrots at the supermarket, \$1.50 represents the value of the fertilizer that was used to grow these carrots.

The example above suggests an alternative (but equivalent) definition of GDP. Define **value-added** as the value of a good or service net of the cost of any intermediate inputs used to produce it. Then one can define the GDP as the total valued-added. In the context of the example above, we have \$150 of fertilizer and \$50 of flour (both of which use no intermediate inputs), together with the value added in the production of vegetables (\$50) and bread (\$50); the sum of which is \$300.

Moving along, observe that the definition above makes reference to the **domestic** product. The domestic product is to be distinguished from the **national** product; otherwise known as the **Gross National Product** (GNP). The difference is as follows. The GDP measures the value of output produced within the borders of a domestic economy, whether or not this production takes place with foreign-owned factors of production. The GNP, on the other hand, measures the value of output produced by the factors of production owned by the 'citizens' of a domestic economy, whether or not these factors of production reside on domestic soil or not. For countries like Canada and the United States, the difference between GDP and GNP is relatively small. For countries like Turkey and Mexico, on the other hand–with many citizens living and working abroad–the discrepancy between GDP and GNP can be significant.

Finally, consider the term **gross** in the definition of GDP. Here, the gross domestic product is to be distinguished from the **net** domestic product (NDP). The NDP is defined as the *GDP net of capital consumption.* **Capital consumption** simply refers to the value of capital that is consumed (i.e., destroyed or depreciated) in the act of production.

A case could be made that the NDP is a better measure of actual production. For example, if construction workers destroy \$20,000 worth of equipment in the process of building a \$100,000 house, most people would probably agree that \$80,000 constitutes a better measure of the value added to the economy. Environmentalists are particularly fond of an NDP measure modified to include 'environmental degradation' and 'resource depletion' as components of capital depreciation.

Exercise 1.1: Consider the example above, of an economy that produces \$200 in vegetables, \$150 in fertilizer, \$100 in bread, and \$50 in flour over the course of one year. As before, assume that the entire amount of fertilizer and flour is consumed in the process of producing vegetables and bread. But imagine now that the fertilizer and flour were not produced this year; that is, suppose that they were produced last year and brought over into this year as inventory (capital). Does the fertilizer and flour in this example fit the definition of an intermediate good? Compute the GDP and NDP for this economy.

3. Consumption vs. Investment

Consider all the millions of goods and services produced in an economy over the course of some time interval. Economists have found it useful to divide this vast and heterogeneous flow of output into two categories: **consumption** goods (and services); and **investment** goods (and services). What are the distinguishing characteristics of these two types of output and why is it useful to make such a distinction?

Consumption represents that part of the output flow that is consumed (i.e., destroyed) for the purpose of augmenting **current** material living standards. By 'current,' I mean over the course of a given time-interval, like one month or one year. Investment represents that part of the output flow that is destined to augment the **future** production of output (and ultimately, future material living standards). Investment is sometimes also referred to as the production of **new** capital goods and services. Note that new capital goods are to be distinguished from **old** capital goods–or, the **stock** of existing capital–which is presumably employed in the production of current output (along with other factors of production).

Consumption: That branch of the output flow that is consumed (destroyed) for the purpose of augmenting current material living standards.

One should keep in mind that the distinction between consumption and investment is not always so clear-cut; and in particular, the distinction may depend on the time-interval under consideration. Suppose, for example, that I hire the kid next door to mow my lawn one sunny afternoon. The kid's labor (together with his lawnmower capital) is used to produce 'lawn-enhancement services.' I pay the kid $10 and let him have a sip of my beer (when his mother isn't looking). Now, this $10 in lawn-enhancement services–does it constitute output in the form of consumption or investment?

The answer depends on the time-interval under consideration. Imagine that the benefit I derive from the freshly-cut lawn lasts for one week (the lawn needs to be mowed again after this period of time). If the time-interval under consideration is one day, then one could well argue that the lawn-enhancement service constituted a form of investment that generated a flow of consumption for several periods (i.e., days). On the other hand, if the time-interval under consideration is one week (or more), then one might argue that the lawn-enhancement service simply constituted consumption (as the output depreciates fully after one week).

Thus, one way to distinguish to between consumption and investment is to fix the time-interval under consideration and ask whether newly-produced output is expected to last longer than this time interval. If the answer is no, then the output constitutes consumption. If the answer is yes, then the output constitutes a form of investment that augments the stock of existing capital (with capital generating a flow of future services).

Investment: That branch of the output flow that augments the existing stock of capital.

The most common time-intervals used in macroeconomic analysis are one quarter (3 months) and one year. With these time-intervals, a large class of goods and services can be clearly categorized as either consumption or investment. The construction of a new house, a new piece of machinery, a new car, for example, would seem to constitute new capital goods (additions to the existing stock of capital). The same could be said of many medical procedures (from hernia operations to breast implants) and education services (to the extent that students can be expected to hold on to what they have learned beyond the final exam). In contrast, the production of perishable food products, transportation services, haircuts, shelter services, etc., would seem to constitute consumption. However, there remain other forms of output that are not so easily classified; for these objects, a judgement call must be made.

It is important to keep in mind that the term 'investment' as it is used here differs from the way it is commonly used in everyday language. Imagine, for example, that you are currently living in a rented apartment but decide to purchase a home. Most people would regard this purchase as an 'investment' in real estate. But whether this purchase is counted as investment in the IEA depends on whether the home you purchased is an old home or a new home. An old home is considered to be part of the existing (residential) capital stock. The purchase of an old home simply represents a change of ownership in the existing stock of capital and hence is not counted as investment for the economy as a whole. A new home, on the other hand, represents a new addition to the existing stock of residential capital; a new home is counted as investment for the economy as a whole.

The IEA definition of investment generally differs from the way people commonly understand the term.

Why is the distinction between consumption and investment important? This distinction is important because the manner in which an economy divides its output flow across consumption and investment ultimately determines the 'long-run' living standards of its inhabitants. Consider, for example, an economy consisting of farmers producing perishable food products year after year at some given level. This level of production determines living standards now and off into the indefinite future. Suppose now that some of these farmers become construction workers employed in the production of greenhouses and irrigation systems. This diversion of labor necessarily entails a temporary decline in food production (and hence current living standards) as the new capital that is constructed takes time to be productive. But in the 'long-run,' the new capital has the effect of enhancing agricultural output and hence future living standards beyond the initial level. In this way, investment entails a sacrifice of current consumption in exchange for higher levels of future consumption.

On the Concept and Measurement of Capital

The term **capital** appears to mean different things to different people. To an economist, capital refers to a **durable** factor of production and **inventory**. It does not, in particular, refer to financial assets, which simply represent claims to future objects.

The most obvious form of capital is what is called **physical capital**. Examples of physical capital include objects like machinery, computers, buildings, land, automobiles, highways, sewage systems, and inventories of goods. Even physical objects such as these are difficult to measure. For example, in measuring the stock of residential capital, is it appropriate to count the number of houses and apartments, or the square footage of living space? And is a 3000 square foot home made of brick the same thing as a 3000 square foot home made out of rice paper? Furthermore, how does one add together a house and a printing press to arrive at an aggregate measure of physical capital?

In principle, perhaps the only way to measure capital consistently is by market value. In this way, we could say that a brick home valued at $400,000 and a printing press valued at $100,000 together make up $500,000 worth of capital. The 'problem' with this approach is that the value of capital may vary even without any change in its physical quantity. On the other hand, perhaps it makes sense to think of more valuable capital as constituting more 'effective' capital.[2]

In practice, it appears that the aggregate stock of physical capital is measured using a 'perpetual inventory method.' The way this is done is to take the investment flow (measured at market value) in each of various asset classes, applying a constant depreciation rate (that varies with asset class) and then adding the results across investment years and asset classes. A major problem with this method is that it values the existing capital stock at **book value** (historical market value of past investment) instead of current market value. Presumably, this is done because it is difficult to ascertain the current market value of all forms of capital.

Unfortunately, the measurement of capital problem is in fact much worse than this. The reason for this lies in the fact that physical capital is not the only–or even most important–form of capital in an economy. One could reason

[2] See, for example: Black, Fischer (1995), *Exploring General Equilibrium*, The MIT Press, London, England (pp. 31–34).

by analogy that every human being is a kind of durable machine. In a very real sense, each person can be thought of as an owner-operator of **human capital** that generates a (potential) stream of labor services over time.

Like a machine or a house, we need to be maintained and (at times) repaired. In the IEA, maintenance and repair is considered a form of investment. In contrast, all personal expenditures on food, clothing, shelter and medical services are treated as consumption. Does this make sense? This is to say nothing about the investments that people make to improve their skill (again, education is treated as a form of consumption instead of investment), or the investments that parents make in raising their children.

What is the value of human capital? In theory, the market value of a person's human capital is the present value of one's lifetime wage stream net of the present value of any direct investments in human capital (the same principle holds for the valuation of any form of capital). In practice, direct claims on human capital are rarely exchanged in markets, making the value of human capital more difficult (but not impossible) to estimate.

Another important form of 'intangible' capital exists in the form of (disembodied) **technology**. The modifier 'disembodied' here is used in reference to 'knowledge' or 'technological know-how' that exists separately from what may be embodied in either physical or human capital. Some examples that come to mind here may include things like 'organizational capital' (the way production activities and distribution networks are organized, or other 'best-practice' techniques), or even the introduction of new products (e.g., the sudden availability of computers is not the same thing as having more factory space).

As you may have guessed, the value of these 'intangible' objects is often difficult to measure. Nevertheless, this does not diminish their potential importance. Firms can and do spend significant resources toward 'figuring things out.' The most obvious example is R&D expenditure. R&D spending is clearly a form of investment, even if the value of what is produced by such spending is difficult to measure. But for some strange reason, R&D spending (and other forms of investment in 'intangible' capital) is not counted as investment in the IEA. It is, however, counted as part of the GDP. Implicitly then, R&D spending is counted as form of consumption.

The upshot of all this is that the IEA essentially ignores human and intangible capital, so that care must be exercised in interpreting the measured stock of capital as the true stock of productive capital. Likewise, one must be careful in interpreting measured investment as reflecting the true level of investment in an economy (e.g., see Figure 1.1).

Figure 1.1

Big Companies Go Intangible

Companies are putting more emphasis on R&D and less on capital investment. Since 2000, the "intangibility index"—the ratio of R&D to capital spending, multiplied by 100—has risen for 9 of the 10 biggest U.S. companies that report R&D

COMPANY	INTANGIBILITY INDEX* 2000	LATEST**
EXXONMOBIL	5.1	4.4
GE***	73.6	100.7
MICROSOFT	429.1	761.6
PROCTER & GAMBLE	62.9	89.0
PFIZER	211.0	295.4
JOHNSON & JOHNSON	183.8	239.2
ALTRIA	32.0	42.3
CHEVRONTEXACO	2.2	2.9
INTEL	58.4	88.4
IBM	95.6	129.9
ALL 10	56.8	79.1

OVERALL	PERCENTAGE CHANGE 2000-LATEST**
R&D SPENDING	+42.1%
CAPITAL SPENDING	+2.1

*Capital spending for oil companies includes expenditures for exploration as well.
**Latest year for which R&D and capital spending are both available.
***Excluding GE Capital Services

Data: Company reports, *BusinessWeek*

4. How the GDP is Calculated

> *Statistics are a lot like sausages: both are best enjoyed when one does not see how they are made.*

As the son of Italian immigrants, I have had the pleasure of participating in the process of sausage-making from the ground up (including the initial slaughter). Believe me, this is an experience you can definitely afford to miss (I have never looked at my breakfast sausage in quite the same way ever since). And while I have never worked at a statistical agency, I have friends who have. From what I can gather, they now view 'official statistics' much in the same way I view my breakfast sausage.

The IEA report two measures of GDP, both of which should add up to the same number in theory (but in practice differ by a relatively small number called a 'statistical discrepancy'). These two measures are based on two different approaches: an income approach, and an expenditure approach. Below, I discuss each approach in turn.

The Expenditure Approach

The expenditure approach to computing the GDP relies on the following fact: *Everything that is produced must also be purchased.* At first blush, this might seem like a strange thing to assert. What if, for example, a manufacturer produces an automobile that is not delivered to a dealer? In this case, the newly produced automobile is treated as a purchase of **inventory** by the manufacturer. The value of this inventory investment is based on the **market price** of the automobile (i.e., the market value of similar automobiles that are sold on the market).

Thus, the **expenditure approach** calculates GDP as *the total spending on all domestically-produced final goods and services.* The GDP calculated in this manner is sometimes referred to as the **Gross Domestic Expenditure** (GDE). Mathematically, this is done as follows. Let x_{it} denote the quantity of good (or service) i that is sold at date t at market-price p_{it}. Then the value of this expenditure (measured in dollars) is simply $p_{it}x_{it}$. If there are Q such goods and services, then the GDE is given by:

$$GDE_t \equiv \sum_{i=1}^{Q} p_{it}x_{it}. \tag{1}$$

Again, note that expenditures on intermediate goods and services are not included in this calculation (why not?).[3]

[3] The notation $x \equiv y$ means that x is *equivalent* (or *by definition*) equal to y. Note that

Actually, this is not quite how it's done. The way it is done in practice is to first compute total spending on all (newly produced) final goods and services, whether or not they are domestically produced. Of course, this figure will include expenditures on imports, which are goods and services that are not produced domestically. To arrive at the GDE then, one must subtract from this figure the total value of all imported goods and services.

Largely as the result of an historical accident, the national income and product accounts organize the expenditure components of the GDP into four broad categories that depend on the sector in which the expenditure is undertaken. This classification is somewhat arbitrary in that there is no unique way in which to define 'sector.' Nevertheless, the way this is done in practice is to define four sectors as follows: [1] the **household sector**; [2] the **business sector**; [3] the **government sector**; and [4] the **foreign sector**. Sometimes, [1] and [2] are combined to form the **domestic private sector**. The government sector includes all levels of government (i.e., federal, provincial, state, local, etc.). The foreign sector includes both foreign private and government agencies.

Having defined sectors in this way, let H_t denote household spending; let B_t denote business spending; let G_t denote government spending; and let X_t denote foreign sector spending (on domestically-produced goods and services). As the spending on H_t, B_t and G_t includes imports, we have to subtract off the value of these imports M_t to calculate spending on domestically-produced output. Using these expenditure categories, the GDE may equivalently be calculated as:

$$GDE_t \equiv H_t + B_t + G_t + X_t - M_t. \tag{2}$$

Since this is probably not your first macro class, you've likely seen something similar to (2) before. It doesn't quite look right though, does it? This is because every macroeconomics textbook in existence (that I am aware of) uses slightly different notation; and instead writes (2) in the following way:

$$GDE_t \equiv C_t + I_t + G_t + X_t - M_t. \tag{3}$$

Obviously, (2) and (3) are equivalent if we define $C_t \equiv H_t$ and $I_t \equiv B_t$. There is nothing wrong in using whatever notation we wish, as long as the notation does not detract from clear thinking.

Unfortunately, the widely-used notation in (3) *does* at times appear to detract from clear thinking. Let me explain. Get your hands on any macro text currently on the market and flip to the section on national income accounting. Now look for the expenditure identity (3). In the discussion that surrounds this identity, you will invariably find statements asserting that C_t denotes consumption expenditure and I_t denotes investment expenditure. These statements are misleading (a product of bad notation) because C_t in fact represents household

this is not the same as stating $x = y$. The latter is an *equation* while the former is an *identity*. For example, if x denotes 'supply' and y denotes 'demand,' then it is not true that $x \equiv y$. However, it is true that $x = y$ at a market-clearing price.

sector spending on both consumption and investment (in the form of durables and human capital investments) and I_t represents only one component of investment (i.e., each of C_t, G_t and X_t also include expenditures on investment goods and services).

Table 1.1 provides the expenditure components of GDP based on (2) for Canada. A number of observations are in order here. First, as remarked earlier, note that one should refrain from interpreting the subtotal C as consumption. That is, a significant component of household expenditures on goods are in the form of durable and semi-durable goods. Furthermore, one can guess that a significant component of the purchase of services is also in the form of investment services (broadly defined). Second, again as remarked earlier, note that the IEA only appears to count business sector spending I in the form of additions to physical capital (and not any investments in intangible capital). Third, note that a portion of government purchases G is in the form of new capital goods. It is probably the case, however, that some of the spending categorized as 'current' goods and services might be better labelled as investment. Fourth, note that expenditures on exports X generally consist of both consumption and investment goods and services.

Table 1.1 Gross Domestic Product: Expenditure-Based

Canada 2005 (millions of dollars)

Household Sector	
Durable and semi-durable goods	164,815
Non-durable goods	189,213
Services	407,934
Subtotal (C)	**761,962**
Business Sector	
Residential structures	89,595
Non-residential structures	63,938
Machinery and equipment	91,354
Inventory investment	9,469
Subtotal (I)	**254,356**
Government Sector	
Current goods and services	262,369
Investment	35,156
Subtotal (G)	**297,525**
Foreign Sector	
Exports of goods	443,401
Exports of services	64,855
Subtotal (X)	**518,256**
Imports of goods	386,749
Imports of services	77,281
Deduct: Subtotal (M)	**464,030**
Gross Domestic Expenditure	**1,368,069**

Source: Statistics Canada, CANSIM table 380-0017.

Exercise 1.2: Can you think of an example of an individual who might belong to (i.e., makes purchases that would be reflected in) each of the four sectors described in Table 1.1?

The Income Approach

The income approach to computing the GDP relies on the following fact: *Every purchase of a good or service must constitute income to the agent (or agency) selling it.* The GDP calculated in this manner is sometimes referred to as the **Gross Domestic Income** (GDI). In principle, this can be done in any one of a number of ways.

The most obvious way to calculate aggregate income would be to define an individual (or household) as the basic economic unit. In general, each individual has multiple sources of income, including income on domestically-employed human capital w (wages), income on domestically-employed capital d (dividends, retained earnings, interest income on bonds), and income on assets employed in the foreign sector, f (this could include wages earned outside of the country). In the computation of GDP, this latter source of income is left out (although, it is included in the measure of GNP). Thus, if there are N domestic residents, one could compute:

$$GDI_t \equiv \sum_{i=1}^{N}(w_{it} + d_{it}). \tag{4}$$

Note that taxes and transfers to and from the government are not included in this measure. Why is this? Let's think about it. Suppose that the government collects taxes τ_{it} from individual i. Note that this tax measure includes taxes from all sources, including sales taxes, property taxes, and income taxes (including corporate income, at least, on that fraction of the domestic business sector owned by domestic residents). What does the government do with these taxes? It uses them to pay the wages and salaries of government sector employees, which shows up in w_{jt} for some government worker j. It also uses these taxes to make transfers to individuals in the economy. Note that these transfers do not count toward the GDP as they do not constitute any production of new good or service (they serve simply to redistribute output across members of society). Nevertheless, you should keep the following in mind:

> *The fact that government transfers are not counted in the measure of GDP does **not** imply that government transfer programs have no effect on GDP.*

When one computes the GDP in this manner, it is interesting to note that most of the income generated in the economy accrues to human capital. For many economies, the income share of human capital ranges between 65–75%.

Naturally, this common-sense way of reporting income is not the way the IEA typically does things. Instead, the IEA reports the breakdown in income according to Table 1.2:

Table 1.2 Gross Domestic Product: Income-Based
Canada 2005 (millions of dollars)

Wages and Salaries	678,925
Corporate Profits (before tax)	193,936
Government Business Enterprises (before tax)	13,370
Net Interest Income	61,240
Net Income (farm)	1,551
Net Income (unincorporated farm)	84,666
Inventory Valuation Adjustment	–442
Taxes less subsidies on factors of production	59,961
NDP (at factor cost)	1,093,207
Taxes less subsidies on products	94,750
Capital Cost Allowance	181,427
Gross Domestic Income	**1,369,384**

Source: Statistics Canada, CANSIM table 380-0001.

As you can see from Table 1.2, it is somewhat cryptic (not to mention, amusing). Let's try to figure out what's going on here.

First, I am guessing that *wages and salaries* refers to wage income net of taxes. The reason I believe this is because there is also a category called *taxes less subsidies on factors of production*. Presumably, factors of production here refers only to human capital, as I notice that the income generated by the physical capital owned by the business and government sectors is reported on an after-tax basis. Note that corporate profits include earnings that are both retained and distributed as dividends, and is net of depreciation costs (which is why the capital cost allowance is added later on).

There is an issue here as to how the income of unincorporated (non-farm) businesses is treated. Presumably, this is just lumped in with wages and salaries, although properly speaking, at least a part of this should actually be recorded as capital income. It is rather amusing to see that there are separate categories for the farm sector (why?). Presumably, the net income from farms represents capital income. However, the net income of unincorporated farms includes both wage income and capital income (no attempt is made to separate out these components).

The NDP represents the net (of depreciation) domestic production. At *factor cost* means that the incomes were calculated net of 'indirect' taxes (like sales taxes). Thus, to compute the GDP, both capital consumption and indirect taxes must be added to the NDP figure.

The Income-Expenditure Identity

So far, we have established that $GDP \equiv GDI$ and $GDP \equiv GDE$. From these two equivalence relations, it follows that $GDE \equiv GDI$. In other words, aggregate expenditure is equivalent to aggregate income, each of which are equivalent to the value of aggregate production. Again, the way to understand why this must be true is as follows. First, any output that is produced must also be purchased (additions to inventory are treated purchases of new capital goods, or investment spending). Hence the value of production must (by definition) be equal to the value of spending. Second, since spending by one individual constitutes income for someone else, total spending must (by definition) be equal to total income.

The identity $GDI \equiv GDE$ is sometimes referred to as the *income-expenditure identity*. Letting Y denote the GDI, most introductory macroeconomic textbooks express the income-expenditure identity in the following way:

$$Y \equiv C + I + G + X - M. \tag{5}$$

Note that since the income-expenditure identity is an identity, it *always* holds true. A natural inclination is to suppose that since the identity is always true, one can use it to make *theoretical* or *predictive* statements. For example, the identity seems to suggest that an expansionary fiscal policy (an increase in G) must necessarily result in an increase in GDP (an increase in Y). In fact, *the income-expenditure identity implies no such thing.*

> *The Income-Expenditure Identity does* ***not*** *imply that an increase in G leads to an increase in Y.*

To understand why this is the case, what one must recognize is that an identity is *not a theory* about the way the world works. In particular, the income-expenditure identity is nothing more than a *description* of the world; i.e., it is simply categorizes GDP into to its expenditure components and then exploits the fact that total expenditure is *by construction* equivalent to total income. To make predictions or offer interpretations of the data, one must *necessarily* employ some type of theory. As we shall see later on, an increase in G *may* or *may not* lead to an increase in Y, depending on circumstances. But whether or not Y is predicted to rise or fall, the income-expenditure identity will always hold true.

5. Other Measurement Issues

We've already talked a bit about some of the measurement problems concerning the classification of output into its consumption and investment components. That discussion, however, was predicated on the assumption that the concept of 'output' was well-defined and consistently measured. As we dig deeper into the sausage-making machinery, however, we find that both of these assumptions need to be viewed with caution.

From the definition of GDP, we know–in principle, at least–that the GDP is supposed to represent some measure of the 'value' of what an economy produces in the way of (final) goods and services; or 'output,' for short. The easiest and most consistent way of aggregating the value of different goods (or factors of production) is by adding up their values on the basis of market prices–which are usually denominated in units of the national currency. Of course, such an exercise first presumes the existence of markets for different forms of output and factors of production; and second, presumes that prices and the quantities exchanged in these markets are somehow observable to national statistical agencies.

One might suppose that in the so-called developed world, that there a sufficient number of markets to price most forms of goods and services. But even in this best-case scenario, statisticians are confronted with a number of conceptual and practical problems.

The Government

Consider, for example, government 'purchases' of output. Some of this output is purchased from the private sector (e.g., military purchases from private defense contractors). But a large component of government 'purchases' is in the form of output that it produces itself and then transfers to the private sector at zero (or close to zero) prices; e.g., medical services, law enforcement services, education services, etc. How does one measure the market value of services such as these that are not sold on markets?

Since market prices do exist for private medical care and private education, one approach would involve trying to **impute** the value of this government production using the market prices of **close substitutes** available in the private sector. This method of imputation may not available for a large class of non-marketed goods and services, however (e.g., what is the market value of the services provided by the court system)? In practice, the way this is handled by the IEA is simply to assume that the market value of government production is equal to its factor cost. For example, if a judge is paid \$150,000 per annum by the government, then this \$150,000 figure is **assumed** to be the market value of the judicial services produced by the judge's labor input.

This method of imputing the market value of government production may not be a bad approximation at most times and for most countries. Nevertheless, one should be aware of the potential pitfalls with this method. To see what can go wrong, consider the following stark example. Imagine an economy in which labor is the only factor of production and assume that workers are either employed in the private or government sector. The private sector produces and sells (at market prices) output Y_p at the wage cost W_p, which yields profit $\Pi_p = Y_p - W_p$. The government produces output Y_g at the wage cost W_g and simply 'gives' Y_g away, financing its wage bill with a tax T.

Using conventional IEA methods, the imputed value of government production is calculated as $Y_g = W_g$. Calculating GDP by the expenditure approach in this economy would yield $GDE = Y_p + Y_g$. Calculating GDP by the income approach would yield $GDI = W_p + \Pi_p + W_g$. You should be able to convince yourself that $GDE = GDI$.

Now, imagine in fact that the market value of Y_g is zero; that is, government production is considered to be a complete waste. In this case, the *true* GDP in this economy is given by Y_p. However, the *measured* GDP continues to be Y_p+Y_g. In this case, government production should in fact be counted as a **transfer** of resources (from private sector employees to government sector employees); it should not be counted as adding value to the economy as a whole.

Note that this type of problem is absent in computing the GDP generated by the private sector. Suppose, for example, that the business sector makes a 'mistake' in producing output that nobody values. In this case, the market value of Y_p is equal to zero. The expenditure-based GDP number is now given by Y_g; and the income-based GDP number is now given by $W_p+\Pi_p+W_g$. Since $Y_p = 0$, the business sector now makes a loss (negative profits) equal to $\Pi_p = -W_p < 0$. Hence, $GDI = W_g$ and $GDE = Y_g$, with $GDI = GDE$ measuring the true value-added in this economy.

Home Production

Time-use studies reveal that out of the total amount of time available to individuals, only a relatively small fraction of this is devoted toward activities that produce goods and services that are sold in markets (or given away by governments). In other words, there is likely a significant amount of time (labor) devoted toward activities that generate **home production**, defined to be goods and services that are produced and consumed (or invested) within a household (and hence, not exchanged on any market). In addition to labor, households also generally have available a stock of capital that is likewise employed in home production (e.g., owner-occupied housing and consumer durables).

To see what sort of issues arise here, consider the following examples. Imagine that you own a house that you rent out at \$1500 per month. Then this \$1500 is counted toward the GDP, since it constitutes capital income for you (and generates \$1500 per month worth of shelter services, which is consumed by your tenants). Imagine now that you decide to move into this home. Then you no longer report \$1500 in rental income on your tax return. However, it still remains the case that the home is generating \$1500 worth of shelter services that are now consumed by yourself. Consistency demands that the IEA impute the market value of these shelter services as valued-added. In fact, this is what many statistical agencies do.

In practice, however, statistical agencies often treat the valuation of home production inconsistently. Consider, for example, what happens if you own and operate an automobile. Your purchase of the automobile is recorded as an expenditure on a consumer durable. This durable good generates a flow of transportation services, the value of which should be recorded as valued-added. In fact, this is typically not the case. In contrast, if a taxi company purchases the same automobile, it is treated as investment and the revenue the taxi company earns from this asset is recorded as capital income.

Similarly, consider two households, one of which is 'traditional,' and the other which is 'modern.' By a traditional household, I mean one in which the father goes to work and spends a considerable amount of time around the home engaged in activities like mowing the lawn, cleaning the gutters, painting the house, repairing his automobile, etc. As well, the mother stays at home, raising the kids, cooking meals, cleaning the house, etc. By a modern household, I mean one in which both parents are employed and contract out extensively for services that 'self-produced' by a traditional family (e.g., they hire a nanny to cook, clean, and look after the kids; they hire contractors to effect home renovations, car repairs, and maintain the home, etc.).

Imagine that the two households described above are more or less similar in age structure, number of kids, education level of the parents, and other attributes. Then the true valued-added generated by each household is likely to be similar. But the measured value-added of the modern household is likely to be much higher, as many of its time-use activities are formally exchanged on markets (together with the fact that the IEA does not impute a value to household production in the form of raising kids, home maintenance, etc.).

The simple example described above warns us to be careful in interpreting time-series evidence of the growth in GDP as reflecting an increased level of production. In particular, to the extent that 'household structure' changes over time (from traditional to modern), much of the growth in GDP may simply reflect a measurement phenomenon (rather than reflecting true growth in production). A similar caveat is in order when one is making cross-country comparisons of GDP, especially between developed and underdeveloped countries. Much of the output that is produced in underdeveloped countries is likely to take the form of home production and hence not counted toward official GDP measures.

Exercise 1.3: According to this website: **www.globalissues.org/ TradeRelated/ Facts.asp**, half the world (nearly 3 billion people) manage to live on less than \$2 a day. While these unfortunate souls are undoubtedly poor by any measure, explain why the \$2 a day figure likely overstates the true extent of their poverty.

One might argue that conceptually, the GDP should measure the market value of 'marketable' output, even if it is not actually 'marketed.' That is, if

I decide to clean the gutters of my home one fall morning, the IEA should (in principle) count this cleaning service toward the GDP. While I did not market out for this service, it is clearly a service that I could have contracted out for.

This then raises the question of whether there are goods or services that are 'non-marketable.' In fact, one could argue that there are. An immediate example that comes to mind is sleep. One can produce and consume sleep and individuals clearly value sleep (a minimum amount of which is necessary to maintain the health of one's human capital). It is impossible, however, to contract out for sleep services; i.e., I cannot get someone to sleep for me. The same might be said of learning. You cannot get someone to learn the contents of this text for you. Most forms of leisure activities appear to be non-marketable as well; e.g., having someone take a vacation on my behalf just doesn't seem right.

Of course, one could–in principle, at least–attempt to impute a market value for non-marketable production as well. For example, the value of the time one spends producing leisure could be valued at the opportunity cost of this time (i.e., the wage that is foregone by consuming time in the form of leisure). For better or worse, the production/consumption of leisure is not viewed as contributing to GDP (even conceptually). Since people obviously do value leisure, however, the GDP cannot be considered the sole determinant of what determines individual well-being.

The Underground Economy

The underground economy refers to economic activity that is beyond the scope of government regulation and measurement. Underground activity typically takes place in well-defined markets, so that it is relatively easy to measure the market prices goods and services transacted in these markets. It is more difficult–if not impossible–however, to measure the volume of transactions, since they are purposely hidden. Underground activities may either be legal or illegal. If they are legal, they are hidden primarily for the purpose of evading taxes. For example, if you would like drywall installed in your basement, a drywall contractor may offer you two prices depending on whether you are willing to pay by cash or cheque (you will get a cheaper price if you pay with cash). If they are illegal, they are obviously hidden to avoid legal ramifications. In some jurisdictions, for example, purchasing and selling sex and certain forms of drugs is illegal.

By their very nature, underground economies are difficult to measure, so that any estimate of their size is necessarily imprecise. Nevertheless, some estimates do exist. According to one *Economist Magazine* article, for example, the underground economy in Italy is estimated to be between !5-27% of (measured) GDP.[4] According to this article, underground businesses are widespread

[4] www.economist.com/countries/Italy/profile.cfm?folder=Profile-Economic%20Structure

in agriculture, construction and services.

Once again, the presence of such unmeasured output should lead us to view official GDP numbers with a fair amount of caution as they likely understate the true value of production by a considerable margin. On the other hand, while the level of GDP may be understated, its growth rate may not be–at least, to the extent that underground (and other unmeasured) activity remains a relatively constant proportion of measured activity. Likewise, cross-country comparisons of GDP are likely to be more meaningful if the set of measured activities is more or less the same. We have no *a priori* reason, however, to believe that either of these conditions are met in reality.

Exercise 1.4: Imagine that a government suddenly enacts into legislation an oppressive tax regime on its citizens (e.g., a 100% tax on all income). Explain why the economy's measured GDP is likely to fall by much more than the true level of GDP.

6. Nominal versus Real GDP

GDP was defined above as the value of output (income or expenditure). The definition did not, however, specify in which units 'value' is to be measured. In everyday life, the value of goods and services is usually stated in terms of *market prices* measured in units of the national currency (e.g., Canadian dollars). For example, the dozen bottles of beer you drank at last night's student social cost you \$36 (and possibly a hangover). The 30 hours you worked last week cost your employer \$300; and so on. If we add up incomes and expenditures in this manner, we arrive at a GDP figure measured in units of money; this measure is called the **nominal GDP**.

If market prices (including nominal exchange rates) remained constant over time, then the nominal GDP would make comparisons of GDP across time and countries an easy task (subject to the caveats outlined above). Unfortunately, as far as measurement issues are concerned, market prices do not remain constant over time. So why is this a problem?

The value of either income or expenditure is measured as the product of prices (measured in units of money) and quantities. It seems reasonable to suppose that material living standards are somehow related to quantities; and not the value of these quantities measured in money terms. In most economies (with some notable exceptions), the general level of prices tends to grow over time; such a phenomenon is known as **inflation**. When inflation is a feature of the economic environment, the nominal GDP will rise even if the quantities of production remain unchanged over time. For example, consider an economy that produces nothing but bread and that year after year, bread production is equal to 100 loaves. Suppose that the price of bread ten years ago was equal to \$1.00 per loaf, so that the nominal GDP then was equal to \$100. Suppose further that the price of bread has risen by 10% per annum over the last ten years. The nominal GDP after ten years is then given by $(1.10)^{10}(\$100) = \$260$. Observe that while the nominal GDP is 2.6 times higher than it was ten years ago, the 'real' GDP (the stuff that people presumably care about) has remained constant over time.

Thus, while measuring value in units of money is convenient, it is also problematic as far as measuring material living standards. But if we can no longer rely on market prices denominated in money to give us a common unit of measurement, then how are we to measure the value of an economy's output? If an economy simply produced one type of good (as in our example above), then the answer is simple: Measure value in units of the good produced (e.g., 100 loaves of bread). In reality, however, economies typically produce a wide assortment of goods and services. It would make little sense to simply add up the level of individual quantities produced; for example, 100 loaves of bread, plus 3 tractors, and 12 haircuts does not add up to anything that we can make sense of.

So we return to the question of how to measure 'value.' As it turns out, there is no unique way to measure value. How one chooses to measure things depends on the type of 'ruler' one applies to the measurement. For example, consider the distance between New York and Paris. How does one measure distance? In the United States, long distances are measured in 'miles.' The distance between New York and Paris is 3635 miles. In France, long distances are measured in 'kilometers'. The distance between Paris and New York is 5851 kilometers. Thankfully, there is a fixed 'exchange rate' between kilometers and miles (1 mile is approximately 1.6 kilometers), so that both measures provide the same information. Just as importantly, there is a fixed exchange rate between miles *across* time (one mile ten years ago is the same as one mile today).

The phenomenon of inflation (or deflation) distorts the length of our measuring instrument (money) over time. Returning to our distance analogy, imagine that the government decides to increase the distance in a mile by 10% per year. While the distance between New York and Paris is currently 3635 miles, after ten years this distance will have grown to $(1.10)^{10}(3635) = 9451$ miles. Clearly, the increase in distance here is just an illusion (the 'real' distance has remained constant over time). Similarly, when there is an inflation, growth in the nominal GDP will give the illusion of rising living standards, even if 'real' living standards remain constant over time.

There are a number of different ways in which to deal with the measurement issues introduced by inflation. Here, I describe one approach that is commonly adopted by statistical agencies. Following the discussion surrounding the expenditure-based GDP measure, we have GDP given by (1); which I reproduce here for convenience:

$$GDE_t \equiv \sum_{i=1}^{Q} p_t^i x_t^i.$$

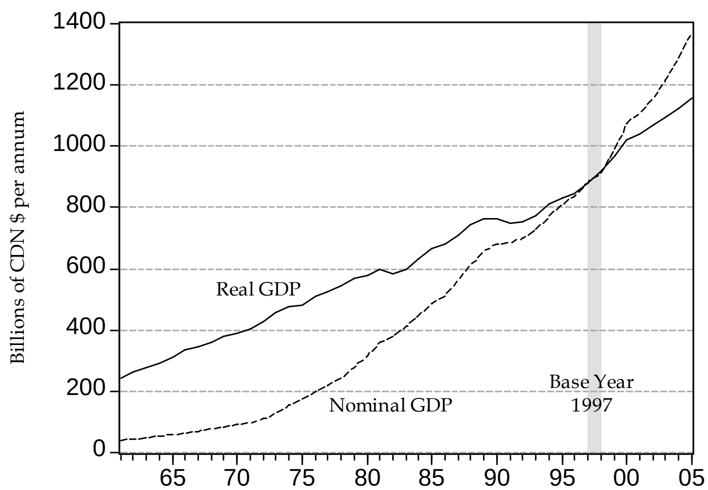
As this measure relies on current (i.e., date t) prices (whether actual or imputed), it is sometimes referred to as the **GDP measured at current prices**; or simply, the **nominal GDP**.

Now, choose one year arbitrarily (e.g., $t = 1997$) and call this the *base year*. Then, the **real GDP** (RGDP) for any year t is calculated according to the following formula:

$$RGDP_t \equiv \sum_{i=1}^{Q} p_{1997}^i x_t^i.$$

This measure is called the **GDP measured at base year prices**. In other words, the value of the GDP at date t is now measured in units of 1997 dollars (instead of current, or date t dollars). Note that by construction, $RGDP_{1997} \equiv GDE_{1997}$.

Figure 1.2
Nominal vs. Real GDP
Canada 1961-2005

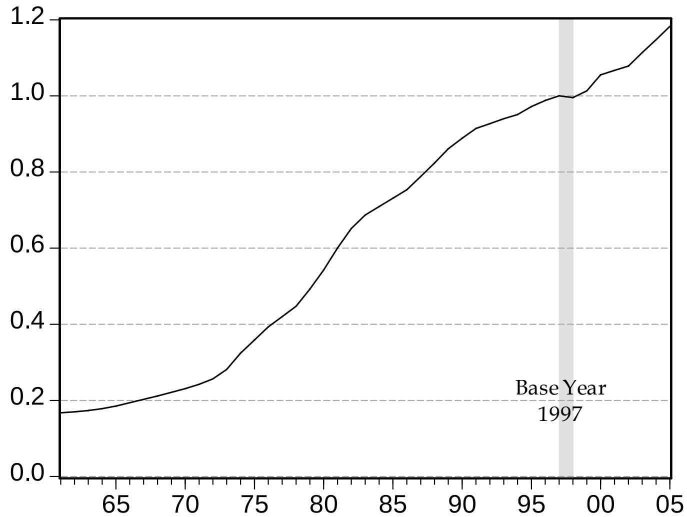


Source: CANSIM II; series V3860085 (real GDP); V646937 (nominal GDP).

As a by-product of this calculation, one can calculate the average level of prices (technically, the **GDP Deflator** or simply, the **price level**) P_t according to the formula:

$$P_t \equiv \frac{GDE_t}{RGDP_t}.$$

Figure 1.3
GDP Deflator (1997 = 1)
Canada 1961-2005



Note that the GDP deflator is simply an index number; i.e., it has no economic meaning (in particular, note that $P_{1997} = 1$ by construction). Nevertheless, the GDP deflator is useful for making comparisons in the price level across time. That is, even if $P_{1997} = 1$ and $P_{1998} = 1.10$ individually have no meaning, we can still compare these two numbers to make the statement that the 'average' level of prices rose by 10% between the years 1997 and 1998.

The methodology just described above is not fool-proof. In particular, the procedure of using base year prices to compute a measure of real GDP assumes that the structure of relative prices remains constant over time. To the extent that this is not true (it most certainly is not), then measures of the **growth rate** in real GDP can depend on the arbitrary choice of the base year.[5] Finally, it should be noted that making cross-country comparisons is complicated by the fact that nominal exchange rates tend to fluctuate over time as well. In principle, one can correct for variation in the exchange rate, but how well this is accomplished in practice remains an open question.

Real per capita GDP

In general, the real GDP of any economy may rise (or fall) owing to: [1] a

[5] Some statistical agencies have introduced various 'chain-weighting' procedures to mitigate this problem.

rise (or fall) in population; and/or [2] a rise (or fall) in the output produced per person. To get a sense of how material living standards for the 'average' person in an economy, it makes sense to divide an economy's total real GDP by population size. The resulting number is called the **real per capita GDP**; or more commonly, the **real per capita income**.

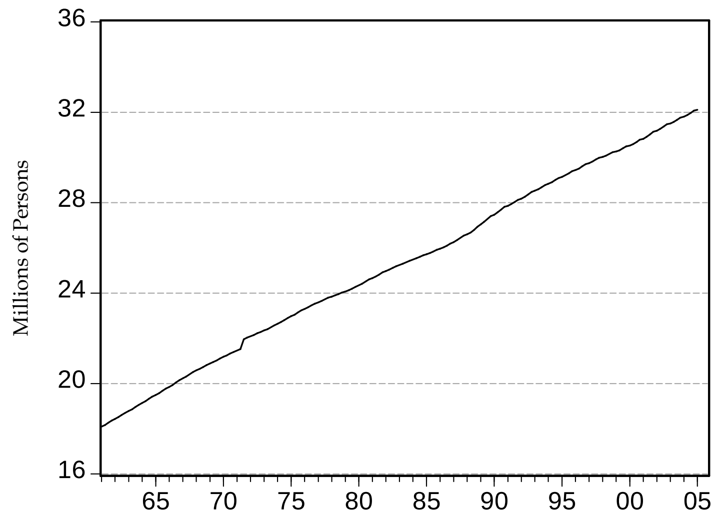
It cannot be stressed enough that extreme caution should be exercised in interpreting real per capita GDP as a measure of material living standards or as a measure of economic welfare. First, one should keep in mind all of the measurement issues discussed at length above. Second, as far as material living standards are concerned, theory suggests that consumption is likely to constitute a better measure (or perhaps even wealth, to the extent that consumption is related to wealth). One would not want to judge the material living standards of a student with zero income, for example, solely on the basis of his or her income. Finally, there is good reason to believe that economic welfare depends not only on one's consumption flow, but also on other things (e.g., leisure time spent with one's family and friends).[6]

With the appropriate caveats in place, let us examine the behavior of real per capita GDP for Canada. Figure 1.4 plots the evolution of Canada's population and Figure 1.5 plots the real per capita GDP.

---

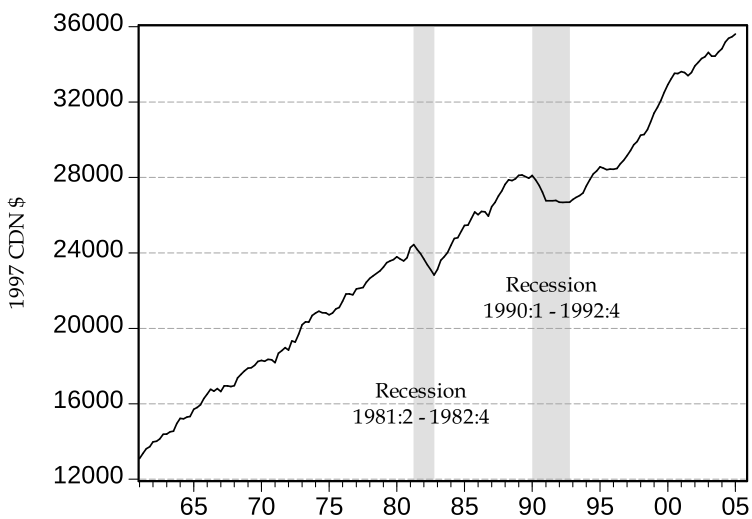
[6] For an interesting discussion on the history of work and leisure, see: Hill, Roger B. (1999). *History of Work Ethic*, www.coe.uga.edu/~rhill/workethic/hist.htm

Figure 1.4
Population of Canada
1961:1 - 2005:1



Source: Statistics Canada, table 510005

Figure 1.5
Real per capita GDP
Canada 1961:1 - 2005:1



Source: CANSIM II, Series V1992067 (divided by population).

7. Growth and Business Cycles

The pattern of economic development displayed in Figure 1.5 for Canada is typical among many countries, especially for those that occupy the so-called 'developed world.' The most striking feature of Figure 1.5 is the trend rate of growth in real per capita income. In 1961, income per capita was approximately \$13,100. By the end of the sample in 2005, income per capita had grown by a factor of 2.7 (to \$35,600). This represents an average annual growth of approximately 2.2%.

Now, 2.2% may not sound like a large number to you. And you may be tempted into thinking that it really does not matter very much whether an economy grows at 1.2%, 2.2%, or 3.2%. In fact, even seemingly small differences in long-run growth rates such as these can translate into huge differences in the level of income over time. The reason for this lies in the **power of compound interest**.

To appreciate the power of compound interest, imagine that there are three economies A, B, and C that are currently generating \$10,000 in per capita income. Economy A grows at $g = 1.2\%$, economy B at $g = 2.2\%$, and economy C at $g = 3.2\%$ per annum. What will be the level of per capita income at the end of 20 years? The answer is provided in Table 1.3.

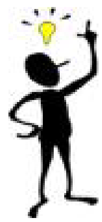
Table 1.3: Power of Compound Interest

	Initial GDP	GDP after 20 years	Change
Economy A ($g = 1.2$)	\$10,000	\$12,694	27%
Economy B ($g = 2.2$)	\$10,000	\$15,453	55%
Economy C ($g = 3.2$)	\$10,000	\$18,776	88%

A useful formula to keep in mind is the so-called **Rule of 72**. This rule states that if an economy grows at a rate of g% per annum, then the number of years it takes to double income is given by n:

$$n = \frac{72}{g}. \tag{6}$$

Thus, an economy that grows at 2% per annum will increase its living standards by 100% in 36 years. An economy that grows at 4% per annum will double its living standards in only 18 years.



Any factor (including government policy) that may affect an economy's long-run growth rate by even a small amount can ultimately result in very large differences in living standards over prolonged periods of time.

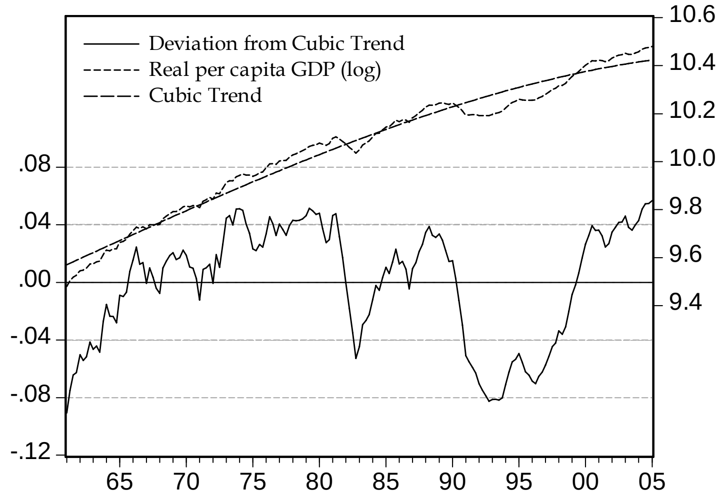
Since our current high living standards depend in large part on past growth, and since our future living standards (and those of our children) will depend on current and future growth rates, understanding the phenomenon of growth is of primary importance. The branch of macroeconomics concerned with the issue of long-run growth is called **growth theory**. A closely related branch of macroeconomics, which is concerned primarily with explaining the level and growth of incomes across countries, is called **development theory**.

Traditionally, macroeconomics has concerned itself more with the issue of 'short run' growth, or what is usually termed the **business cycle**. The business cycle refers to the cyclical fluctuations in GDP around its 'trend,' where trend may defined either in terms of levels or growth rates. From Figure 1.5, we see that while per capita GDP tends to rise over long periods of time (at least, in Canada and some other countries), the rate of growth over short periods of time can fluctuate substantially. In fact, there appear to be (relatively brief) periods of time when the real GDP actually falls (i.e., the growth rate is negative). When the real GDP falls for two or more consecutive quarters (six months), the economy is said to be in **recession** (i.e., the shaded regions in Figure 1.5).

It is important to keep in mind that while it is tempting to dichotomize a pattern of economic development like Figure 1.5 into its 'trend' and 'cycle' components, there is in fact no *a priori* reason to believe that such a decomposition makes any sense theoretically (although, one could certainly perform such a decomposition statistically). Because what I've said here is important and not widely appreciated, let me elaborate.

When viewing a diagram like Figure 1.5, the natural inclination is to draw a smooth line through the data and interpret this as 'trend.' The difference between the actual data and the trend line is then interpreted as 'cycle.' Unfortunately, there is no unique or obvious way to detrend time-series data. In Figures 1.6, I display the trend and cycle components of Canadian GDP using one particular method (a cubic trend).

Figure 1.6
Real per capita GDP
Cubic Trend and Cycle
Canada 1961:1 - 2005:1



There is nothing inherently wrong in detrending time-series data. The mistake that people commonly make, however, is to assume a smooth trend line–estimated on historical data–actually represents a trend that can be expected to prevail into the foreseeable future. To put things another way, people commonly make the mistake of assuming that a smooth trend line represents something 'real' or 'fundamental' about the way an economy functions. In fact, a smooth trend line may be nothing more than a statistical illusion.

To make what I am saying more concrete, consider the following argument. Let y_t denote the (log) real per capita GDP and assume that you and I know underlying **data generating process** (DGP) for the economy. Imagine further that this DGP is given by:

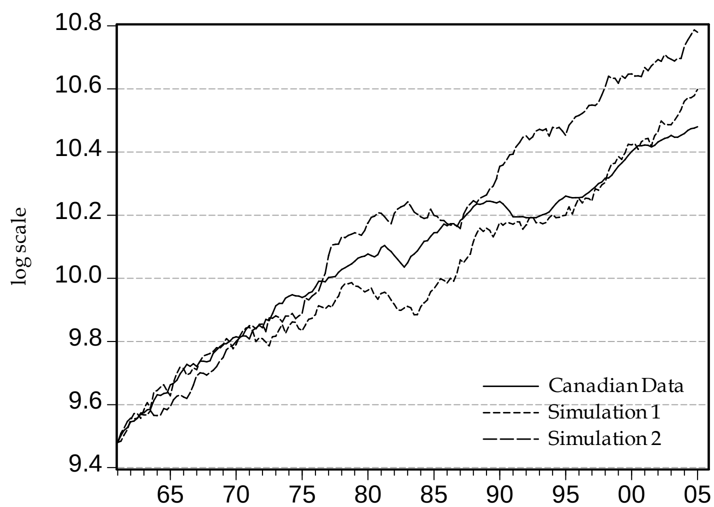
$$y_{t+1} = \gamma + y_t + e_{t+1}, \tag{7}$$

where e_{t+1} is a random variable, representing an unforecastable 'shock' to the economy's GDP. We can assume that e_{t+1} takes one of two values, each of which is determined by the flip of a coin. Or, to be slightly more sophisticated, we can assume that e_{t+1} is determined by a draw from a Normal distribution with mean μ and standard deviation σ. If we set $\mu = 0$, then the expected value of e_{t+1} as of date t is given by $E_t e_{t+1} = 0$. In fact, the expected value of e_{t+j} for any $j = 1, 2, 3, ...$ is given by $E_t e_{t+j} = 0$.

The DGP in (7) is an example of what is called a **random walk with drift**. The 'drift' parameter γ represents the expected rate of growth of (log) GDP. We can simulate an equation like (7) by assigning parameter values and generating the 'shocks' e_t using a random number generator. Suppose, for example, that we set $y_{1961:1} = 9.48$ (its value for the Canadian economy in 1961). Let $\gamma = 0.0055$ (which generates an annual expected growth rate of 2.2%. As well, choose a standard deviation $\sigma = 0.015$. Figure 1.7 displays two simulated time-series for y_t (along with the actual Canadian data).

One thing that should strike you from viewing Figure 1.7 is how the simulated series resemble the actual data; in fact, it would be very hard to tell (without knowing beforehand) which series was generated by our model (7) and which was generated by the economy.

Figure 1.7
Log GDP: Actual and Simulated



Now, imagine that we handed our simulated data to an econometrician and asked him or her to estimate a smooth trend line the way that was done in Figure 1.6. One could certainly do this; and we would be left with a diagram similar to Figure 1.6. But does the estimated trend line represent anything 'fundamental' about the manner in which our model economy functions? In particular, is there any reason to believe that future GDP levels will revert back to the estimated trend line? Can we use the estimated trend line to forecast the future level of GDP? The answers to these questions is *no*.

The reasons for why we can be so certain of this are twofold: [1] unlike the econometrician, we know the true DGP generating our simulated series; and [2] given that we know the true DGP, we can compute the theoretical trend and show that does not correspond to a smooth line.

The theoretical (i.e., true) trend behavior displayed by our model economy can be calculated as follows. Using (7), we can deduce that:

$$\begin{aligned} y_{t+2} &= \gamma + y_{t+1} + e_{t+2}; \\ &= 2\gamma + y_t + e_{t+1} + e_{t+2}. \end{aligned}$$

Similarly,

$$\begin{aligned} y_{t+3} &= \gamma + y_{t+2} + e_{t+3}; \\ &= 3\gamma + y_t + e_{t+1} + e_{t+2} + e_{t+3}. \end{aligned}$$

Continuing on in this way, we have:

$$y_{t+n} = n\gamma + y_t + e_{t+1} + e_{t+2} + ... + e_{t+n},$$

for any arbitrary $n > 0$.

Now, suppose we are at date t and wish to estimate the future level of GDP at date $t+N$, where N is some number large enough to be considered the 'long-run;' e.g., $N = 40$ quarters (ten years) off into the future. Then we can ask the question: what is the expected value of y_{t+N}, given what we know at date t? The answer is given by:

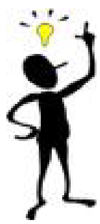
$$E_t y_{t+N} = N\gamma + y_t. \tag{8}$$

This 'long-run' ($N-$period ahead) forecast of GDP can be thought of the model's 'trend' for the (log) level of GDP. In other words, the 'long-run' level of GDP is determined by the **current** level of GDP y_t plus an expected growth term $N\gamma$.

Note that since y_t changes over time, so does the trend level of GDP. In fact, since y_t is a random variable, our model displays what is known as a **stochastic trend**. Needless to say, a stochastic (random) trend line is not going to look anything like the smooth deterministic trend line drawn in Figure 1.5. There is, in fact, no reason to believe that y_t will ever revert back to a smooth trend line estimated with historical data. The estimated trend line is simply a figment of the econometrician's imagination (we know this, but he doesn't).

Why is this important to understand? It is important for the following reasons. First, as mentioned earlier, there is a tendency for people to believe that a smoothly drawn line through the data represents something 'fundamental' about the way an economy functions in the long-run. In other words, there is a tendency to believe that an economy will eventually revert back to some given trend behavior. As the example above demonstrates, such a belief is not necessarily correct (and can easily be incorrect, given how similar the simulated series in Figure 1.7 resemble the actual data).

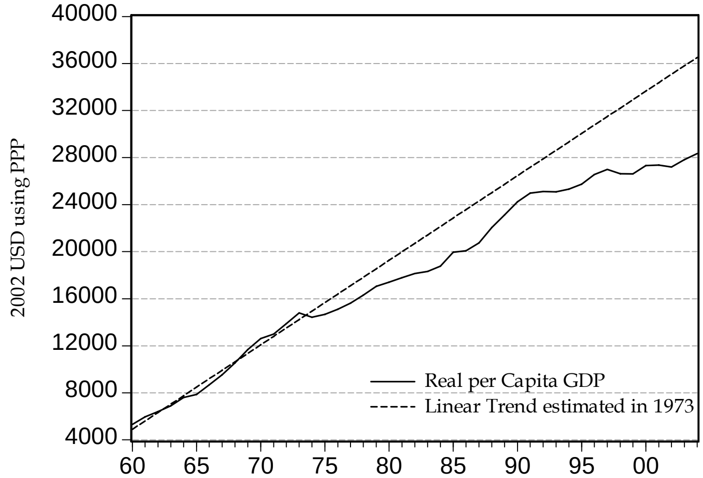
Second, by assuming that the economy does possess a smooth trend, one is implicitly assuming that growth and business cycle phenomena are independent of each other. That is, one is easily led to the conclusion that we can use a growth theory to understand trend behavior and a business cycle theory to understand deviations from trend, with each theory bearing no relation to one another. In fact, it may very well be the case that the so-called 'business cycle' is nothing more than a by-product of the process of economic development (as suggested by our model). In other words, we may be wrong in thinking (as people commonly do) that we can understand the business cycle without understanding the process of growth itself.



Fluctuations in GDP may largely be a by-product of a random growth process (a shifting trend); i.e., the 'business cycle' may be inextricably linked to the process of economic development itself.

The basic lesson here is to be careful in assuming that an economy has a smooth trend and that GDP will eventually return to this trend. To demonstrate the potential pitfall of this commonly held view, consider Figure 1.8, which plots the real per capita GDP for Japan from1960–2004. Imagine that an econometrician living in the year 1973 wants to estimate the 'trend' for Japanese GDP based on the historical data 1960-73. The dashed line (a simple linear trend) in Figure 1.8 appears to fit the historical data reasonably well. Unfortunately, the forecast of GDP far off into the future would have been off a tad.

Figure 1.8
Real per capita GDP
Japan 1960-2004



Source: U.S. Department of Labor BLS
www.bls.gov/fls/flsgdp.pdf

> *The trend is your friend...*
> *...until it ends.*

8. Schools of Thought

The reason for why aggregate economic activity fluctuates the way it does, even in relatively stable institutional environments, remains largely an unresolved puzzle. It should come as no surprise then to learn that there are many different hypotheses that offer different interpretations of observed patterns. At the end of the day, the lines of debate are drawn across the following two questions:

- *What are the primary **shocks** that are the ultimate source of aggregate fluctuations?*
- *What is the **mechanism** by which an economy responds to any given shock?*

Strictly speaking, a **shock** refers to a 'surprise' event that is determined by God or nature (i.e., an event that is beyond the control of any economic agent or agencies). A tsunami that wipes out a significant fraction of a region's stock of human and physical capital constitutes a possible example. The sudden appearance of new technology–like the internet–may be another.

Unfortunately, the interpretation of shock events is not entirely unambiguous. For example, some religious groups contend that the December 26, 2004 tsunami that afflicted southeast Asia was in fact brought forth by God as a punishment for the region's sins (sex and drugs). According to this interpretation, the inhabitants (and tourists) in southeast Asia brought the tsunami on by their own debaucherous behavior. This view requires that we take as **exogenous** (i.e., unexplained) two things: [1] God's law; and [2] preferences for debaucherous activities (that violate God's law). Nevertheless, one might still argue that while the tsunami itself should have been expected, the exact date of its arrival could not have been forecasted. Thus, the actual arrival of the tsunami is still usefully interpreted as a shock.

The same sort of argument can be made with respect to a technology shock. That is, let us take as exogenous two things: [1] the law of nature governing the process of discovery; and [2] preferences for higher living standards. Then one might reasonably argue that the idea behind the internet was in fact the product of human behavior (e.g., R&D activity). As with the tsunami, however, no one can reasonably be expected to forecast the exact arrival date of any technological advancement. When knowledge is discovered then, it comes as a shock.

In a sense, any economic theory constitutes an **explanation** of how a set of **endogenous** variables Y is determined in relation to a set of **exogenous** variables X. Thus, in abstract terms, any theory can be thought of taking the following form:

$$X \rightarrow_L Y,$$

where $\rightarrow_L$ denotes the logic underlying the explanation. A shock then can be thought of as some exogenous change in X and denoted ΔX. The theory then provides an explanation for the mechanism by which a change in X might be expected to influence the endogenous variables Y; i.e.,

$$\Delta X \rightarrow_L \Delta Y.$$

At issue then is what to include in X and how to think about ΔX (assuming that few people will argue with $\rightarrow_L$ or the form of logic to be used in connecting assumptions with predictions). Ultimately, one would hope for a theory that could explain everything in terms of an X that was 'truly' exogenous. Unfortunately, the nature of economics (and of science in general) is such that a 'grand unifying theory' of this form is unlikely to found anytime soon. In the meantime, we have to make due with what must be considered only 'partial' explanations that will (hopefully) be improved upon as the science progresses.

Thus, there is at present a long list of candidates for what might be included in ΔX. Some theories assert the existence of government spending shocks or monetary policy shocks, as if the behavior of the government or its affiliated agencies is beyond comprehension (i.e., determined by God or nature). The *Bank of Canada*, for example, emphasizes domestic shocks that arise from political uncertainty (e.g., will Québec separate from Canada or not) and international shocks like the 1973 OPEC oil crisis and the 1997 Asian financial crisis.[7] Economic commentators and analysts on television are fond of pointing to price shocks (e.g., the stock market, interest rates, inflation, exchange rates) as if these objects too are somehow not determined by conscious human behavior in reaction to more fundamental disturbances.

In many cases, it can make sense to view particular events such as sudden price changes or financial crisis as an exogenous shock, even if we know (or suspect) that prices and financial market behavior are not truly exogenous. For example, we may want to frame a question in the following way: How might the domestic economy react **given** the financial crisis in Asia? The answer to such a question, while useful for some purposes, is ultimately unsatisfying as it leaves unexplained the crisis itself. Economists have different hypotheses concerning the ultimate source of such disturbances, and these different views help define various schools of thought.

At the risk of oversimplifying, one might usefully categorize macroeconomic theory into two broad schools of thought, each of which is characterized primarily by the particular set of shocks and mechanisms that tend to be emphasized. I label the first school **conventional wisdom**, as variants of this view are held so widely among market analysts, politicians, central bankers, and a good part of the academic community. I label the second school **neoclassical**; this view is not nearly so widely-held, but is nevertheless influential among academic economists.

Conventional Wisdom

The conventional wisdom owes its intellectual debt primarily to the work of **John Maynard Keynes**, whose views on the business cycle were shaped to a large extent by the events of the Great Depression.[8] The primary legacy of this view is twofold: [1] that shocks are ultimately the result of exogenous changes in private sector expectations (*animal spirits*); and [2] that market economies are sufficiently dysfunctional as to make well-designed government stabilization policies desirable.

The way these ideas have evolved into conventional wisdom are as follows. First, growth is explained as the product of a relatively smooth process of technological development, so that one can infer from the data a relatively stable

[7] See: www.bankofcanada.ca/en/monetary_mod/factors/index.html

[8] Keynes, John M. (1936). *The General Theory of Employment*, Interest and Money, MacMillan, Cambridge University Press.

trend that determines an economy's 'long-run' fundamentals. The business cycle then, constitutes fluctuations around this trend (with movements in GDP eventually reverting back to trend). The trend level of GDP is sometimes referred to as **supply** or **potential GDP**, with the actual level of GDP referred to as **demand**.[9]

Having identified a relatively stable trend (supply) and then observing that actual GDP (demand) fluctuates around trend, one is led to the conclusion that business cycles are caused by **demand shocks** (i.e., unexplained and random changes in desired spending patterns emanating from various sectors of the domestic and foreign economy). While the root cause of these shocks is not usually discussed, it seems clear enough from the language used to describe them that they are thought to be the product of exogenous (and irrational) swings in market sector expectations (animal spirits). A strong quarter, for example, might be explained as resulting from the 'strength of the consumer;' which in turn may lie in the behavior of 'consumer confidence' (high expectations of future earnings). Similarly, business sector behavior may be described as being the product of 'irrational exuberance' (high expectations concerning the future return to investment).

To the extent that demand shocks are 'irrational,' they have adverse consequences that can last a long time. A bad investment today, for example, will have implications for GDP many periods into the future. These shocks are further exacerbated by various market imperfections–for example, in the form of 'sticky' nominal prices and wages–that prevent markets from adjusting rapidly to shocks (which explains why 'supply' is not usually equated to 'demand'). Given this interpretation of the cycle, it should come as no surprise that this view also advocates the use of various government stabilization policies (active monetary and fiscal policy) to mitigate the adverse consequences of the cycle.[10]

Neoclassical View

The neoclassical view is closer in spirit to those expressed by another great economist, **Joseph Schumpeter**.[11] The primary legacy of this view is that **technology shocks**–the very shocks that contribute to the general rise in living standards–are at the same time responsible for generating the fluctuations that are commonly interpreted as the business cycle.

[9] This language has even found its way into the IEA. For example, the expenditure-based measure of GDP is often labelled 'final demand.' The implication, of course, is that the final demand computed in this manner does not necessarily measure 'final supply.'

[10] It is interesting to note that Keyne's (1936) own views differ significantly from those that evolved from his work. In particular, while he emphasized the role of 'animal spirits,' he viewed these exogenous changes in expectations as being rational in the sense of constituting 'self-fulfilling prophesies.' Further, the concept of 'sticky' prices or wages played no role in his theory; except peripherally and as a mechanism that potentially *dampened* the adverse consequences of demand shocks.

[11] Schumpeter, Joseph A. (1939). *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*, New York, McGraw-Hill.

According to the neoclassical view then, the distinction between 'growth' and 'cycles' is largely an artificial one. Almost everyone agrees that long-run growth is the product of technological advancement. But unlike the conventional wisdom, which views trend growth as being relatively stable, the neoclassical view asserts that there is no God-given reason to believe that the process of technological advancement proceeds in such a 'smooth' manner. Indeed, it seems more reasonable to suppose that new technologies may appear in 'clusters' over time. These 'technology shocks' may cause fluctuations in the trend rate of growth through what Schumpeter called a process of **creative destruction**. That is, technological advancements that ultimately lead to higher productivity may, in the short run, induce cyclical adjustments as the economy 'restructures' (i.e., as resources flow from declining sectors to expanding sectors). Further, there is no guarantee that all new technologies work out exactly as planned or expected. What may have looked promising at one date, may in fact turn out to be a disaster later on (resulting in an observed negative technology shock).

As with the conventional wisdom, the neoclassical view admits that sudden changes in private sector expectations may lead to sudden changes in desired household and business sector spending. But unlike the conventional wisdom, these changes are interpreted as reflecting the 'rational' behavior of private sector decision-makers **in response** to perceived real changes in underlying economic fundamentals (i.e., technology shocks, or other real factors). In other words, changes in market sentiment are the *result* and not the *cause* of the business cycle. It is important to keep in mind when evaluating this perspective that the concept of 'rational' expectations does not imply that individuals never make 'mistakes.' It simply means that expectations are formed in the 'best' way possible, using whatever relevant information is currently at one's disposal. More often than not, actual outcomes will differ from those that are expected.

According to the neoclassical view, the business cycle is an unfortunate but largely unavoidable product of the process of economic development. Market imperfections play little or no role in exacerbating economic fluctuations; indeed, even a well-functioning 'planned' economy (if such an object were to exist) would exhibit similar fluctuations. Given this interpretation, it should come as no surprise to learn that the policy implication here is that government attempts to stabilize the cycle are likely to do more harm than good.

9. The Plan Ahead

In the chapters that follow, I plan to lay out–hopefully, in easily digestible bits and pieces–the various key elements that constitute modern macroeconomic theory. This endeavor is not meant to be an exercise in pure theory; throughout the book I try to demonstrate how the theory can be used to interpret and understand various aspects of real-world economies.

The book is designed so that, by the end of it, a conscientious (and patient) reader will have a reasonably good idea of basic theory (and how the various bits and pieces I present can ultimately be tied together in a more advanced theoretical treatment), together with an idea as to how modern macroeconomic theory can be applied toward interpretation, prediction, and the evaluation of policy.

Problems

1. While Americans constitute a relatively small fraction of the world's population (less than 5%), they spend approximately 20% of the world's income. This fact is sometimes used as evidence of American 'greed.' Provide a different interpretation of this fact based on your knowledge of the relationship between aggregate expenditure and output.

2. We often read that 'the consumer' drives the economy because consumption accounts for 60% of GDP. On the other hand, it is also true that 'the laborer' accounts for 75% of GDP; yet we seldom (if ever) hear of stories relating to how the GDP depends on the supply of labor. Why do you think this may be the case?

3. Why do you think it is important to distinguish between consumption and investment goods?

4. Explain why government transfers are not counted as a part of an economy's GDP (it will be useful to first define the GDP).

5. Explain why 'overpaid' government employees will lead to an overstatement of GDP, whereas 'overpaid' private sector employees will not.

6. Explain the conceptual difference between a statistical trend and a theoretical trend.

7. Consider two economies A and B that each have a real per capita GDP equal to \$1,000 in the year 1900. Suppose that economy A grows at 2% per annum, while economy B grows at 1.5% per annum. The difference in growth rates does not seem very large, but compute the GDP in these two economies for the year 2000. In percentage terms, how much higher is the real GDP in economy A compared to economy B?

CHAPTER 2

Output and Employment

1. Introduction

A central feature of the business cycle is the **comovement** between output (real per capita GDP) and employment (or hours allocated to work activities per capita). In the short-run, output and employment tend to move in the same direction. In fact, much of the change in GDP over the cycle is attributable to changes in the level of employment. This makes a lot of sense since, as more individuals work to produce output (or as employed individuals work longer hours), one would expect the level of output to increase. But understanding this fact alone does not help us understand the business cycle, as it does not explain **why** employment should change in the first place.

Most economists would probably agree that the cyclical variation in employment is driven by fluctuations in the **demand for labor**. There is, unfortunately, considerably less agreement on what forces are responsible for generating the volatility in labor demand. Since labor productivity and real wages tend to by **procylical** (i.e., move in the same direction as GDP), some economists stress the role played by **productivity shocks** (recall the discussion in Chapter 1). The basic idea here is that temporal variation in the productivity of labor is a natural phenomenon in a growing economy. When productivity is high (relative to trend), the business sector demands more labor to exploit the high return to labor. This shift in labor demand puts upward pressure on real wages, which serves to draw more individuals into the labor force. The reverse holds true when productivity is low (relative to trend).

The primary goal of this chapter is to formalize the intuition above in terms of an explicit (i.e., mathematical) theory. Developing a formal model will prove useful for a number of reasons. First, it will allow us to check whether the intuition expressed above survives a logical analysis. (There are times when simple intuition only holds under some very specific conditions–or perhaps not at all). Second, we can use the logic contained in the theory to help us evaluate the potential role for government policy. Third, the simple theory developed here will serve as useful groundwork for the more elaborate theories to be developed later on in the text.

To this end, we will construct a **model economy**, populated by individuals that make economic decisions to achieve some specified goal. The decisions that people make are subject to a number of constraints so that inevitably, achieving any given goal involves a number of trade-offs. If these trade-offs fluctuate over time owing to any sort of exogenous shock, then individuals are likely to change their behavior accordingly. The question here is whether exogenous changes in productivity might generate changes in behavior that imply business cycle activity that is qualitatively similar to what is observed in reality (i.e., procylical employment and real wages).

2. A Simple Model

The model I consider here is a very simple one indeed; in particular, it makes a lot of simplifying assumptions. Many of these assumptions will appear to be highly **unrealistic**. You should resist the natural inclination to judge a model solely on the basis of its assumptions. In particular, one might note that we use unrealistic models every day for useful ends. The common roadmap as an abstract representation of the countryside is one example. We judge a roadmap not on the fact that it (unrealistically) abstracts from atmospheric conditions; we judge a roadmap on its ability to help guide us through unknown territory. The same principle should be used to evaluate economic models (abstract representations of the real economy).

In any case, you can rest assured of two things: [1] the model can be easily extended in a number of interesting (and more complicated) ways–some of which we will explore later on; and [2] the basic forces highlighted in the simple model continue to hold in much more general (and realistic) environments.

We begin by stating a number of simplifying assumptions. Since employment behavior plays an important role in the business cycle, we want to think of how to model the labor market. To this end, we want to model a household sector (from which stems the supply of labor) and a business sector (from which stems the demand for labor). So to begin, let us assume that the economy consists only of these two sectors; i.e., assume that there is no government or foreign sector. From our knowledge of the income-expenditure identity, we know that this assumption implies $G = X = M = 0$, so that $C + I = Y$. If we assume further that all output is in the form of consumer goods and services, then $I = 0$ and $C \equiv Y$. In other words, all income in this model will take the form of claims against domestically-produced consumer goods and services. In short, we are dealing here with a **closed economy**, with **no government**, **no foreign sector**, and **no investment**. These assumptions will be relaxed in later chapters.

Let us think next of the people that occupy our hypothetical world. We want to think of an economy consisting of a 'large' number of people, each of whom belong to the household sector. In reality, people obviously differ along many dimensions. On the other hand, people also seem to share many things in common, including a general desire to advance their material well-being. Our strategy here will be to focus on these common attributes and downplay the differences. For simplicity, we take this to the extreme by assuming the existence of a **representative household** (i.e., we assume that households are all identical along economically relevant dimensions).[12]

Let us now think about the business sector. We want to think of the business sector as consisting of a 'large' number of competitive firms. It is important to note that firms are not people; they are simply legal entities (operated by people) that organize production. Again for simplicity, we will assume that all firms are identical so that there exists a **representative firm**. Assume that firms are owned by members of the household sector (to which all individuals belong) and that firms are motivated by a desire to maximize shareholder wealth.

Finally, we want to consider an assumption that will simplify decision-making considerably. In particular, we consider here what is called a **static model**. The word 'static' should not be taken to mean that the model is free of any concept of time. What it means is that the decisions focussed on here have **no intertemporal dimension** (which allows us to abstract from financial markets). The restriction to static decision-making allows us, for the time-being, to focus on **intratemporal decisions** (such as the allocation of time across competing uses over the course of a year). As such, one can interpret the economy as generating a sequence of static outcomes over time.

2.2 The Household Sector

The representative household has **preferences** defined over two objects: [1] a basket of consumer goods and services (consumption), which we denote by c; and [2] a basket of home-produced goods and services (leisure), which we denote by l.[13] Let (c, l) denote a **commodity bundle** (i.e., a particular quantity of consumption and leisure); this is also sometimes called an **allocation**. The set of all conceivable commodity bundles (allocations) is called the **commodity space**.

Household preferences are defined over the commodity space. What this means in plain language is that we assume that households can **rank** different commodity bundles by making statements like: *I prefer* (c_1, l_1) *to* (c_2, l_2), or *I am indifferent between* (c_1, l_1) *and* (c_2, l_2). Under some weak conditions, we (as theorists) can **represent** such preferences with a mathematical relation called a **utility function**. In particular, let $u(c, l)$ denote the rank attached to any given commodity bundle (c, l). Then the statement *I prefer* (c_1, l_1) *to* (c_2, l_2) can be represented by $u(c_1, l_1) > u(c_2, l_2)$ and the statement *I am indifferent between* (c_1, l_1) *and* (c_2, l_2) can be represented by $u(c_1, l_1) = u(c_2, l_2)$. This is not rocket-science.

Note that by specifying the household's preferences explicitly (by way of a utility function), we are being very explicit about what **motivates** household

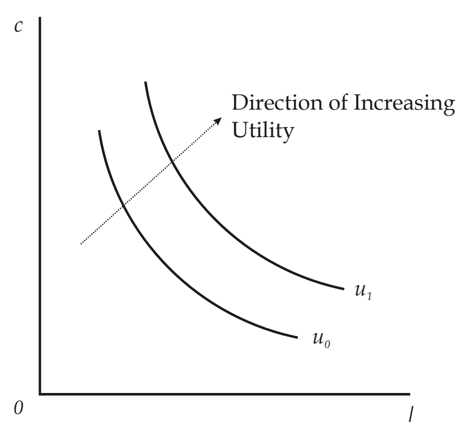
[12] Again, the student should note that one could easily extend the model to incorporate heterogeneity among households. The primary cost of doing this is some added mathematical complexity (we would have to keep track of distribution functions).

[13] Note that the value of home-produced output (leisure) is not counted as a part of the GDP.

behavior. We are suggesting that households care about the level of broad-based living standards, both in the form of consumption and leisure. Given a choice, households will presumably choose the (c, l) that they rank most highly; i.e., households are motivated by the desire to **maximize utility** $u(c, l)$. In plain language, we are just assuming that households desire to do the best they can according to their preferences. This is not an unreasonable assumption; and in particular, it is difficult to think of what one might replace it with.[14]

To gain predictive power, we need to make a few (standard) assumptions regarding preferences. First, assume that **more is preferred to less**, so that the utility function is increasing in both c and l. Second, assume that both c and l are **normal goods**. What this will imply is that as a household becomes wealthier, it will demand more of both c and l (holding the relative price of these two goods fixed). Third, assume that preferences are **transitive**. What this means is that a household prefers A to B and B to C, then it also prefers A to C. Finally, assume that the utility function is continuously differentiable and that it is strictly concave in each argument. This latter assumption implies that the household experiences a **diminishing marginal utility** of consumption and leisure as their levels are increased. Taken together, these assumptions allow us to represent preferences diagrammatically in the commodity space by way of **indifference curves**; i.e., see Figure 2.1.

FIGURE 2.1
Indifference Curves



In Figure 2.1, u_0 and u_1 simply represent two different numbers that assign a ranking to each point on their respective curves. If we fix a utility-ranking at some number u_0 (e.g., $u_0 = -3$), then the associated indifference curve is defined to be all the combinations of (c, l) that generate this rank; i.e., $u_0 = u(c, l)$.

[14] One alternative, employed in evolutionary economics, is to assume that people are born with pre-programmed behavioral rules subject to the forces of natural selection.

By construction, a household is indifferent between all the commodity bundles located on a given indifference curve. Note that the commodity bundles located on the indifference curve associated with u_1 are preferred (i.e., yield higher welfare) to those located on the lower indifference curve; i.e., $u_1 > u_0$.

Transitivity implies that the indifference curves of a given utility function can never cross. (Prove this).

A concept that we will make great use of throughout the text is the so-called **marginal rate of substitution** (or MRS, for short). Essentially, the MRS refers to the **slope of the indifference curve** at any point in the commodity space (actually, it is the *absolute value* of this slope). Notice that the MRS is a function, since it depends on where one is positioned on the indifference curve; to emphasize this, we can write $MRS(c, l)$.

The MRS has an important economic interpretation. In particular, $MRS(c, l)$ provides a measure of the **household's relative valuation of consumption and leisure** at any bundle (c, l). For example, consider some allocation (c_0, l_0) with associated utility rank $u_0 = u(c_0, l_0)$. How can we use this information to measure a household's relative valuation of consumption and leisure? Imagine taking away a small bit Δ_l of leisure from this household. Then clearly, $u(c_0, l_0 - \Delta_l) < u_0$. Now, we can ask the question: How much extra consumption Δ_c would we have to compensate this household such that they are not made any worse off? The answer to this question is given by the Δ_c that satisfies the following condition:

$$u_0 = u(c_0 + \Delta_c, l_0 - \Delta_l).$$

For a very small Δ_l, the number Δ_c/Δ_l gives us the slope of the indifference curve in the neighborhood of the allocation (c_0, l_0). It also tells us how much this household values consumption relative to leisure; i.e., if Δ_c/Δ_l is large, then leisure is valued highly (one would have to give a lot of extra consumption to compensate for a small drop in leisure). The converse holds true if Δ_c/Δ_l is a small number.

Now, sit back and relax for a moment. What we have done so far is very simple. We've assumed a world populated a large number of identical households. Each household cares about its material living standard, in the form of consumption and leisure. Generally speaking, the more the better. But households are also **willing to substitute** consumption for leisure (and vice-versa). You might be willing to sacrifice a lot of leisure in exchange for living in a big house, for example. How willing you are to do this depends on the nature of your preferences. We (as theorists) can represent this willingness by way of a diagram with appropriately shaped indifference curves (with the shape influencing the MRS). Furthermore, we can use these indifference curves to ascertain which allocations are preferred relative to others.

Before I stop talking about preferences, I want to stress once more why we (as theorists) go through all of the trouble of modeling them. There are at least two important reasons for doing so. First, one of our goals is to try to predict household behavior. In order to predict how households might react to any given change in the economic environment, one presumably needs to have some idea as to what is motivating their behavior in the first place. By specifying the objective (i.e., the utility function) of the household explicitly, we can use this information to help us predict household behavior. Note that this remains true even if we do not know the exact form of the utility function. All we really need to know (at least, for making qualitative, rather than quantitative predictions) are the general properties of the utility function (e.g., more is preferred to less, etc.). Second, to the extent that policymakers are concerned with implementing policies that improve the welfare of individuals, understanding how different policies affect household utility (a natural measure of economic welfare) is presumably important.

By modeling preferences explicitly, we take a stand on what ultimately motivates decision-making. Information concerning preferences can then be used to help predict behavior and evaluate the welfare consequences of policy and other exogenous events.

Now that we have modeled the household objective, $u(c, l)$, we must now turn to the question of what **constrains** household decision-making. Households are endowed with a fixed amount of time, which we can measure in units of either hours or individuals (assuming that each individual has one unit of time). Since the total amount of available time is fixed, we are free to normalize this number to unity. Likewise, since the size of the household is also fixed, let us normalize this number to unity as well.

Households have two **competing uses** for their time: work (n) and leisure (l), so that:

$$n + l = 1. \tag{9}$$

Equation (9) is referred to as a **time constraint**. Since the total amount of time and household size have been normalized to unity, we can interpret n as either the fraction of time that the household devotes to work or the fraction of household members that are sent to work at any given date.

In the model, households will generally have two sources of income: labor income and non-labor income.[15] Let w denote the **real wage** (i.e., the amount

[15] Note that the income that workers earn here is in the form of privately-issued claims

of output that can be purchased with one unit of labor), so that labor income is given by wn. Denote non-labor income by d (i.e., the dividends that would be accruing to the household sector via their ownership of firms in the business sector). For now, we simply view (w, d) as parameters (i.e., exogenous variables) that are beyond the control of the household. Later, I will describe how these variables are determined by market forces.

Since this is a static model, all income earned is consumed (i.e., none of it is saved). The household's **budget constraint** is therefore given by:

$$c = wn + d.$$

By combining the time constraint (9) with the equation above, we can rewrite the budget constraint in terms of consumption and leisure:

$$c + wl = w + d. \tag{10}$$

In the equation above, $w + d$ is sometimes referred to as **full income**; i.e., the combined value of the household's time endowment and non-labor income. Out of this full income, the household makes purchases of consumption and leisure, with the **price of leisure** (measured in units of consumption) given by the real wage.

I remarked earlier that the shape of the indifference curve (in particular, the MRS) reflects the household's **willingness** to substitute consumption for leisure. Observe now that the budget constraint reflects the household's **ability** to substitute consumption for leisure. The interaction between the **willingness and ability to substitute across commodities** is a concept that plays a central role in economic analysis, so it will be helpful to keep it in mind always.

Now that we have described what motivates and constrains household choices, we are in a position to deduce their behavior. Consider an arbitrary (w, d), which the household views as beyond its control. Given this (w, d), the household is assumed to choose its most preferred allocation (c, l) that at the same time respects its budget constraint. In mathematical terms, the choice problem can be stated as:

$$\boxed{\text{Choose } (c, l) \text{ to maximize } u(c, l) \text{ subject to: } c + wl = w + d.}$$

This problem has a **solution** (representing the household's optimal choice). Without saying what the solution is, we can denote it by a pair of choices (c^D, l^D), where c^D can be thought of as 'desired consumer spending' (**consumer demand**) and l^D represents the **demand for leisure**. Note that since the total time endowment is fixed, the demand for leisure automatically implies a **supply**

against output. Think of these claims as coupons that are redeemable in merchandise.

of labor, $n^S = 1 - l^D$. In terms of a diagram, the optimal choice is displayed in Figure 2.2 as allocation A (This figure is drawn for the case in which $d = 0$).

FIGURE 2.2
Household Choice

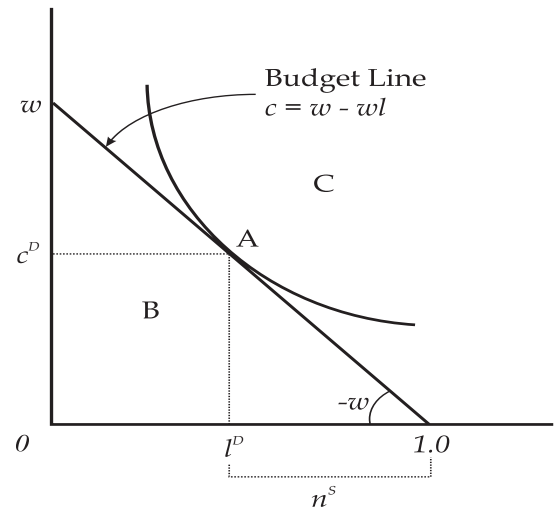


Figure 2.2 contains several pieces of information. First note that the budget line (the combinations of c and l that exhaust the available budget) is linear, with a slope equal to $-w$ and a y-intercept equal to $w + d$ (with $d = 0$ here). The y-intercept indicates the maximum amount of consumption that is budget feasible, given the prevailing real wage w. In principle, allocations such as point B are also budget feasible, but they are not optimal. That is, allocation A is preferred to B and is affordable. An allocation like C is preferred to A, but note that allocation C is not affordable. The best that the household can do, given the prevailing wage w, is to choose an allocation like A.

As it turns out, we can describe the optimal allocation mathematically. In particular, one can prove that only allocation A satisfies the following two conditions at the same time:

$$\begin{aligned} MRS(c^D, l^D) &= w; \\ c^D + wl^D &= w + d. \end{aligned} \tag{11}$$

The first condition states that, at the optimal allocation, the slope of the indifference curve must equal the slope of the budget line. The second condition states that the optimal allocation must lie on the budget line. Only the allocation at point A satisfies these two conditions simultaneously.

Exercise 2.1: Using a diagram similar to Figure 2.2, identify an allocation that satisfies $MRS = w$, but is not on the budget line. Can such an allocation be optimal? Now identify an allocation that is on the budget line, but where $MRS \neq w$. Can such an allocation be optimal? Explain.

Notice that since we have assumed that the household makes its choice conditional on some prevailing pattern of wages and dividends (w, d), it follows that the optimal choice will, in general, depend on these parameters. At times, we may wish to emphasize this dependence by writing the solution explicitly as a function of the underlying parameters; e.g., $c^D(w, d), l^D(w, d)$ and $n^S(w, d)$.

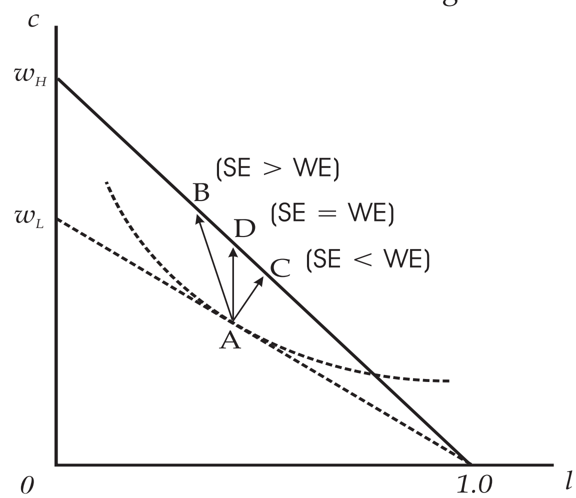
Exercise 2.2: Suppose that preferences are given by the utility function $u(c, l) = \ln(c) + \lambda \ln(l)$, where $\lambda > 0$ is a preference parameter. For these preferences, one can demonstrate that $MRS(c, l) = \lambda c/l$. Use this information, together with the conditions in (11) to solve explicitly for consumer demand, the demand for leisure and (from the time constraint) the supply of labor.

The theory developed here makes clear that the allocation of time to market work (the supply of labor) should depend on the return to work relative other potential uses of time (in this simple model, leisure is the only alternative). The return to work here is given by the real wage. Intuitively, one would expect that an exogenous increase in the real wage might lead a household to reduce its demand for leisure (and hence, increase labor supply). While this intuition is not incorrect, it needs to be qualified. We can discover how by way of a simple diagram.

Figure 2.3 depicts how a household's desired behavior may change with an increase in the return to labor. Let allocation A in Figure 2.3 depict desired behavior for a low real wage, w_L. Now, imagine that the real wage rises to $w_H > w_L$. Figure 2.3 (again, drawn for the case in which $d = 0$) shows that the household may respond in three general ways, represented by the allocations B,C, and D. In each of these cases, consumer demand is predicted to rise. However, the effect on labor supply is, in general, ambiguous. Why is this the case?

An increase in the real wage has two effects on the household budget. First, the price of leisure (relative to consumption) increases. Our intuition suggests that households will respond to this price change by substituting into the cheaper commodity (i.e., from leisure into consumption, with the implied increase in labor supply). This is called the **substitution effect**. Second, household wealth (measured in units of output) increases. Recall that both consumption and leisure are assumed to be normal goods. The logic of the model therefore implies that the demand for both consumption and leisure should rise along with wealth with the increase in leisure coming at the expense of labor. This is called the **wealth effect**. Both of these effects work in the same direction for consumption (which is why consumer demand must rise). However, these effects work in opposite directions for labor supply (or the demand for leisure). The ultimate effect on labor supply evidently depends on which effect dominates; i.e., see Figure 2.3.

FIGURE 2.3
Household Response to an
Increase in the Real Wage



Since theory alone cannot be used to ascertain the effect of wage changes on labor supply, the issue becomes an empirical one. As it turns out, the prevailing empirical evidence suggests that labor supply responds positively to an increase in the real wage (although, there is some debate as to how strong this effect is quantitatively). In what follows then, let us assume that the substitution effect dominates the wealth effect.

2.3 The Business Sector

The representative firm operates a **production technology** that utilizes labor and capital to generate output (in the form of consumer goods and services). To make things simple, assume that there is no capital, so that labor is the only factor of production.[16] The prevailing technology is represented by a linear production function:

$$y = zn;$$

where $z > 0$ is a parameter that indexes the efficiency of the production process, and y denotes the level of output. Assume that z is determined by forces that are beyond the control of any individual or firm (i.e., z is exogenous to the model). Notice that with this linear production technology, z corresponds both to the **marginal product of labor** ($\Delta y/\Delta n = z$) and the **average product of labor** ($y/n = z$). Henceforth, I will refer to z as 'productivity' and exogenous changes in z as 'productivity shocks.'

In formulating its production and hiring choices, the representative firm is assumed to take the prevailing market wage w as given (we made the same assumption for households). If a firm hires n worker-hours, then it incurs a wage bill equal to wn.[17] The employment of n worker-hours generates zn units of output, which is then 'sold' to the household sector.[18] The difference between revenue zn and cost wn constitutes profit, which is subsequently handed over to shareholders (i.e., the households) in the form of a dividend payment d. The objective of each firm is to maximize shareholder value (profit):

$$d = (z - w)n. \tag{12}$$

In mathematical terms, the choice problem facing a representative firm can be stated as follows:

Choose (n) to maximize $(z - w)n$ subject to $0 \leq n \leq 1$.

As with the household's problem, the firm's choice problem has a **solution**. Let us denote this solution by n^D (for the firm's desired labor input, or **labor demand**). The solution to this particular problem is very simple and depends only on (z, w); i.e.,

[16] This assumption is relaxed in Appendix 2.1.

[17] Remember that, since there is no 'money' in this model economy, firms must pay workers in the form of 'coupons' redeemable in the output produced by the business sector.

[18] That is, workers take their coupons home and then 'shopping' by redeeming the coupons for merchandise.

$$n^D = \begin{cases} 0 & \text{if} \quad z < w; \\ n & \text{if} \quad z = w; \\ 1 & \text{if} \quad z > w; \end{cases}$$

where n in the expression above is any number in between 0 and 1. In words, if the return to labor (z) is less than the cost of labor (w), then the firm will demand no workers. On the other hand, if the return to labor exceeds the cost of labor, then the firm will want to hire all the labor it can. If the return to labor equals the cost of labor, then the firm is indifferent with respect to its choice of employment (the demand for labor is said to be **indeterminate** in this case). With the demand for labor determined in this way, the **supply of output** (again, in the form of consumer goods and services) is simply given by $y^S = zn^D$. With this hiring and production program in place, the firm expects to generate a profit $d = y^S - wn^D$.

Notice that the demand for labor depends on both w and z, so that we can write $n^D(w, z)$. Labor demand is (weakly) decreasing in w. That is, suppose that $z > w$ so that labor demand is very high. Now imagine increasing w higher and higher. Eventually, labor demand will fall to zero. The demand for labor is also (weakly) increasing in z. To see this, suppose that initially $z < w$. Now imagine increasing z higher and higher. Eventually, labor demand will be equal 1. In short, our theory predicts that the demand for labor will be decreasing in the real wage and increasing in productivity.

2.3 Households and Firms Together

So far, we have said nothing about how the real wage (the relative price of output and leisure) is determined. In describing the choice problem of households and firms, we assumed that the real wage was beyond the control of any individual household or firm. This assumption can be justified by the fact that, in a competitive economy, individuals are small relative to the entire economy, so that individual decisions are unlikely to influence market prices.

But market prices do not fall out of thin air–ultimately, they are determined by the **collective** behavior of households and firms. In other words, we view market prices as being determined by conditions of **aggregate supply** and **aggregate demand**.[19] So, it is now time to bring households and firms together and describe how they interact in the market place. The outcome of this interaction is called a **general equilibrium**.

The economy's general equilibrium is defined as an allocation (c^*, y^*, n^*, l^*) and a price system (w^*) such that the following is true:

1. Given w^*, the allocation (c^*, n^*, l^*) maximizes utility subject to the budget

[19] Note that for an economy populated by representative agencies, computing aggregates is very simple. In particular, if there are N agents who choose x, then the aggregate is simply given by Nx. In the analysis here, we have normalized $N = 1$.

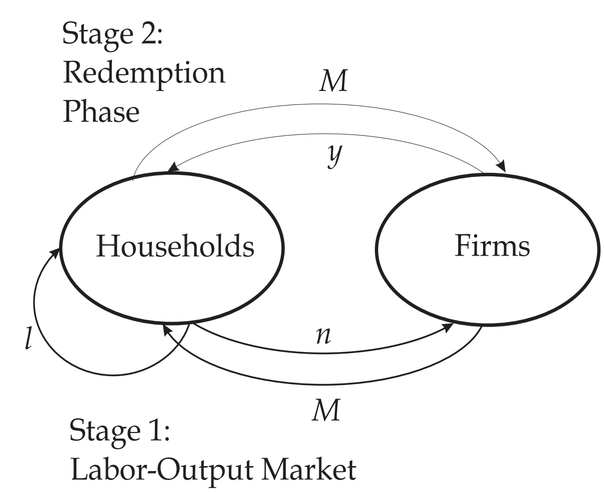
constraint [households are doing the best they can];

2. Given (w^*, z), the allocation (y^*, n^*) maximizes profit [firms are doing the best they can];
3. The price system (w^*) clears the market [$n^S(w^*) = n^D(w^*, z)$ or $c^D(w^*) = y^S(w^*, z)$].

In words, the general equilibrium concept is asking us to interpret the world as a situation in which all of its actors are trying to do the best they can (subject to their constraints) in competitive markets. Observed prices (equilibrium prices) are likewise interpreted to be those prices that are **consistent** with the optimizing actions of all agents taken together.

Before we examine the characteristics of the general equilibrium, it is useful to summarize the pattern of exchanges that are imagined to occur in each period; i.e., see Figure 2.4. One can imagine that each period is divided into two stages. In the first stage, workers supply their labor (n) to firms in exchange for coupons (M) redeemable for y units of output. The real GDI at this stage is given by y. In the second stage (after production has occurred), households take their coupons (M) and redeem them for output (y). Since M represents a claim against y, the real GDE at this stage is given by y. And since firms actually produce y, the real GDP is given by y as well.

FIGURE 2.4
Pattern of Exchange

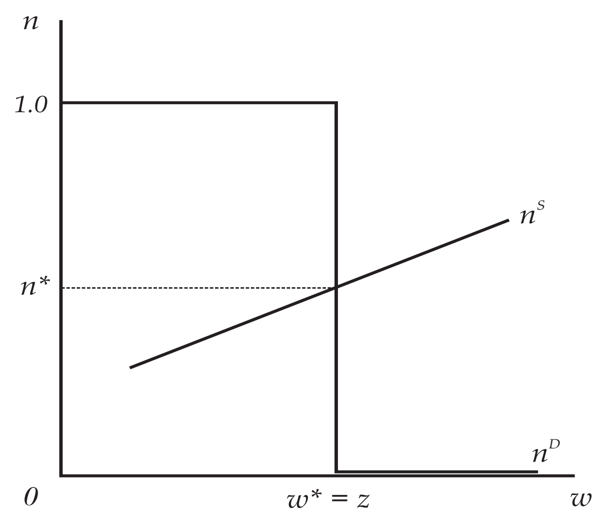


Now let us proceed to describe the general equilibrium in more detail. From the definition of equilibrium, the **equilibrium** real wage w^* must satisfy the labor market clearing condition:

$$n^S(w^*) = n^D(w^*, z).$$

Figure 2.5 provides a diagrammatic representation of the labor market.

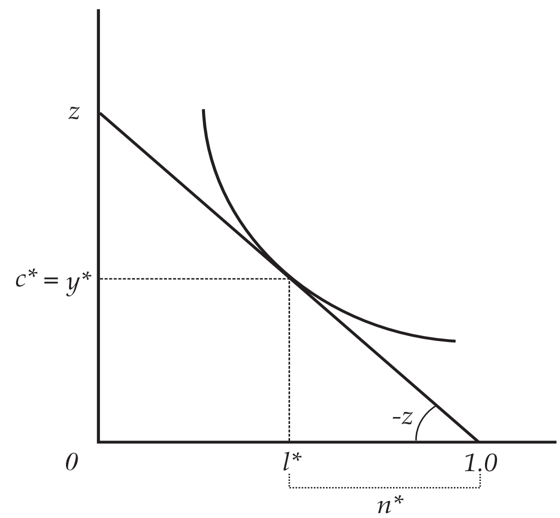
FIGURE 2.5
Equilibrium in the Labor Market



Normally, one might expect the labor demand function in Figure 2.5 to slope downward smoothly from left to right. It has this peculiar 'stepped' feature here because of the linear nature of the production technology. In Appendix 2.1, I consider a slight modification to the production technology (by introducing capital and assuming a diminishing marginal product of labor) that generates a more 'normal' looking demand function. But for present purposes, this 'stepped' function simplifies things considerably, without detracting from basic intuition.

In general, the equilibrium real wage is determined by both labor supply and demand (as in Appendix 2.1). However, in our simplified model (featuring a linear production function), we can deduce the equilibrium real wage solely from labor demand. In particular, recall that the firm's profit function is given by $d = (z - w)n$. For n^* to be strictly between 0 and 1, it must be the case that $w^* = z$ (so that $d^* = 0$). That is, the real wage must adjust to drive profits to zero so that the demand for labor is indeterminate. With w^* determined in this way, the equilibrium level of employment is then determined entirely by the labor supply function; i.e., $n^* = n^S(w^*)$. The general equilibrium allocation and equilibrium real wage is depicted in Figure 2.6.

FIGURE 2.6
General Equilibrium



Exercise 2.3: Confirm that the allocation and price system depicted in Figure 2.6 satisfy the definition of a general equilibrium.

The theory developed here makes predictions concerning the determination of output and employment. It also provides an explanation for consumer spending (in this model, it is equal to GDP) and the amount of time allocated to non-market activities (leisure). Finally, it provides an explanation for what determines the real wage. These equilibrium values $(y^*, n^*, c^*, l^*, w^*)$ constitute the model's **endogenous variables**. The theory relates the determination of these endogenous variables to the underlying structure of the economy as summarized by the set of **exogenous variables**; i.e., preferences (u) and technology (z). We are now in a position to ascertain how the model's endogenous variables are predicted to change in response to any given exogenous shock.

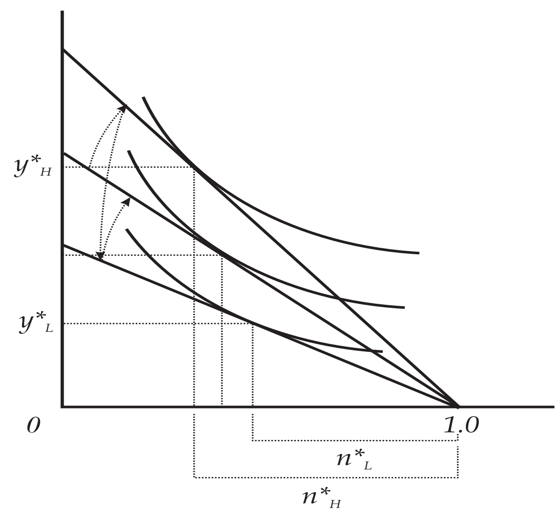
Exercise 2.4: Consider the general equilibrium allocation depicted in Figure 2.6. Is the real GDP maximized at this allocation? If not, which allocation does maximize GDP? Would such an allocation also maximize economic welfare? If not, which allocation does maximize economic welfare? What are the policy implications of this finding?

3. Understanding Business Cycles

It is an empirical fact that productivity, as measured by z, tends to grow and fluctuate. Why this happens is a very interesting question, but is not our concern here. Instead, the question we are interested in answering is the following: Given that productivity fluctuates over time, how might an economy, consisting of goal-oriented individuals and firms, be expected to respond to these fluctuations? As it happens, our theory is well-equipped to answer this question (which is not to say that the answer is entirely satisfactory).

Imagine that in our model economy, productivity fluctuates randomly between three levels $z_H > z_M > z_L$ (high, medium, and low). We can use Figures 2.5 and 2.6 to deduce that these productivity shocks will generate procyclical movements in employment and wages; i.e., see Figure 2.7.

FIGURE 2.7
Business Cycles: Productivity Shocks



Exercise 2.5: Using a diagram similar to Figure 2.5, demonstrate what effect productivity shocks have on the equilibrium in the labor market.

As I mentioned earlier, the tendency for employment and wages to move in the same direction as GDP over the cycle is a feature of the data. We can use our theory to provide us with **one interpretation** of this observed behavior. The interpretation offered by this model corresponds closely to the intuition that was expressed earlier. In particular, when productivity is high (relative to trend), the business sector demands more labor to exploit the high return to labor. This shift in labor demand (see Exercise 2.5) puts upward pressure on real wages, which serves to draw more individuals into the labor force. The reverse holds true when productivity is low (relative to trend).

The model developed above is far too simple in many respects. But at least it provides us with a starting point. And while the model obviously abstracts from many interesting dimensions that are likely important for understanding the real world, the economic forces it emphasizes are not completely crazy. In particular, the return to labor does appear to fluctuate in reality; and it is not implausible to suppose that individuals and firms alter their behavior in light of how this return changes over time.

What I would like to stress here is not the model *per se*; but rather, the **methodology** we employed. Our method involved being explicit about what motivates households and firms and how they are constrained in achieving their goals. We characterized the choices being made as being **individually rational** in the sense that supply and demand functions are the product of **optimizing behavior** (agents trying to do the best they can, according to their objectives, and subject to their constraints). We viewed the interaction of agents as occurring in **competitive markets**. This is the methodology of modern macroeconomic theory.[20]

3.1 Policy Implications

One of the benefits of our method is that it is ideally suited for addressing questions concerning government policy. Policy questions can be divided into two categories: **positive** and **normative**. A positive question is concerned with **prediction**; i.e., how **does** policy X affect variable Y? A normative question is concerned with **welfare**; i.e., how **should** government policy be designed to improve the economic well-being of people? I will address some positive questions in the next chapter; here I am interested in studying our model's normative implications.

Many people have a view that the business cycle is a 'bad' thing. This view is understandable for economic downturns, since a recession is characterized by declining incomes, wages, and employment. But this view is also expressed often during periods of economic expansion; one often hears, for example, how an economy is sometimes in danger of 'overheating.' Since we (in the developed world, at least) operate primarily in a market-based economy, there is a widespread perception that these recurring boom and bust episodes are somehow inherently linked to the market system itself. Evidently, the market is like a wild beast–useful for some purposes–but only if placed in shackles and guided by the hand of a skilled trainer. A natural candidate for the role of trainer is the government; the shackles often recommended take the form of **stabilization policies**.

[20] There is a branch of the literature that replaces the assumption of competitive markets with alternative specifications; e.g., monopolistically competitive markets, or search markets in which prices are determined by bilateral bargaining considerations.

Is there a role for stabilization policy in our model economy? From Figure 2.7, we can see clearly that our representative household is hurt by a recession (lower indifference curve). On the other hand, our model suggests that welfare increases during a cyclical boom (higher indifference curve). Is there any way to judge whether these equilibrium allocations are in some sense inefficient? A natural efficiency criterion is the concept of Pareto optimality. An allocation is said to be **Pareto optimal** if it is impossible to find a feasible allocation that improves the welfare of some person without harming the welfare of others. So, another way to approach this question is to ask whether the economy's general equilibrium (as depicted in Figure 2.7) is Pareto optimal. In other words, could a benevolent government (one that works in the interest of our representative household) do any better (according to our Pareto criterion) than what a competitive market delivers?

As it turns out, the equilibrium of our model economy is Pareto optimal. This is not a general result, but it happens to hold true here (and can continue to hold in much more complicated environments). The policy implication is rather startling: There is **no role for a government stabilization policy**. In other words, even a benevolent government would choose to vary output and employment in accordance with Figure 2.7.[21]

In fact, we can go further and state that–far from improving the welfare of households–government stabilization policies are likely to do the exact opposite. To show this formally, consider the following example. Imagine that the economy initially begins at point A in Figure 2.8. Imagine further that the economy then experiences a negative productivity shock (brought about, for example, by an extremely harsh winter).[22] Natural market forces result in a recession (a decline in output, employment, and the real wage); i.e., the new equilibrium is given by point B. The movement from allocation A to B is clearly associated with a decline in welfare. But can anything be done to improve matters?

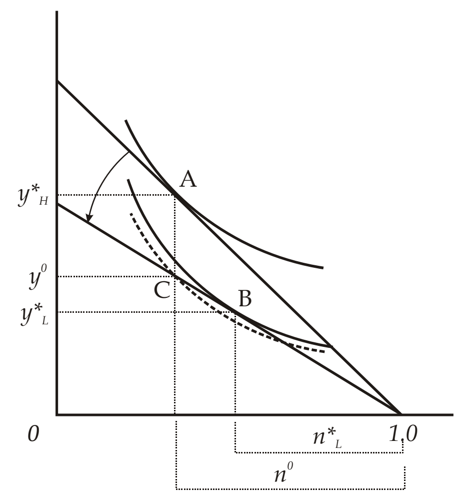
[21] The way one can prove this result is to formulate the choice problem of a benevolent government:

$$\max u(c,l) \text{ subject to: } c = y = z(1-l).$$

The solution to this problem corresponds to the equilibrium allocation.

[22] Having worked for many years in the construction sector, I can attest first-hand to the effect that winter has on labor productivity!

FIGURE 2.8
Government Stabilization Policy



One policy that the government may consider is to stabilize the level of employment at its initial level (i.e., the level associated with point A in Figure 2.8).[23] But as the government has no control over the decline in productivity, allocation A remains infeasible. The best the government can do here is the allocation given by point C. On the surface, such a policy may be deemed a success. After all, note that the level of GDP and employment is higher at point C relative to point B. But upon closer inspection, we see that (in this model, at least) the level of welfare is lower at point C.

What is going on here? The intuition is very simple. We hypothesized that the economy is subject to **exogenous** shocks to labor productivity. Individuals may not like the fact that productivity fluctuates, **but given that it does**, they respond to the implied changes in incentives (i.e., the return to work *vis-à-vis* other activities) in a natural and perfectly understandable manner. In particular, it makes perfect sense for firms to scale back their demand for labor when labor productivity falls (and expand their desired workforce when productivity rises). Likewise, it makes perfect sense for households to adjust the way they allocate time across competing activities when the relative return to different activities changes. In our model, individuals respond in precisely the correct way to changing incentives, so that collectively, **individual choices result in a socially desirable outcome**. There is no need for a government to alter the behavior of households or firms in any way.

How seriously should one take this conclusion? Arguably, it is a conclusion

[23] One way the government could do this is to offer firms a wage subsidy that is financed via a lump-sum tax on households. Another alternative is for the government to simply 'command' workers to continue working as hard as before.

that should be taken very seriously by anyone who advocates the desirability of stabilization policies. This is not to say that one needs to believe that the model or its conclusions are necessarily correct. There may very well be some role for a well-designed government stabilization policy. Nevertheless, the analysis here makes clear that it is **not obvious** why or under what circumstances stabilization policy is desirable. The simple observation that economic aggregates fluctuate and that individuals are made worse off during a recession is not sufficient evidence to justify government intervention (the model developed above makes this point very clear).

Because government intervention invariably implies some degree of coercion, one could argue that it is incumbent upon those who advocate interventionist policies to explain (to all those who treasure individual freedom) why such intervention is necessary. The value of being precise in our modeling of behavior is that it forces us to **think more deeply** about the ultimate rationale for government policy. This deeper understanding will presumably translate into better policy advice concerning the exact nature of a well-designed government intervention.

5. Summary

Labor is the most important factor of production. In market economies, labor is exchanged in markets where households supply labor in exchange for claims against output, and where the business sector demands labor in exchange for claims against their output. The price of labor (the real wage) depends on the supply and demand for labor.

Market economies are characterized by fluctuations in the labor input and corresponding fluctuations in the production of output. Most economists agree that these fluctuations stem largely from movements in the demand for labor; but there is considerably less agreement about what actually causes these movements.

One interpretation–the one studied in this chapter–is that exogenous changes in factor productivity might be responsible for causing these movements in labor demand. When labor productivity is high, firms demand more labor at any given wage; and the upward pressure on wages induces households to substitute out of non-market activities into market activities. Output (GDP) expands owing to the increase in both productivity and the increased flow of resources reallocated toward market production. The resulting changes in output and employment are the byproduct of households and firms reacting in a natural way to changing economic circumstances; there is no a prior reason to believe that these fluctuations are in any sense brought about by "irrational" behavior.

According to this view of the business cycle, there is no role for government stabilization policies. It makes no sense to stabilize output and employment when the government is in no position to influence productivity.

Appendix 2.1
A Model with Capital and Labor

The model of time allocation developed in this chapter assumed that labor is the only factor of production. If there are constant returns to scale in production, then this motivates the existence of a production function that takes the linear form $y = zn$.

The model is easily extended to include two (or more) factors of production, which we can think of as capital (k) and labor (n). In this case, the production function can be written in a general form as $y = zF(k, n)$. A specific functional form for F is the *Cobb-Douglas* specification: $F(k, n) = k^{1-\theta} n^{\theta}$, where $0 \leq \theta \leq 1$ is a parameter that indexes the relative importance of labor in production. Our earlier model is just the special case in which $\theta = 1$.

As this is a static model, assume that the amount of capital is fixed in supply; for example, suppose that $k = 1$. The production function can then be written as $y = zn^{\theta}$. What this tells us is that output is an increasing function of labor (as before). However, as employment expands, output does not expand in a linear manner (as it did before). In particular, output expands at a declining rate. The reason for this is because as more labor works with a given amount of capital, the average (and marginal) product of labor declines.

For this production function, the marginal product of labor is given by:

$$MPL(n, z) = \theta z n^{\theta - 1}.$$

The **marginal product of labor** tells us the extra output that can be produced with one additional unit of labor. On a diagram, the MPL can be depicted as the **slope of the production function**. Observe that when $\theta = 1$, we have $MPL(n, z) = z$ (i.e., the MPL is a constant). When $\theta < 1$, then the MPL declines (the slope of the production function becomes flatter) as n increases. As well, note that for a given level of n, an increase in z implies an increase in the MPL.

The choice problem facing a typical firm is given by:

$$\boxed{\text{Choose } (n) \text{ to maximize } d = zn^{\theta} - wn.}$$

The solution to this choice problem is a desired labor input (labor demand) function n^D, which happens to satisfy the following condition:

$$MPL(n^D, z) = w.$$

As an exercise, you should try to solve for the labor demand function n^D and show that it is smoothly decreasing in w (unlike the 'step' function in the chapter). Once we know n^D, we can easily calculate the supply of output $y^S = z(n^D)^\theta$ and the planned dividend payment $d = y^S - wn^D$. Note that unlike before, firms here will generally earn a positive profit $d > 0$, which reflects the return to the capital used in production.

The household's choice problem remains as before, except that now households have two sources of income (wage and dividend income). Given some arbitrary market wage w and dividend income d, the solution satisfies:

$$\begin{aligned} MRS(c^D, l^D) &= w; \\ c^D &= w(1 - l^D) + d. \end{aligned}$$

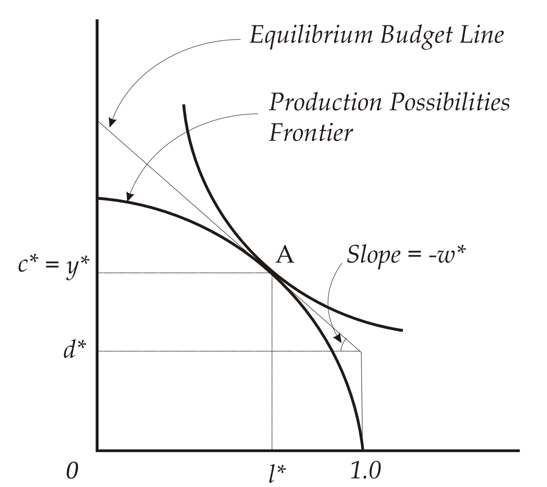
Once we know l^D, we can infer the labor supply function from the time-constraint: $n^S = 1 - l^D$.

Since the market wage w has been arbitrarily chosen at this stage, it is generally not a market-clearing wage. The next step then is to impose the **market-clearing condition**:

$$n^D = n^S.$$

The assumption here is that the real wage adjusts in order to ensure that the market clears. The model's general equilibrium is depicted as Point A in Figure 2.10.

FIGURE 2.10
General Equilibrium



The general equilibrium of this model economy differs from the one in the text in only two minor ways. First, the equilibrium real wage is equal to $w^* =$

$z(n^*)^{\theta-1}$ (it now varies with the equilibrium level of employment). When $\theta = 1$, we once again have $w^* = z$. Second, firms actually earn 'profit' (generate dividends) in this model. This profit represents the return to the capital that is used in production. Since households own the equity in the business sector, they are also the ultimate owners of capital. The dividend payment reflects this ownership in the capital stock.

The key thing to note here is that the main conclusions derived in the body of the chapter are in no way affected by this more realistic production structure.

Exercise 2.6: Consider a model economy populated by a representative household with preferences $u(c,l) = c + \lambda \ln(l)$; so that $MRS(c,l) = \lambda/l$. There is also a representative firm with technology $y = zn^{1/2}$; so that $MPL(n,z) = (1/2)zn^{-1/2}$. Solve for the competitive equilibrium allocation and real wage rate as a function of z and other parameters. How does this economy respond to exogenous changes in z?

Let me show you how to answer the question above. First, begin with the household's choice problem. For a given (w,d), the household maximizes utility subject to its budget and time constraints. We know that the optimal choice is characterized in the following way:

$$\begin{aligned} \lambda/l^D &= w; \\ c^D &= w(1-l^D) + d. \end{aligned}$$

From the first equation, we see that $l^D(w) = \lambda/w$. That is, a greater taste for leisure (an increase in λ) means that more leisure is demanded. But more importantly, note that an increase in w means that less leisure is demanded. In terms of labor supply, we have $n^S(w) = 1 - \lambda/w$. That is, a higher w implies that more labor is supplied (hence, for these preferences, the substitution effect dominates the wealth effect). The consumer demand function can be solved for by plugging in $l^D(w)$ into the second equation above; i.e., $c^D(w,d) = w - \lambda + d$. That is, a higher w or d implies greater consumer demand.

The second step is to solve for the firm's choice problem. For a given (w,z), the firm maximizes wealth (d) subject to its technology constraint and the cost of labor. We know that the optimal choice here is characterized by:

$$\frac{1}{2}z(n^D)^{-\frac{1}{2}} = w.$$

Rearranging terms (multiplying both sides by $n^{1/2}$), we have:

$$(n^D)^{\frac{1}{2}} = \frac{1}{2}zw^{-1}.$$

Now, square both sides to derive the labor demand function:

$$n^D(w,z) = \frac{1}{4}\left(\frac{z}{w}\right)^2.$$

Note that labor demand is decreasing in w and increasing in z. The supply of output can then be derived by plugging the labor demand function into the production function; i.e.,

$$y^S(w,z) = z\left[\frac{1}{4}\left(\frac{z}{w}\right)^2\right]^{1/2} = \left(\frac{z^2}{2w}\right).$$

Hence, for a given (w,z), the firm plans to distribute profit equal to:

$$d(w,z) = y^S(w,z) - wn^D(w,z).$$

The final step is to construct the equilibrium. In particular, we have to find a wage w^* such the following is true: $n^S(w^*) = n^D(w^*,z)$, $y^S(w^*,z) = c^D(w^*,d^*)$, and $d^* = d(w^*,z)$. As it turns out, we can focus on $n^S = n^D$, since if this holds, then we can verify that the other conditions will hold as well. Using what we have derived above, the labor-market clearing condition is given by:

$$1 - \lambda w^{-1} = \frac{1}{4}z^2 w^{-2}.$$

Multiply both sides of this equation by w^2 to derive:

$$w^2 - \lambda w - \frac{1}{4}z^2 = 0.$$

From high-school algebra, you should recognize this as a quadratic equation in w. What this means is that there will be two solutions for w^*; but in general, only one of these solutions will make economic sense. Using the quadratic formula, we can solve for these solutions as follows:

$$w = \frac{\lambda \pm \left(\lambda^2 + z^2\right)^{1/2}}{2}.$$

Since negative wage rates do not make sense, we can safely assume that the solution is:

$$w^*(z) = \frac{\lambda + \left(\lambda^2 + z^2\right)^{1/2}}{2}.$$

This implies that w^* is an increasing function of z (which is consistent with what we derived in the body of the chapter).

Once we solve for $w^*(z)$, we are basically done. The equilibrium level of employment can be derived by plugging w^* into either the labor supply or labor demand function (it does not matter, since they are both equal to each other in equilibrium). Doing so results in an expression for $n^*(z)$. You should be able to verify that n^* is an increasing function of z (and explain why this is the case). By plugging n^* into the production function, you can then derive an expression for $y^*(z) = c^*(z)$. Again, you should be able to verify that the equilibrium level of production (and consumption) is an increasing function of z. Finally, if you want, you can also derive an expression for $d^*(z)$ and show that the equilibrium dividend payment is also an increasing function of z. All of this (except for the last result) is consistent with what we studied in the body of the chapter.

Problems

1. What is the definition of a theory? Explain the difference between an exogenous and endogenous variable.

2. Is the fact that a theory makes an unrealistic assumption sufficient to reject it as a plausible interpretation of the data?

3. Give two reasons why it is important for an economic model to have the preferences of individuals stated explicitly.

4. Provide a justification for the hypothesis that individuals maximize utility (subject to their constraints). Is it possible to test such an hypothesis?

5. Judging by the theory developed in this chapter, an exogenous increase in productivity appears to increase both the supply of and demand for labor. Clarify what is meant by this (in particular, be careful to distinguish between movements along a curve and shifts in a curve). A diagram would be helpful.

6. You are asked to write a brief newspaper column explaining the nature of the business cycle. Obviously, you cannot use mathematical notation or make any use of diagrams. Furthermore, your audience will generally not be familiar with economic jargon. Write your article based on the economic interpretation provided by the theory developed in this chapter.

7. Is it obvious that evidence of a business cycle justifies a call for government stabilization policies?

8. Provide a critique of the view that productivity is determined by exogenous factors.

CHAPTER 3

Uncertainty and Expectations

1. Introduction

It seems obvious that many–if not most–decisions must be made in the context of some uncertainty over how the future will unfold. You attended class today because you **expected** some return from doing so. At some point in the future, you will realize an **actual** return that either meets, exceeds, or falls short of your expectations. Likewise, in the business sector, firms must decide on how many workers to hire, or much to invest, prior to knowing the exact outcome of such decisions. Similar considerations are at work in other sectors of the economy.

The question of how expectations are formed and what role they play in determining business cycles presents a number of challenges for theorists and empiricists alike. One important issue involves the question of how welfare is to be measured. For example, should welfare be measured in an **ex ante** (expected) sense, or an **ex post** (actual) sense? Most economists are inclined to adopt an ex ante criterion. For example, when a person buys insurance, this presumably increases expected (ex ante) utility. If the insured event fails to occur, actual (ex post) utility will be lower than if the person had not purchased the insurance; that is, there may be some regret that money was wasted on purchasing the insurance policy. It seems clear enough, however, that this ex post measure is not the correct way to measure the cost and benefit of insurance. In this section, I touch on some of the issues involved surrounding the question of expectations.

2. Decision-Making Under Uncertainty

Let us begin by modifying the simple model developed in Chapter 2 in a way that takes uncertainty (over the return to labor) more seriously. To this end, imagine that z can take one of two values: $z = z_H$ or $z = z_L$ with $z_H > z_L$ (i.e., productivity is either 'high' or 'low'). Let $p(z, s)$ denote the probability of z occurring, conditional on receiving some **information** s (a signal that is correlated with productivity). This signal takes one of two values: $s = g$ or $s = b$ (i.e., the signal is either 'good' or 'bad'). Assume that a good signal implies that z_H is more likely than z_L; and assume that the opposite is true with a bad signal. Mathematically, we can state this assumption as:

$$\begin{aligned} p(z_H, g) &> p(z_L, g); \\ p(z_H, b) &< p(z_L, b). \end{aligned} \tag{13}$$

Let me now be explicit about the **structure of information** and the **timing** of events. Assume that a period begins with the arrival of some information s. Next, *prior to the realization of actual productivity*, households and firms make their employment choices. Finally, the actual level of productivity z is determined. At this final stage, production and consumption takes place given the actual z, and given the level of employment as determined in the previous stage.

The probability structure $p(z, s)$ is assumed to be known by all agents in the economy. Another way of stating this is that decision-makers understand the **fundamentals** governing the random productivity parameter. These fundamentals are viewed as being determined by God or nature (i.e., they are exogenous). In other words, for better or worse, this is just the way things are: the world is an uncertain place and people must somehow cope with this fact of life.

So how might individuals cope with this uncertainty? It seems reasonable to suppose that they form **expectations** over z, given whatever information they have at their disposal. Since individuals (in our model) are aware of the underlying fundamentals of the economy, they can form **rational** expectations. Let $z^e(s)$ denote the expected value for z conditional on having the information s. Then it is easy to calculate:

$$\begin{aligned} z^e(g) &= p(z_H, g)z_H + p(z_L, g)z_L; \\ z^e(b) &= p(z_H, b)z_H + p(z_L, b)z_L. \end{aligned} \tag{14}$$

Given the probability structure described in (13), it is clear that $z^e(g) > z^e(b)$. In other words, if people observe the 'good' signal, they are 'optimistic' that productivity is **likely** to be high (this is not guaranteed, of course). Conversely, if people receive the 'bad' signal, they are 'pessimistic' and believe that productivity is **likely** to be low. As information varies over time (i.e., as good and bad signals are observed), it will appear as if the 'mood' or 'confidence' of individuals varies over time as well. **These apparent mood swings, however, have nothing to with psychology**; i.e., they reflect entirely rational changes in expectations that vary as the result of changing fundamentals (information).

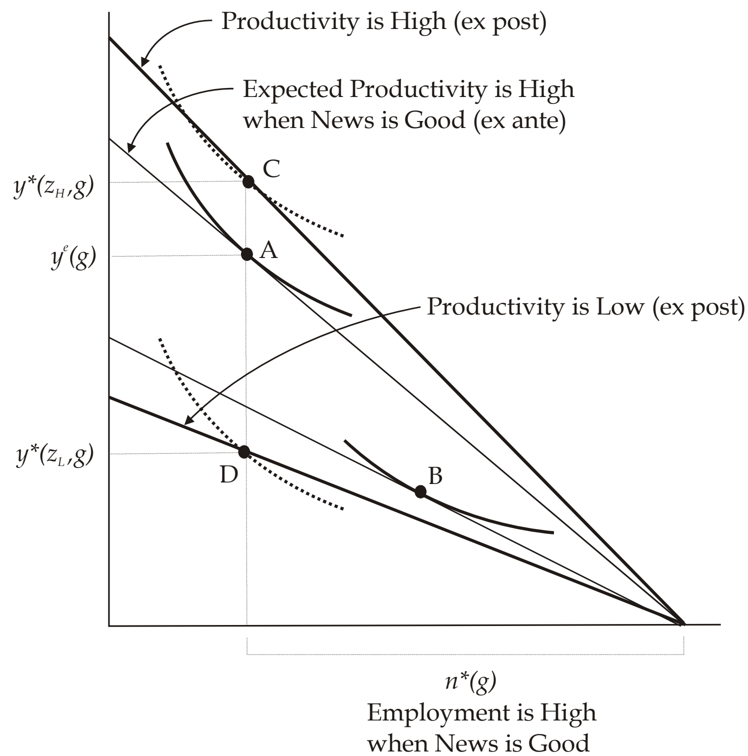
In the model, firms always earn (in equilibrium) zero profits–both in expected and actual terms (before and after uncertainty is resolved). What this implies is that the expected wage must be given by $w^e(s) = z^e(s)$, with the actual wage given by $w^*(z) = z$. Of course, given the timing of events, employment decisions must be based on the **expected** wage. Since actual employment in our model is determined entirely by labor supply, household decision-making determines the level of employment and the (expected) level of consumer demand; i.e.,[24]

$$\begin{aligned} MRS(c^e(s), 1 - n^*(s)) &= w^e(s); \\ c^e(s) &= w^e(s)n^*(s). \end{aligned} \tag{15}$$

[24]Instructors: For this mathematical characterization to hold exactly, we must assume that **certainty equivalence** holds (e.g., assume that preferences are quadratic).

In this model, observing a good signal will result in an employment boom (since the expected return to labor is high). Conversely, observing a bad signal will result in an employment bust (since the expected return to labor is low). The economics here are exactly the same as what has been described in Chapter 2. The only difference here is that **actual** GDP (consumption) will in general differ from **expected** GDP (consumption). In particular, actual GDP is given by $y^*(z,s) = zn^*(s) = c^*(z,s)$.

FIGURE 3.1
Ex Ante and Ex Post Outcomes



To illustrate how things might work here, consider Figure 3.1. Imagine that agents begin the period by observing $s = g$ (good news). Conditional on this good news, expected productivity is high; and the equilibrium level of employment $n^*(g)$ is determined by point A. Note that while productivity is expected to be high, it may subsequently turn out to be either high (z_H) or low (z_L). If it turns out to be high, then the ex post allocation is given by point C; if it turns out to be low, then the ex post allocation turns out to be point D. Either way, the actual (ex post) level of GDP is not going to be equal to the expect (ex ante) GDP. The discrepancy between actual and expected GDP here is not a product of people making 'mistakes;' it is a product of having to make decisions in an uncertain world.

Exercise 3.1: In Figure 3.1, we assumed that $s = g$. If instead it turns out that $s = b$ (bad news), then the equilibrium level of employment is determined by point B. Using Figure 3.1, depict the ex ante and ex post levels of GDP.

Exercise 3.2: Note that in this model economy, it is possible that Nature produces a sequence of actual productivity realizations that are identical over extended periods of time. At the same time, it is possible that the signals (news) people receive fluctuate randomly. Over such a sample period, the economy will display cyclical fluctuations in GDP and employment, even though actual productivity remains constant. Are the fluctuations produced in this manner 'inefficient' in any sense? Explain.

In the model developed here, expectations play an important role in determining resource allocation. When people are optimistic, the economy is likely to boom; when people are pessimistic, the economy is likely to bust. However, note that expectations themselves are not the cause of these boom-bust cycles. As with the simple model developed earlier, **the business cycle is the product optimal decisions in reaction to changing fundamentals**. In other words, if individuals are optimistic, it is because they have good reason to be so (i.e., they have received information that leads them to revise upward their estimate of productivity). Likewise, a sudden wave of pessimism is not the product of some unexplained psychological depression; instead it is the product of a rational expectation in response to 'bad news.' **Expectations in this model simply reflect the changing nature of economic fundamentals (information)**.

3. Irrational Expectations

The nature of expectations expressed in the previous section does not appear to reflect the view held by numerous commentators and policymakers (including central bankers). The prevailing view appears to be one in which **psychology** plays a prominent role in the formation of expectations. According to this view, exogenous (i.e., unexplained) changes in expectations themselves constitute an important source of shocks for an economy. These expectation shocks are sometimes called **animal spirits**–a colorful phrase coined by the famous economist John Maynard Keynes.[25]

There are a couple of ways in which to formalize this notion of animal spirits. One way is to simply assume that expectations are **irrational** (i.e., the product

[25]Keynes, John M. (1936, pp. 161-62). *The General Theory of Employment, Interest and Money*, MacMillan, Cambridge University Press.

of random changes in psychological sentiment that bear no obvious relation to an economy's underlying fundamentals). One crude way to model this is to assume that while actual productivity (z) remains constant over time, expected productivity (z^e) fluctuates for no apparent reason. Imagine, for example, that corporate executives wake up one morning feeling 'optimistic' (i.e., $z^e > z$). Then the **expected** return to production is high, which stimulates the demand for labor. The result is an expansion in employment. Unfortunately, since actual productivity is in fact given by z, these high expectations invariably turn out to be unfounded. Nevertheless, the end result is a level of employment and GDP that is higher than 'normal.' The opposite holds true if corporate executives instead wake up in the morning feeling 'pessimistic' (i.e., $z^e < z$).

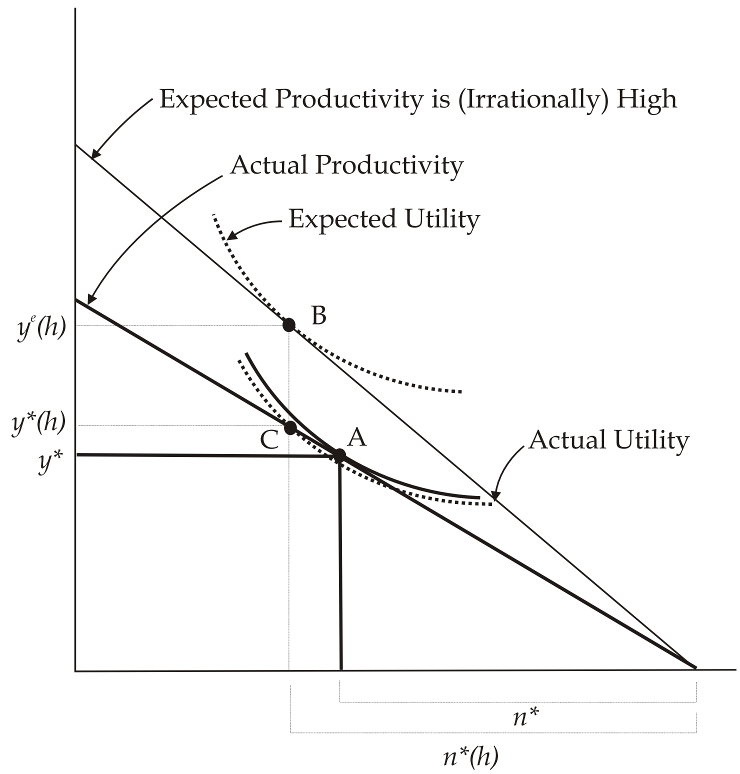
Let us consider an example. Assume that the true level of productivity is fixed at some level z. If people had "rational" expectations, then they would set $z^e = z$. The resulting equilibrium allocation is depicted by point A in Figure 3.2. Under rational expectations, the equilibrium level of GDP would always remain at y^* and the equilibrium level of employment would always remain at n^*. Using commonly employed language, we can think of y^* as denoting "potential GDP" and n^* as denoting "full employment."

Now, assume instead that people (for some unexplained psychological reason) do not believe that z is a constant (despite the fact that actual productivity remains unchanged forever). For simplicity, assume that agents are either optimistic or pessimistic; that is, $z^e = h$ or $z^e = l$ (where $h > z > l$). I assume the same timing structure as before; that is, people first make their employment decisions based on expected productivity, followed by production and consumption choices that are based on actual productivity.

Figure 3.2 depicts the case in which $z^e = h$ (irrational exuberance). Given peoples' expectations, the ex ante equilibrium allocation is depicted by point B. That is, the equilibrium level of employment is above full employment, and the expected level of GDP is above potential. The economy appears to be "overheating" in this case. Of course, these high expectations will inevitably be dashed by the ensuing reality (actual productivity is lower than what is initially expected). The ex post allocation is depicted by point C.

There are several interesting things to note about the way this model economy functions. First note that optimistic expectations lead to an employment boom. Moreover, while actual GDP turns out to be lower than expected, it is still higher that it would have been under rational expectations (point A). One might be tempted to conclude that since actual output and employment turns out to be higher than it would otherwise have been (if people had rational expectations), then people in this economy must be made better off when they are optimistic. The theory, however, suggests otherwise. In particular, note that the indifference curve passing through point C is **lower** than the one passing through point A.

FIGURE 3.2
Animal Spirits: Irrational Expectations



Exercise 3.3: Redraw Figure 3.2 for the case in which $z^e = l$ (irrational pessimism) and demonstrate that: [1] employment is below full employment; [2] expected GDP is lower than potential; [3] actual GDP turns out to be higher than expected, but still below potential; and [4] welfare is lower relative to what is achieved under rational expectations.

The "mood swings" described above are irrational because expectations do not conform with the economy's underlying fundamentals. As a rule, (academic) economists do not take this view very seriously, since it is hard to imagine that corporate executives (and other economic actors) would not eventually learn from experience. Nevertheless, such a view appears common in many circles. For example, we are often warned in the financial pages of newspapers that the economy is in danger of "overheating." Even the fabled (now former) Federal Reserve Chairman Alan Greenspan has, in the past, warned us of "irrational exuberance" in the market place.

What are we to make of this? While the idea of irrational expectations might seem crazy to an economist, the concept does appear to offer us an insight as to what motivates many policy recommendations. In particular, if private sector decisions are influenced by a degree of irrational mood swings, then one can plainly see the benefit of a well-designed stabilization policy. In the context of the model described above, the government should act to stabilize output and employment at their fundamental levels. Of course, this policy prescription presumes that government decision-makers themselves are not prone to forming irrational expectations. Exactly how it is that the government has the presence of mind to view things clearly, while private decision-makers do not, is never fully explained. Perhaps the answer is too embarrassing to contemplate.

4. Self-fulfilling Prophesies

The previous section describes one way to think about the concept of animal spirits. As I have alluded above, this treatment is viewed skeptically by (academic) economists, as it relies on the notion of irrational expectations. There is, however, a more sophisticated way in which to formalize the concept of animal spirits. This more sophisticated treatment does not abandon the assumption of rational expectations. According to this view, expectation shocks can become **self-fulfilling prophesies**.

A self-fulfilling prophesy is a situation where an exogenous change in expectations can alter economic reality in a way that is consistent with the change in expectations. Some of us may have had experiences that fit this description. For example, you wake up on the morning before an exam and (for some unexplained reason) expect to fail the exam regardless of how hard you study. With this expectation in place, it makes no sense to waste time studying (you may as well go to the bar and at least enjoy the company of friends). Of course, the next day you write the exam and fail, confirming your initial expectation (and rationalizing your choice of visiting the bar instead of studying). But suppose instead that you woke up that fateful morning and (for some unexplained reason) expect to pass the exam with a last-minute cram session. The next day you write the exam and pass, confirming your initial expectation (and rationalizing your decision to study rather than drinking).

According to this story, what actually ends up happening (pass or fail) depends critically on what sort of 'mood' you wake up with in the morning before the exam. Your mood is uncontrollable. But given your mood, you can form expectations rationally and act accordingly. Your actions and outcomes can be consistent with your initial expectations.

This example is probably not the best one since it relates one's expectation only to oneself. But the same idea can apply to how one's expectations are formed in relation to the behavior of others. A useful tool for examining the strategic interaction of expectations is the theory of games (**game theory**).

Consider the following two-person **coordination game**. For concreteness, imagine that the problem involves deciding on which side of the road to drive on (we are considering a world where there are no laws dictating which side to drive on). Each person can make one of two choices: R or L. The two people have no way of directly communicating with each other beforehand, but we allow for the possibility that their choices might be conditioned on a publicly observable variable, commonly referred to as a **sunspot variable**. Assume that the sunspot variable blinks on or off with equal probability. By assumption, a sunspot variable has no effect on the economy's fundamentals; that is, it simply serves as a coordination device.

One equilibrium of this game is as follows. When the sunspot blinks on, everyone chooses R; and when the sunspot blinks off, everyone chooses L. This behavior is an equilibrium because every person behaves rationally, given his or her expectation of the behavior of others. And indeed, the outcome is not a bad one from a social perspective, as head-on collisions are avoided in this equilibrium. The equilibrium is also rather odd in some respect. In particular, note that the economy 'fluctuates' between R and L; even though there is **no fundamental reason** for why this should be the case (the fluctuations are triggered by an exogenous sunspot variable). The outcome I have described here has the flavor of a self-fulfilling prophesy. The 'expectation shock' in this case is the realization of the sunspot variable. If everyone wakes up in the morning expecting that others will choose R, then everyone will choose R, so that expectations turn out to be consistent with behavior.

Let me now consider a slightly more general version of the coordination game described above. Instead of R and L, imagine instead that the two choices are P or O (for 'pessimistic' and 'optimistic'). Assume that there is a big payoff in this economy if everyone coordinates on O; and that there is a small payoff if everyone coordinates on P. Moreover, assume that any lack of coordination results in disaster. Formally, I can represent this game in the following way:

Player 2	P	O
Player 1		
P	**5**	**0**
	5	**0**
O	**0**	**10**
	0	**10**

The boldfont numbers in the table above represent the individual payoffs associated with the choices that society makes. Observe that if both players choose O, there is a 'big' joint payoff (each person gets **10**). If both players choose P, there is a 'small' joint payoff (each person gets **5**). Any lack of coordination (one player chooses P and the other O) results in a very small payoff (each person gets **0**).

As with the earlier coordination game, this game features two (pure strategy Nash) equilibria. One equilibrium has both players coordinating on P; and the other has both players coordinating on O. (If you have not taken any game theory, ask your instructor to explain this.) But this game differs from our 'roadside' game in that society clearly prefers one equilibrium over the other. As with our earlier example, this economy may exhibit expectation (sunspot) driven fluctuations between O and P. It is also possible for society to become 'stuck' in the bad equilibrium for a long period of time.

Let me now try to cast this idea within the context of our simple macroeconomic model. Imagine that the existing technology is such that, from the perspective of an individual business, the return to hiring labor is high when all other firms are producing at high levels (employing many workers). Conversely, the return to labor is low when other firms are producing at low levels (employing few workers). We can model such a technology as follows:

$$y = \begin{cases} z_H n & \text{if} \quad \mathbf{n} \geq \mathbf{n}_C; \\ z_L n & \text{if} \quad \mathbf{n} < \mathbf{n}_C; \end{cases}$$

where $0 < \mathbf{n}_C < 1$ denotes some 'critical' aggregate level of employment and $\mathbf{n}$ denotes the level of **aggregate** employment. Note that it is very important here to distinguish between the level of employment at an individual firm n and the aggregate level of employment $\mathbf{n}$ (interpreted here as the average number of workers employed at all other firms in the economy). The distinction is important because an individual firm can control n but cannot control $\mathbf{n}$. Nevertheless, $\mathbf{n}$ has an important influence on the productivity of each individual firm. That is, if $\mathbf{n} \geq \mathbf{n}_C$, then an individual firm produces according to $y = z_H n$; and if $\mathbf{n} < \mathbf{n}_C$, then an individual firm produces according to $y = z_L n$ (where $z_H > z_L$).

Now, imagine that an individual firm must make a decision as to how many workers to hire based on an **expectation** of whether other firms are likely to hire many or few workers; that is, on whether $\mathbf{n} \geq \mathbf{n}_C$ or $\mathbf{n} < \mathbf{n}_C$. In this case, it is clear that for each individual firm, the expected productivity of its enterprise will depend on its own forecast of how all other firms in the economy are likely to behave. If an individual firm is 'optimistic' (i.e., it expects $\mathbf{n} \geq \mathbf{n}_C$), then it will (rationally) forecast a high return to hiring labor; that is, it will make sense to hire a lot of labor. On the other hand, if a firm is 'pessimistic' (i.e., it expects $\mathbf{n} < \mathbf{n}_C$), then it will (rationally) forecast a low return to hiring labor; that is, it will make sense to scale back on hiring.

Let us now turn to characterizing the behavior of households. As with individual firms, individual households too must try to forecast $\mathbf{n}$, as $\mathbf{n}$ will ultimately determine whether the equilibrium wage turns out to be high (z_H) or low (z_L). If a household is optimistic, then its individual labor supply n will satisfy the usual condition:

$$MRS(z_H n, 1 - n) = z_H;$$

where here, I have used the fact that $c = z_H n$ (anticipating that $d = 0$) and $l = 1 - n$. On the other hand, if a household is pessimistic, then its individual labor supply n will satisfy:

$$MRS(z_L n, 1-n) = z_L.$$

Thus, optimal behavior for each individual firm and household (n) depends on what each individual agent is expecting in terms of $\mathbf{n}$.

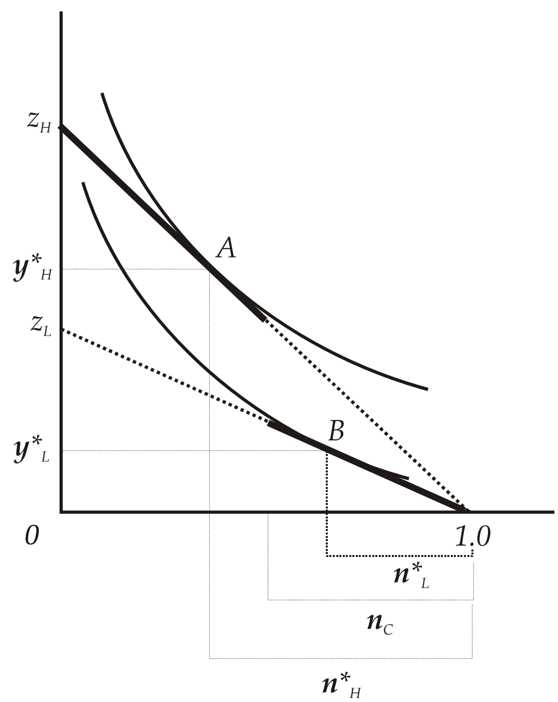
What is the likely outcome of this type of economy? To answer this question, we impose the equilibrium condition: $n = \mathbf{n}$. That is, since all agents are alike, let us assume that they all ultimately **coordinate** on the same individual choices. In this case, all individual labor supply and demand choices n correspond to the average level of employment $\mathbf{n}$.

As with the simple coordination games described above, it appears that multiple equilibria are possible. In particular, we can identify two numbers $(\mathbf{n}_L^*, \mathbf{n}_H^*)$ such that $\mathbf{n}_L^* < \mathbf{n}_C < \mathbf{n}_H^*$, each of which satisfy:

$$\begin{aligned} MRS(z_H \mathbf{n}_H^*, 1-\mathbf{n}_H^*) &= z_H; \\ MRS(z_L \mathbf{n}_L^*, 1-\mathbf{n}_L^*) &= z_L. \end{aligned}$$

These two possible outcomes are displayed in Figure 3.3 as points A and B.

FIGURE 3.3
Animal Spirits: Rational Expectations



In Figure 3.3, the solid (disconnected) line depicts the economy's production technology. That is, productivity is low when aggregate employment is low; and vice-versa when aggregate employment is high. Hence, this economy features a form of **increasing returns to scale** (productivity is weakly increasing in the level of employment). This is to be contrasted with our earlier models, each of which featured constant returns to scale in labor (except for the extension in Appendix 2.1, which featured diminishing returns to scale in labor).

The key thing to note here is that **both** points A and B are consistent with a rational expectations equilibrium. To see this, assume that all agents expect (for some unexplained reason) an aggregate level of employment equal to $\mathbf{n}_H^*$. Then, conditional on this expectation, all agents will rationally choose point A. At this point, the labor-goods market clears (supply is equal to demand). Moreover, the initial expectation turns out to be **consistent** with what actually transpires (expectations are rational in this sense).

On the other hand, assume that all agents expect (for some unexplained reason) an aggregate level of employment equal to $\mathbf{n}_L^*$. Then, conditional on this expectation, all agents will rationally choose point B. At this point, the labor-goods market clears (supply is equal to demand). Moreover, the initial expectation turns out to be **consistent** with what actually transpires (again, expectations are rational in this sense).

In other words, what we have here is an economy where one of two outcomes are possible; with the actual outcome determined entirely on a **non-fundamental** (arbitrary) initial expectation. Both outcomes are the product of a **self-fulfilling prophesy**. If the psychology of the market is optimistic, then the 'high-level' equilibrium will occur; if the psychology of the market is pessimistic, then the 'low-level' equilibrium will occur. Note that households are clearly better off at the high-level equilibrium (point A). Unfortunately, there is nothing here to prevent the low-level equilibrium from occurring. Over time, the economy may fluctuate between points A and B for purely psychological reasons. It is also possible that the economy gets 'stuck' at the low-level equilibrium for prolonged periods of time. This latter possibility appears consistent with Keynes' own view of the cause of the Great Depression:

> "In particular, it is an outstanding characteristic of the economic system we live in that, while it is subject to severe fluctuations in respect of output and employment, it is not violently unstable. Indeed, it seems capable of remaining in a chronic condition of sub-normal activity for a considerable period without any marked tendency either towards recovery or towards complete collapse. Now, since these facts of experience do not follow of logical necessity, one must suppose that the environment and the psychological propensities of the modern world must be of such a character as to produce these results." J.M. Keynes, *The General Theory* (1936).

Exercise 3.4: Consider Figure 3.3. Suppose you expect a low level of aggregate employment and output. Suppose, however, that you decided to 'work hard;' i.e., choose $n = \mathbf{n}_H$. Draw an indifference curve that shows the level of consumption (and utility) you can expect to enjoy and explain why it is not rational to make such a choice. Suppose instead that you expect a high level of aggregate employment and output. Further, suppose that you decide to 'slack off;' i.e., choose $n = \mathbf{n}_L$. Again, draw an indifference that shows the level of consumption (and utility) you can expect to enjoy and explain why it is not rational for you to make such a choice.

As with the earlier interpretation of animal spirits, this view of the world suggests a potential role for government policy. Suppose, for example, the economy appears to be stuck at a 'bad' equilibrium (point B). In this situation, the 'demand' for output appears to be weak and the level of employment is low. Such a scenario may rationalize a large fiscal expenditure on the part of the government. The government might do this by placing orders for output in the market, or by hiring individuals directly to work in government agencies that produce output. Either way, the demand for labor can be increased beyond the threshold $\mathbf{n}_C$. By doing so, or even by simply *threatening* to do so, individual expectations must rationally move from point B to point A. Problem solved.

5. Summary

Expectations play an important role in economic decision-making. There are two broad ways in which to think about the role that expectations play in the economy. The first view argues that expectations are **passive**; that is they are formed endogenously in light of changing economic fundamentals. People may appear to be optimistic or pessimistic over time, but if they are, it is for good reason (i.e., these sentiments reflect the changes in the underlying reality). According to this view, it makes no sense to speak of an 'expectations shock.'

The second view argues that expectations are **active**; that is, they move about for reasons that are unrelated to economic fundamentals. According to this view, it does make sense to speak of an 'expectations shock.' A naive interpretation suggests that expectations are simply exogenous and 'irrational.' A more sophisticated interpretation suggests that an expectations shock can lead to a self-fulling prophesy and, in this sense, be rational. In this latter view, expectations themselves might actually cause economic fundamentals to change. The expectations shock itself, however, is unexplained (or explained as the product of mysterious 'psychological' forces).

The second view also suggests a potential role for government stabilization policies. In the naive view, this would require that the government is somehow able to see the underlying fundamentals unlike the rest of the population. This hardly seems plausible. In the more sophisticated view, the government would have to be sure that any deterioration in economic fundamentals is primarily the product of overly pessimistic expectations. While this scenario is less implausible, in practice it is extremely difficult to differentiate the self-fulfilling expectations hypothesis from the neoclassical hypothesis. Accordingly, any government intervention should proceed with this in mind.

Problems

1. Explain the difference between an ex ante decision and an ex post outcome.

2. Imagine that you choose to place a bet and that it subsequently turns out badly. What question can you now ask yourself to decide whether placing the bet was "rational" in the first place?

3. Does the term "rational expectations" imply that rational people will never make "mistakes?" Explain.

4. According to Ken Little:[26] "Consumer spending accounts for roughly two-thirds of our economy. When consumers are reluctant to spend, our economy is affected and when they open their pocket books, the economy moves." Note that Mr. Little does not explain in his article **why** consumers may be reluctant to spend. Offer an explanation for how this guy must think about the manner in which the macroeconomy functions.

5. If Ken Little can get away with his statement above, then David Andolfatto should be able to get away with the following statement: "Labor income accounts for roughly three-quarters of our economy. When laborers are reluctant to work, our economy is affected and when they choose to work hard, the economy moves." Does this make any more or less sense than Mr. Little's statement? Explain.

6. Some people are adamant that business cycles are caused by 'aggregate demand' shocks (changes in expecations) and not by 'aggregate supply' shocks (changes in productivity). Explain how it is possible for the neoclassical model to generate business cycle behavior that might be misinterpreted as being caused by 'aggregate demand' shocks.

7. You are asked to write a newspaper article explaining under what circumstances a government stabilization policy may be desirable. Do so without the aid of any math or diagrams (and keep the amount of economic jargon to a minimal level).

[26] http://stocks.about.com/od/marketnews/a/ConsumIn122804.htm

CHAPTER 4

Unemployment

1. Introduction

In this chapter, I take a closer look at the labor market. In Chapter 2, we looked at a model that explained the determination of the level of employment. We saw there how exogenous shocks to productivity (or expectations) might generate employment level fluctuations, as changing incentives induced households to substitute time between market and non-market activities (leisure).

While the simple model in Chapter 2 constitutes a useful starting point, it is not capable of explaining several interesting facts about the way in which labor markets work in reality. The first and most obvious fact is that not everyone in an economy is employed in the labor market. That is, over any short interval of time, most (but certainly not all) people are either working "full time" or they are not working at all. In other words, the labor supply choice appears to be **discrete** (whether to work or not) rather than continuous (how many hours to work).

The second interesting property of the labor market is the existence of large **gross flows** of workers moving into and out of employment at the same time (say, over the course of one month). These gross flows of workers moving into and out of employment are largely offsetting, so that **net** changes in the level of employment are small by comparison.

The third interesting property of the labor market is the existence of **unemployed** workers. There are also large gross flows of workers moving into and out of unemployment at all phases of the business cycle.

In this chapter, I make a simple modification to the model studied in Chapter 2 to account for the coexistence of employed and nonemployed workers and the existence of gross flows of workers into and out of employment. I then extend this basic model to develop a theory of unemployment.

2. Transitions Into and Out of Employment

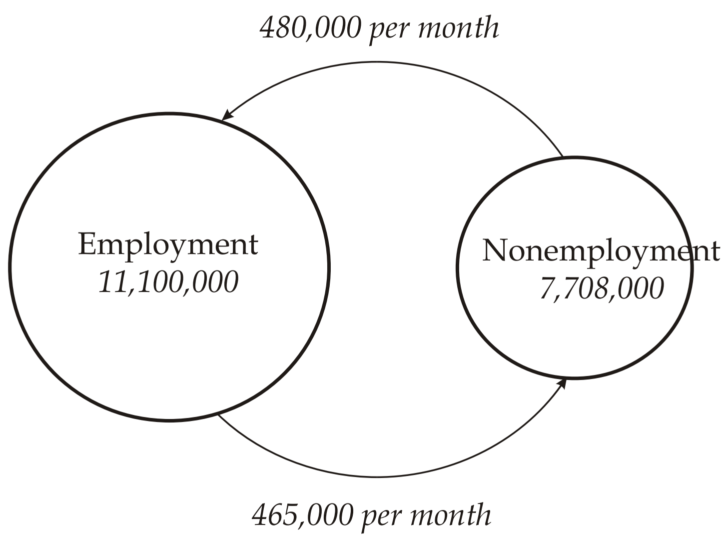
2.1 Evidence

Many countries have statistical agencies that perform monthly labor force surveys that measure various aspects of labor market activity.[27] In these sur-

[27] For example, the *Labour Force Survey* in Canada and the *Current Population Survey* in

veys, a person is labeled as **employed** if they report having done any work during the reference period of the survey (e.g., the previous four weeks). If a person is not employed, they are then labeled as **nonemployed**. Figure 4.1 summarizes the average stocks of employed and nonemployed individuals in Canada over the sample period 1976–1991.[28] The figure also records the average **monthly** flows of workers into and out of employment.

FIGURE 4.1
Average Labor Market Stocks and Flows
Canada 1976 - 1991 (Jones, 1993)



Over the sample period 1976–1991, the **net** monthly change in employment averaged only 15,000 persons (due mostly to population growth and an increase in female labor market participation). Notice how small the net change in employment is relative to the monthly **gross** flows; i.e., in a typical month, **almost one million** individuals flow into or out of employment. The existence of large gross flows that roughly cancel each other out is evidence that individuals are subject to **idiosyncratic shocks** (changes in individual circumstances) that roughly cancel out in the aggregate. In other words, even in the absence of any **aggregate shocks** (a shock that effects most people in the same way–as in Chapter 2), it appears that individuals are subject to a considerable amount of uncertainty in the labor market.

the United States.

[28]See: Jones, Stephen R. (1993). "Cyclical and Seasonal Properties of Canadian Gross Flows of Labour," *Canadian Public Policy*, 19(1): 1–17.

2.2 A Model of Employment Transitions

As in Chapter 2, consider a model economy consisting of a fixed number of people who have preferences defined over consumption and leisure (c, l). Here, I am going to assume (for simplicity) that the utility function for each person takes the following form:

$$u(c, l) = \ln(c) + vl, \tag{16}$$

where $v \geq 0$ can be interpreted as either a preference parameter measuring the value of leisure, or as a parameter measuring the productivity of a home-production technology. An important property of these preferences is that $u(c, l) = -\infty$ when $c = 0$. You can think $c = 0$ as describing a situation where someone is starving; and $u(c, l) = -\infty$ as describing the resulting consequence (a painful death). In short, starving is very painful. You might also note that $\ln(c)$ is an increasing and strictly concave function. Hence, utility is increasing in the level of consumption, but at a decreasing rate (diminishing marginal utility). On the other hand, note that this utility function is linear in leisure (constant marginal utility).

In Chapter 2, we considered the case of a representative agent (identical households). It is time now to modify this simplification. We can do this by assuming (quite plausibly) that different people may value leisure differently; that is, v may differ across people. Doing so means that agents are now **heterogeneous**. Moreover, I allow for the possibility that v may change over time for any given individual. We can describe such an event, if and when it happens, as an idiosyncratic **preference shock**. An example of such a shock might be an unplanned pregnancy (time spent at home is suddenly valued more highly than before).

People might differ in other ways too. One obvious manner in which people differ is in their skill (human capital). If a person's market wage reflects their skill level, then we can model heterogeneity along this dimension as differences in w (the real wage they could command in the labor market). An alternative interpretation of w is that it reflects the quality of a particular job match. For any given person, this parameter may also change over time; an event we might label as an idiosyncratic **wage shock**. An example of such a shock might be a construction worker who breaks his leg in an accident. Another example might be an exogenous increase in the demand for the product being produced by the firm at which the person is employed. Several other interpretations are possible.

Finally, I am going to assume that people generally differ in the level of their non-labor income; which I label here as a (for assets).[29] This point seems obvious and so there is no need to provide examples. And as above, we think of non-labor income changing unpredictably over time for a variety of reasons

[29] At this point, I suspect that some of you are cursing yourselves for criticizing the much simpler representative agent model! But this is the price we pay for added realism.

(winning the lottery, an unfavorable divorce settlement, etc.) We can think of such an event as an idiosyncratic **asset shock**.

To summarize then, I am thinking of a world where people generally differ in terms of an endowment as measured by (w, a, v). This endowment may be subject to idiosyncratic shocks over time. But we can think of these shocks as roughly cancelling out across individuals; so that there is no aggregate uncertainty. To put things another way, for every person that experiences a positive wage shock, there is someone else that experiences a negative wage shock. And likewise for the other shocks.

As in Chapter 2, individuals are endowed with one unit of time. Let n denote the time that an individual allocates to the labor market. Then individuals are assumed to face the budget constraint:[30]

$$c = wn + a. \tag{17}$$

Inserting this budget constraint into (16) together with the time constraint $l = 1 - n$ allows us to rewrite the objective function as:

$$V(n) = \ln(wn + a) + v(1 - n). \tag{18}$$

Hence, the individual's choice problem boils down to choosing an appropriate allocation of time n (just as in Chapter 2). For these preferences, it turns out that $MRS(c, l) = vc$. Hence, the labor supply function satisfies:

$$v\left(wn^* + a\right) = w;$$

or

$$n^* = \frac{1}{v} - \frac{a}{w}.$$

Exercise 4.1: Explain (do not simply describe) how the labor supply function derived above depends on the parameters (w, a, v).

So far, everything that I have done mirrors very closely the theory developed in Chapter 2. But let me now modify the model in a slight manner by assuming that time is **indivisible**. That is, suppose that it is only possible to choose either $n = 1$ or $n = 0$ (employment or nonemployment, respectively). This simple modification to the theory in fact makes solving for the individual labor supply choice much easier. All we have to do in this case is evaluate the utility payoff associated with working $V(1)$ and compare it to the utility payoff associated with not working $V(0)$; see equation (18). If $V(1) \geq V(0)$, then it is optimal to work; if $V(1) < V(0)$, it is optimal to stay at home. Because people generally differ in (w, a, v), they will make different discrete choices concerning their labor supply.

[30] Implicit in this budget constraint is the assumption that individuals cannot save or borrow and that there are no insurance markets.

Let us examine how the labor supply choice depends on w. To do this, imagine for the moment that everyone has the same (a, v); but that they differ in w. For a given w, the utility payoff associated with $n = 1$ is given by:

$$V(1) = \ln(w + a).$$

Note that $V(1)$ is an increasing and concave function in w (the payoff to working is increasing in the wage). For a given w, the utility payoff associated with $n = 0$ is given by:

$$V(0) = \ln(a) + v.$$

Note that this payoff does not depend on the person's w (the payoff to leisure does not depend on what you could have earned by working). Figure 4.2 plots these two values as a function of w.

FIGURE 4.2
Work versus Leisure

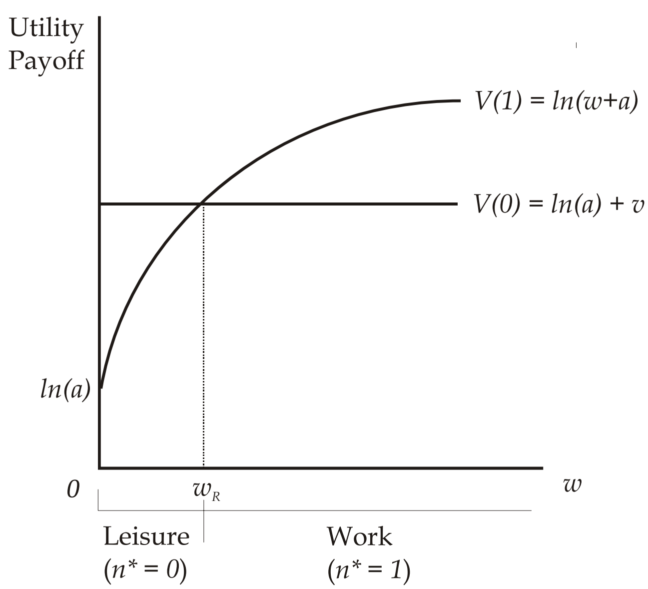


Figure 4.2 generates a very common sense result: holding constant other factors (like a and v), a person is more likely to be employed if they have a higher wage.

Figure 4.2 is useful because it helps us identify a very special wage called the **reservation wage**. The reservation wage is defined to be that wage that would make a person just indifferent between working or not. Using Figure 4.2, we see that this wage (w_R) is defined by the condition that $V(1) = V(0)$; or

$$\ln(w_R + a) = \ln(a) + v. \tag{19}$$

We can solve[31] equation (19) for w_R; i.e.,

$$w_R = (e^v - 1)a. \tag{20}$$

Notice that the reservation wage is a **function** $w_R(a, v)$; i.e., it depends on both a and v. Put differently, because people generally differ in (a, v), each person will generally have their own reservation wage.

It is important to understand that a person's reservation wage is entirely unrelated to their current wage or the wage at which they may become employed. That is, do not confuse a person's w with that person's w_R.

Exercise 4.2: Since you are attending school, most of you are not likely working full time. You can estimate your own reservation wage by asking yourself the following question: If some firm offered you a full time job, what is the minimum wage they would have to offer to induce you to quit school? The answer to this question is your w_R. Now ask yourself what the maximum wage you could earn if you were to quit school immediately and become employed. The answer to this question is your w. (I am going to guess that $w < w_R$).

The reservation wage has a very important economic interpretation. In particular, it represents the price of labor for which an individual is just indifferent between working or not. In other words, it is the minimum wage that would induce an individual to work. As such, the reservation wage is a measure of an individual's **choosiness** over different wage rates. That is, an individual with a high reservation wage is someone who is very choosy, while someone with a low reservation wage is not very choosy. What determines an individual's degree of choosiness over job opportunities? The reservation wage function in (20) tells us that there are two primary factors that determine choosiness: (1) the level of non-labor income (a); and (2) the value of time in alternative uses (v). Choosy individuals are those with either high levels of wealth or those who attach great value to non-market activities. Does this make sense? (It should).

Notice that the individual's labor supply function can also be expressed in terms of their reservation wage; i.e.,

$$n^* = \begin{cases} 1 & \text{if} \quad w \geq (e^v - 1)a; \\ 0 & \text{if} \quad w < (e^v - 1)a. \end{cases}$$

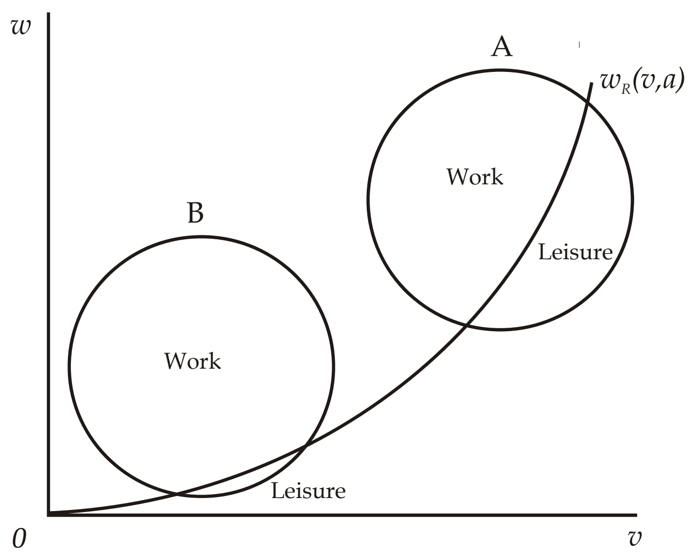
Expressing the labor supply function in this way makes it clear that labor supply tends to be increasing in w, but decreasing in both a and v (higher levels of a and v make people more choosy and therefore less likely to work at any given wage).

To see how the reservation wage function can be used to determine the level of employment in an economy, consider Figure 4.3, which plots the function

[31] Recall the following properties of logarithms: $\ln(e^x) = x\ln(e) = x$ (since $\ln e = 1$); and $\ln(xy) = \ln(x) + \ln(y)$.

$w_R = (e^v - 1)a$ (for a given level of a). In any given economy, people will be distributed in some manner over the space of (w, v) combinations (with each point in this space representing a particular type of person). Now, consider two economies A and B; and imagine that the people of these two economies all have the same level of personal wealth a. In economy A, people are distributed in the upper-right circle depicted in the figure; and likewise, in economy B, people are distributed in the lower-left circle. That is, in economy A, people have better wage opportunities and better leisure opportunities than those people living in economy B. In both economies, the level of employment is determined by the number of people who have $w \geq w_R$.

FIGURE 4.3
The Level of Employment in Two Economies



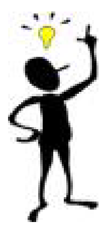
Exercise 4.3: Consider Figure 4.3. Given the way people are distributed over (w, v) in these two economies, which economy features the greater level of employment? If you had the choice of being parachuted at some random point (w, v) in either economy A or B, which would you choose? (Remember that both economies have the same level of non-labor income a).

Our theory also tells us how each person's economic welfare (maximum utility level) depends on their endowment (w, a, v). In particular, an individual's welfare is given by:

$$W(w, a, v) = \max\left\{\ln(w + a), \ln(a) + v\right\};$$

i.e., the maximum of either $V(1)$ and $V(0)$. In Figure 4.2, W is just the 'upper envelope' of the functions $V(1)$ and $V(0)$. According to our theory, the welfare function is (weakly) increasing in w, a and v. What this means is that it is impossible for an increase in any of these parameters to make an individual worse off (and will, in general, make them better off). An important implication of this result is that there is no straightforward way of linking a person's employment status with their level of welfare. Likewise, we cannot generally make statements about how two economies are performing relative to each other simply by looking at employment levels (see Exercise 4.3).

Exercise 4.4: Consider two economies A and B that are identical in every respect except that the individuals in economy A have greater levels of wealth (as measured by the parameter a). According to our theory, which economy is likely to exhibit the higher level of employment? In which economy are individuals likely to be better off? Explain. (Draw a diagram similar to Figure 4.3).



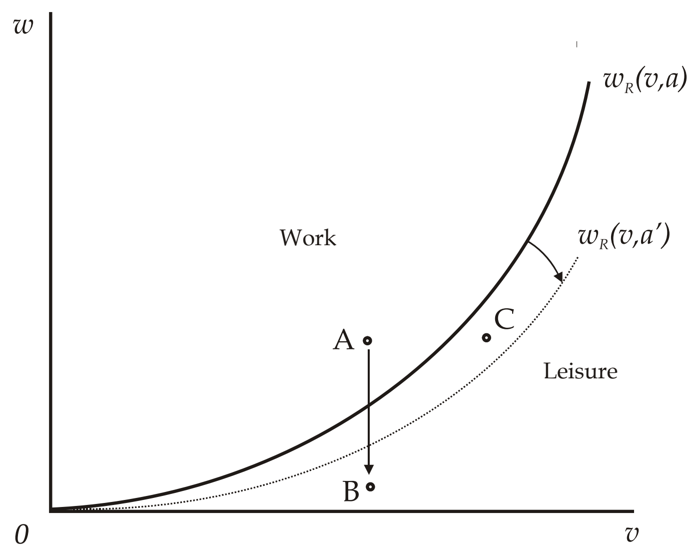
Economic welfare should be evaluated on the basis of broadly-defined consumption; not on how individuals choose to allocate their time across competing activities. There is no a priori reason to believe that high levels of employment necessarily correspond to high levels of social welfare.

Let me conclude this section by describing how our model can generate transitions into and out of employment. The basic idea is simply that individuals may, at various points in time, experience shocks that change their personal circumstances. That is, they may experience a preference shock (a change in v), a wage shock (a change in w), an asset shock (a change in a); or any combination of such shocks. Figure 4.4 illustrates two such examples; one of which results in a transition out of employment, and the other which results in a transition into employment.

Consider a person located at point A in Figure 4.4. This person is currently employed. But now imagine that this person experiences a negative wage shock, so that the person moves from point A to point B. Note that this person still has the opportunity to work at a lower wage. But given his new personal circumstances, he choose to become nonemployed. We can safely conclude that the person is made worse off as a result of the shock. He is not made worse off because he now is nonemployed (he would be made even worse off if we forced him to work at his new wage).

Now consider a person located at point C in Figure 4.4. Imagine that this person initially has wealth level a. In this case, the person is currently nonemployed–even though he has an opportunity to work. But now imagine that this person experiences a negative asset shock, which lowers his wealth to $a' < a$. Her reservation wage falls accordingly (making the person less choosy). As a result, this person will now choose to become employed at the available wage. In this case too, we can conclude that the person has been made worse off as a result of the shock. He is not made worse off because he is now employed (he would be made even worse off if we forced him to remain nonemployed).

FIGURE 4.4
Transitions Into and Out of Employment



This analysis should make clear the danger in evaluating the welfare consequences of observed transitions into and out of employment. In short, we cannot make any conclusive statement about how welfare is affected simply by looking at a person's employment status. Whether a transition into or out of employment results in an increase or decrease in welfare depends critically on the nature of the shock that resulted in a change in behavior. Moreover, it should be clear that people in this model (and presumably, in reality too) will experience changes in their personal circumstances that do not result in any labor market transition; but that nevertheless may affect their well-being in either a positive or negative manner.

3. Unemployment

3.1 Definition

There are so many misperceptions concerning the way in which unemployment is defined and measured, that it will be worthwhile to take some time to set things straight.

First, many people (primarily news reporters) rather carelessly refer to the unemployed as **jobless** (and vice-versa). In fact, jobless individuals are those who are not employed in the market sector. While these individuals are **non-employed**, they are not technically **unemployed**. If one insisted on equating the concept of joblessness with unemployment, then we are done as far as theory is concerned (since the model developed in the previous section provides a theory of joblessness).

Another commonly-held view is that an unemployed worker is someone who is not working, but **wants** to work. This concept makes about as much sense as defining an employed worker as someone who is working, but **wants** leisure.[32] Sometimes, this statement is modified to 'not working, but is willing to work at the market wage.' The problem with this modification is that it is very difficult to identify what 'the' market wage is for any given individual, since people obviously differ along so many dimensions (skill, experience, age, etc.). But in any case, the mere 'wanting' of employment (at any wage) is not sufficient to describe the state of unemployment.

The way unemployment is defined and measured is as follows. A labor market survey first asks a person whether they are working or not. If they are working (or have worked in the reference period of the survey), they are labeled as employed. If they report that they are not working, the survey then asks them what they did with their time by checking the following boxes (item 57 in the Canadian *Labor Force Survey*):

- **57** IN THE PAST 4 WEEKS, WHAT HAS ... DONE TO FIND WORK? (Mark all methods reported):
  - ☐ NOTHING;
  - ☐ PUBLIC employment AGENCY;
  - ☐ PRIVATE employment AGENCY;
  - ☐ UNION;
  - ☐ FRIENDS or relatives;
  - ☐ Placed or answered ADS;

[32] In fact, this latter concept makes more sense to me than the former. Throughout the history of mankind, the want of leisure has more commonly been seen as the relevant social problem (rather than a want of work).

☐ LOOKED at job ADS;
☐ OTHER, Specify in NOTES.

The *Current Population Survey* in the United States asks a similar set of questions. In Canada, if a nonemployed person checks off 'nothing,' then they are labeled **nonparticipants** (or not in the labor force). In the United States, if a person checks either 'nothing' or 'looked at job ads,' they are labeled as nonparticipants. If any other box is checked, then the person is labeled as **unemployed**. Clearly, a person is considered to be unemployed if: (1) they are nonemployed; and (2) if they are 'actively' searching for employment. Note that in Canada, simply 'looking at job ads' is considered to be 'active' job search, while in the United States it is not.

Notice that the survey never actually asks anyone whether they are unemployed or not. Similarly, the survey does not ask whether people 'want' to work but were unable to find work. For that matter, the survey also does not ask people whether they 'want' leisure but were unable to find leisure. Thus, among the group of nonemployed persons, the unemployed are distinguished from nonparticipants on the basis of some notion of **active job search**. Figure 4.5 provides some data for Canada over the sample period 1976–1991.

FIGURE 4.5
Average Labor Market Stocks and Monthly Flows
Canada 1976 - 1991 (Jones, 1993)

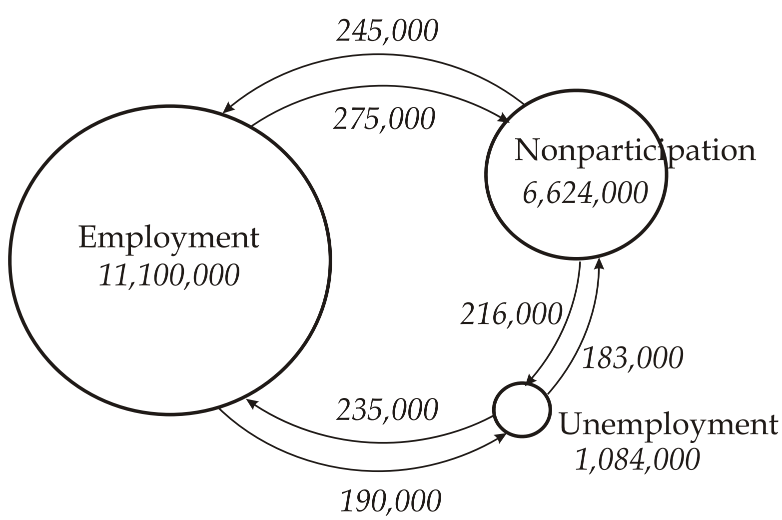


Figure 4.5 reveals a number of interesting facts. First, observe that over half of all individuals who exit employment in any given month become nonparticipants, rather than unemployed (i.e., the exit the **labor force**, which is defined to be the sum of employment and unemployment). Second, note that over half of all individuals who find employment in any given month were **not** unemployed (i.e., they find work as nonparticipants). This latter fact casts some doubt on the empirical relevance of the concept of unemployment (nonemployed persons who actively search for work). On the other hand, note that the monthly probability of becoming employed is much greater for the unemployed $(235/1084) = 0.217$ than for nonparticipants $(245/6624) = 0.037$. This fact lends support to the notion that the unemployed are more intensively engaged in job search activities.

The models that we have studied to this point are ill-equipped to deal with the issue of unemployment (at least, as the concept is defined by labor force surveys). The reason for this is because there is no need for our model individuals to engage in job search activities. In those models, including the one developed in the previous section, everyone knows where to get the best value (highest wage) for their labor. They may not be happy about the going wage for their labor, but given this wage the choice is simply whether to allocate time in the labor market or allocate time to some other activity (like home production or leisure)–whether or not the going wage is 'market-clearing' or not. To explain unemployment, we have to model the reason for why people might willingly choose to allocate scarce time to an activity like job search.

3.2 A Model of Unemployment

To simplify matters, let me modify preferences so that:

$$u(c, l) = c + vl.$$

With utility now linear in consumption, this modification implies that wealth (a) will play no role in decision-making. To see this, observe that the utility payoff associated with work is now $w + a$; and the utility payoff associated with leisure is $a + v$. The reservation wage function is, in this case, given by:

$$w_R = v.$$

All agents for which $w \geq v$ will prefer work to leisure; and all agents for which $w < v$ will prefer leisure to work. Label the former **Type A** individuals and the latter **Type B** individuals.

In the model developed above, it was implicitly assumed that all individuals are aware of the location of their best wage opportunity w. For this reason, an activity like job search literally makes no sense. There can be no unemployment in a world where everyone already knows the location of their best job opportunity.

But let me now extend the model by assuming that people are not generally aware of the location of their best job opportunity; that is, assume now that there is **imperfect information**. Moreover, let me also assume that people have an option to search for a new wage. Assume (for simplicity) that if someone chooses to search for a new job, they have to abandon their current job offer w.[33] This assumption implies that there is a cost to search (with the cost increasing in the value of your current wage offer). The benefit associated with job search is that you may find a superior job offer $w' > w$. On the other hand, you may end up with something worse than before; i.e., $w' < w$. This is to say that the outcome of the search process is **uncertain**; you may get lucky, or you may not.

Let w^e denoted the expected wage associated with the process of job search. Note that the actual wage offer you receive may turn out to be either higher or lower than what you expected. Finally, assume that after the job search process is over (at the end of the period), you have the option of either working at your new wage, or instead employing your time in leisure.

Let us now try to deduce how people are going to behave in this economy. Consider first the type A individuals. We know that these people prefer work to leisure. But do they prefer work to search, now that this option is available to them? The utility payoff to work is simply w and the **expected** utility payoff to search is w^e. Hence, type A individuals will prefer work to search if and only if $w \geq w^e$. We can identify a "reservation search wage" w_R^s by the condition:

$$w_R^s = w^e.$$

Keep in mind that the actual (ex post) wage w' may turn out to be either above or below w^e. When a new wage offer w' is in hand, the person still has the option to either work at the new wage w' or consume leisure, which yields a utility payoff v. Hence, type A people will choose to work at the new wage if and only if $w' \geq v$. Those type A people for whom $w' < v$ will choose leisure and be recorded as being unemployed (they performed no work during the period and engaged in active job search). Note that successful searchers (those for whom $w' \geq v$) will not be recorded as unemployed.

Consider next the type B individuals. We know that these people prefer leisure to work. But do they prefer leisure to search? The utility payoff to leisure is v and the expected utility payoff to search is w^e. Hence, type B individuals will prefer leisure to search if and only if $v \geq w^e$. We can identify a "reservation leisure value" v_R by the condition:

$$v_R = w^e.$$

Again, we have to keep in mind that the actual wage w' may turn out to be either above or below w^e. These people also have the option to either work at

[33] This assumption is called "no-recall" in the search literature. In a marriage market, the analogous assumption would mean that to search for a new mate, you must first abandon your current mate.

their new wage offer w' or consume leisure instead. They will prefer to work at this stage if and only if $w' \geq v$. Those type B people for whom $w' < v$ will choose leisure and be recorded as being unemployed (they performed no work during the period and engaged in active job search). Again note that the successful searchers among this group (those for whom $w' \geq v$) will not be recorded as unemployed.

FIGURE 4.6
Work, Search and Leisure

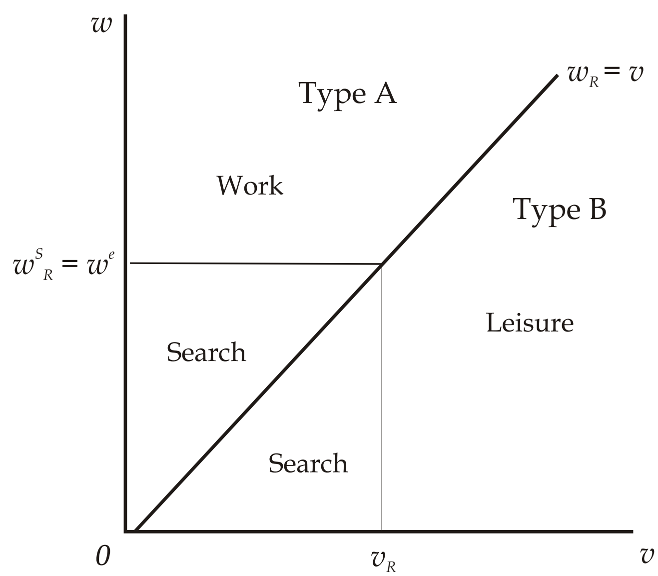


Figure 4.6 reveals the following. The individuals who are most likely to search are those who are presently poorly endowed in terms of **both** their present job opportunity and home opportunity (i.e., low values of w and v). For these individuals, allocating time to search is not very expensive (in terms of opportunity cost). The individuals who are most likely to work are those who currently have a good job opportunity (w) **and** a comparative advantage in working (w is high relative to v). The individuals who are most likely to choose leisure (nonparticipants) are those who have a good home opportunity (v) **and** a comparative advantage in leisure (v is high relative to w).

The model developed above here capable of generating labor market flows between employment, unemployment, and nonparticipation. These flows can be triggered, as before, by changes in individual circumstances (i.e., exogenous changes in w and v). At any point in time, some individuals have sufficiently poor market and nonmarket opportunities that allocating time to search activity makes sense. Note that not everyone who searches would be picked up by the *Labour Force Survey* as being unemployed. In the model, the individuals who would be classified as unemployed are those who search **and** are unsuccessful (over the reference period of the survey).

Finally, we know that individual welfare is generally increasing in both w and v. Since those people who choose to search are those with very low w and v, we can conclude that the unemployed are generally among the least well-off in society. However, it is important to keep in mind that these people are less well-off **not** because they are unemployed but because they are endowed with low w and v. In particular, the concept of 'involuntary unemployment' makes no sense (since people obviously have a choice whether to search or not). On the other hand, it may make sense to think of some people as being involuntarily endowed with poor skills or poor opportunities in the home sector. Since the choice to search is voluntary, it follows that some level of unemployment (single people) is optimal. For example, a government could in principle eliminate unemployment by forcing people to work (a policy adopted in some totalitarian regimes). While measured unemployment would fall to zero, one would be hard pressed to argue that economic welfare must therefore be higher.

3.3 Government Policy

Our model of unemployment assumed that individuals are risk-neutral (the utility of consumption is linear). We could easily extend this model, as in the previous section, by assuming that individuals are risk-averse. If individuals are risk-averse and if they are exposed to uncertainty in how their economic circumstances evolve over time (i.e., random changes in w, a and v), then they will generally want to insure themselves against such risk. Unfortunately, markets that would allow individuals to insure themselves against changes in the value of their human capital (i.e., changes in either w or v), either do not exist or appear to function poorly. The reason for why this is so is a matter of debate. One view holds that because the true value of human capital is known only to the individual (i.e., it is **private information**), individuals may have the incentive to lie about the true value of their human capital just to collect insurance. For example, if a person becomes unemployed, it is not clear whether he is unemployed for 'legitimate' reasons (i.e., a true drop in w),or whether he is simply choosing not to work and simply **reporting** a drop w. Since w is private information, the insurance company has no way of knowing the truth and hence no way of providing an insurance policy that pays for itself.

In such an environment, there may be a role for government provided insurance. Although the government must presumably cope with the same information frictions that afflict private insurance markets, the government does have one advantage over private firms. In particular, the government has the power of coercion so that it can make participation mandatory and collect 'fees' by way of taxes.[34] By operating a well-designed insurance scheme, it is conceivable that in some circumstances, the government may be able to improve economic welfare.

One likely consequence of a government insurance scheme is that it is likely to increase the unemployment rate. But in this model (and perhaps in reality too) an increase in the unemployment rate is not necessarily a bad thing. In particular, absence private insurance markets, it will likely be the case that the equilibrium unemployment rate is too low (unlucky people or people with low wealth levels are compelled to work, instead of searching for a better job).

4. Summary

Aggregate fluctuations (net changes) in the aggregate labor input are not very large compared to the gross flows of workers making labor market transitions over all phases of the business cycle. However, this apparent stability at the aggregate level masks a considerable degree of volatility that occurs at the individual level. Modern labor markets are characterized by large gross flows of workers into and out of employment, as well as large gross flows into and out of other labor market states, like unemployment and nonparticipation. These large gross flows suggest that if policy is to be desired at all, it should likely be formulated in terms of redistributive policies (like unemployment insurance and welfare), rather than aggregate 'stabilization' policies (unless one takes the view that business cycles are caused by 'animal spirits').

It is interesting to note that the notion of 'unemployment' is a relatively modern concept, evidently emerging sometime during the Industrial Revolution (c. 1800). Unemployment is sometimes viewed as the existence 'jobless' workers or 'individuals who want work.' Unemployment rate statistics, on the other hand, define the unemployed at those individuals who are not working by actively searching for work. The distinction is important because nonparticipants are also technically 'jobless.' And the concept of 'wanting work' does not make sense since work is not a scarce commodity. What is scarce are relevant skills (which largely determine the market price of one's labor). To the extent that active job search constitutes a productive investment activity, the notion that measured unemployment represents 'wasted' or 'idle' resources (as is sometimes claimed) is less than useful.

Judgements about the economic welfare of individuals or economies made on the basis of labor market statistics like employment or unemployment must be made with care. Economic well-being is better measured by the level of broad-based consumption. The level of consumption attainable by individuals depends on a number of individual characteristics, including skill, age, heath, work ethic and wealth. The overall level of productivity (technology) and government

[34] While governments make it difficult for private firms to garnishee human capital, the government itself does not restrict itself in this manner.

policies (taxes, trade restrictions, etc.), also have a direct bearing on individual well-being. The choices that individuals make in the labor market are driven primarily by their individual characteristics. Changes in these characteristics may trigger labor market responses that do not vary in any systematic way with their economic welfare.

The individual characteristics that lead individuals to be unemployed are typically such that the unemployed constitute some of the least fortunate members of a society. However, one should keep in mind that most societies are also made up of individuals who may be labelled the 'working poor.' Many nonparticipants are also not particularly well off. By narrowly focussing policies to help the unemployed, one would be ignoring the plight of an even larger number of individuals in need. To the extent that private insurance markets do not work perfectly well, there may be a role for a government 'consumption insurance' policy to help those in need (be they unemployed, employed or nonparticipants).

Appendix 4.1
A Dynamic Model of Unemployment and Vacancies

The model developed in the body of this chapter was designed to explain the determinants of employment and unemployment in the "long run." Over the short-run, the unemployment rate fluctuates (along with other aggregates) over the business cycle. Researchers have also identified another interesting phenomenon; i.e., the fact that job vacancies also fluctuates over the business cycle. A job vacancy can, in some sense, be thought of as an "unemployed job." There appears to be a strong negative relation between unemployment and vacancies over the business cycle. This negative relationship is called the Beveridge Curve.

Figure 4.7
Canadian Unemployment and
Vacancy Rates
1966 - 1988

In what follows, I develop a simple model that explain the Beveridge Curve relationship as the product of search behavior on the part of firms in response to shocks that affect labor productivity. To do so, consider a world with a fixed number of individuals who either work if they have a job or search if they do not (i.e., individuals do not value leisure). The economy is also populated by firms that either produce output (if they have a worker) or expend effort recruiting workers (i.e., if they are vacant). Each firm requires only one worker and a firm-worker pair produce a level of output equal to y (related to labor productivity).

Firm-worker pairs negotiate a wage payment that divides the output y between them. Let θy denote the profit accruing to the firm (so that $(1-\theta)y$ is the wage paid to the worker), where $0 < \theta < 1$ is now a parameter that indexes the bargaining power of the firm. After producing output in the current period, the firm-worker match survives into the next period with probability $(1-s)$, where $0 < s < 1$ is an exogenous probability of separation (the probability that some shock occurs that results in the firm shutting down). If the firm-worker pair survive into the next period, they produce output (and split it) as before. If the firm-worker match breaks down, then the firm becomes vacant and the worker becomes unemployed. If we let r denote the real (net) rate of interest, then the present value of the expected stream of profits generated by a matched firm is given by:

$$J = \frac{\theta y}{r+s}. \tag{21}$$

Exercise 4.5: Explain (do not simply describe) how the value of a firm depends on the parameters $\theta, y,$and s.

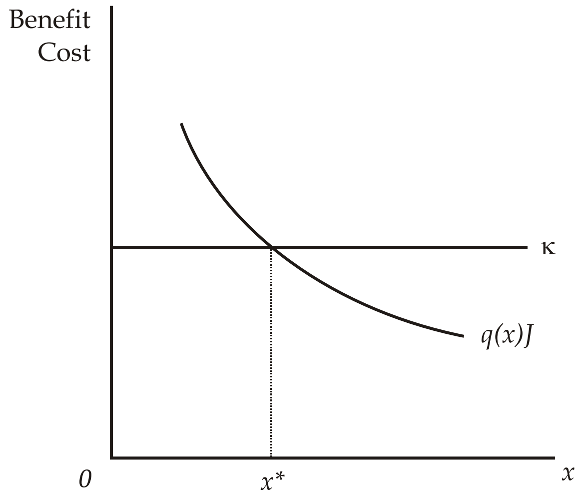
Since there is no centralized labor market, vacant firms and unemployed workers must seek each other out in a 'matching market.' A vacant firm must pay the cost κ to enter this market, but unemployed workers are let in for free (feel free to interpret vacant firms as 'single men,' unemployed workers as 'single women,' and the matching market as a 'nightclub'). Once inside the 'nightclub,' the matching technology works as follows. Let x denote the number of vacant firms (that choose to pay the entrance fee). Then a vacant firm matches with an unemployed worker with probability $q(x)$. Assume that q is a decreasing function of x, which implies that a greater number of vacancies increases competition among searching firms, and so reduces the chances of any given vacancy from 'making contact' with an unemployed worker. An unemployed worker, on the other hand, finds a suitable vacancy with probability $p(x)$, where p is an increasing function of x (the greater the number of men, the better are the chances for the ladies).

If a vacant firm meets an unemployed worker, then they begin to produce beautiful output together. Thus, from the perspective of the vacant firm, the expected gain from paying the recruiting cost κ is given by $q(x)J$. If $q(x)J > \kappa$, then it would be worthwhile for more unmatched firms to incur the cost κ, which would then lead to an increase in the number of vacancies x. But as x increases, the probability of a successful match falls. Imagine that x increases to the point x^* such that an unmatched firm is just indifferent between paying the entrance cost κ and not; i.e.,

$$q(x^*)J = \kappa. \tag{22}$$

The condition above is depicted graphically in Figure 4.8.

FIGURE 4.8
Determination of Vacancies



Exercise 4.6: Using conditions (21) and (22), show that an exogenous increase in productivity (y) leads to an increase in recruiting intensity (vacancies). Depict the change in a diagram similar to Figure 4.8. Explain the economic intuition behind this result.

One final thing we can show is what this theory implies for the evolution of the level of unemployment over time. The level of unemployment at any given point in time t is given by u_t. If we let L denote the labor force, then the level of employment is given by $L - u_t$. Since L is a constant, we are free to set L to any number; e.g., $L = 1$ (so that u_t now represents the unemployment rate). Over time, the unemployment rate must evolve according to:

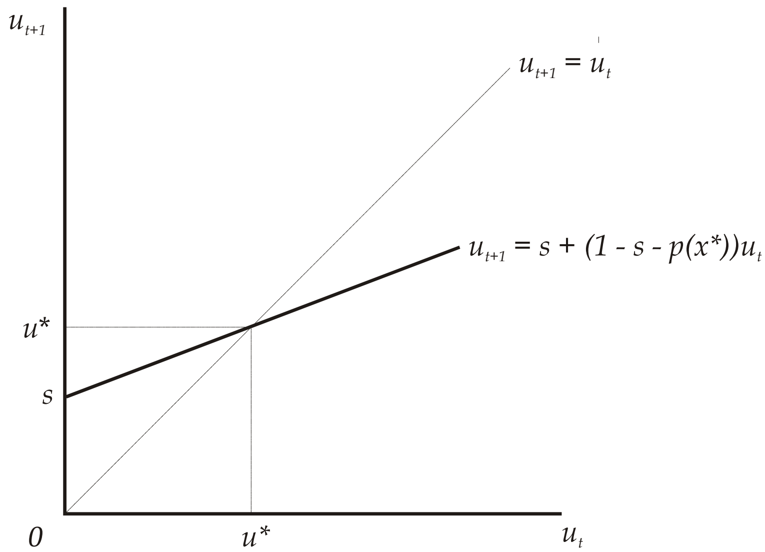
Future Unemployment = Current Unemployment + Job Destruction - Job Creation;

or,

$$\begin{aligned} u_{t+1} &= u_t + s(1-u_t) - p(x^*)u_t; \\ &= s + (1 - s - p(x^*))u_t. \end{aligned} \tag{23}$$

Equation (23) is depicted in Figure 4.9 (assuming that $0 < p + s < 1$).

FIGURE 4.9
Equilibrium Unemployment
Rate Dynamics and Steady State



Exercise 4.7: Use Figure 7.5 to show that for any given initial unemployment rate u_0, that the equilibrium unemployment rate converges to a steady-state unemployment rate u^*.

The unemployment rate u^* in Figure 4.9 is called a 'steady-state' unemployment rate because once this point is reached, the economy will stay there forever (assuming that nothing else changes). Sometimes, u^* is called the **natural rate of unemployment** (NRU). We can use equation (23) to solve for the NRU; i.e., setting $u_{t+1} = u_t = u^*$, we can calculate:

$$u^* = \frac{s}{s + p(x^*)}. \tag{24}$$

Problems

1. You read in the newspaper that 500,000 Canadians have lost their jobs in the last month. Is this a cause for concern?

2. In Figure 4.3, the average wage in economy B is much lower than the average wage in economy A. Imagine that the government in economy B decides to legislate a minimum wage law. Using the model developed above, explain how such a law is likely to influence the level of employment and economic welfare in economy B. Will the law have the effect of increasing the average wage rate among employed workers?

3. The measured unemployment rate in the United States is typically lower than the measured unemployment rate in Canada. This must mean that people in Canada are, on average, worse off than their counterparts in the U.S. Answer true, false, or uncertain; and explain.

4. In many countries, the government levies a payroll tax on firms and uses the proceeds to pay unemployed workers unemployment benefits. Using the theory developed in Appendix 4.1, explain what effect such a payroll tax is likely to have on the natural rate of unemployment (hint: a payroll tax will reduce J).

CHAPTER 5

Consumption and Saving

1. Introduction

To this point, we have focussed primarily on what one might term **intratemporal** decisions and how such decisions determine the level of GDP and employment at any point in time. An intratemporal decision concerns the problem of allocating resources (like time) across different activities **within** a period. However, many (if not most) decisions have an **intertemporal** aspect to them. An intertemporal decision concerns the problem of allocating resources **across** time. For example, deciding how much to consume today can have implications for how much will be available to consume tomorrow. The decision of how much to invest must be made with a view as to how this current sacrifice is likely to pay off at some future date. If a government runs a deficit today, it must have in mind how the deficit is to be paid off in the future, and so on. Such decisions are inherently **dynamic** in nature. To understand how such decisions are made, we need to develop a **dynamic model**.

In this chapter, I focus on the consumption-savings choice of individuals. Since any act of saving serves to reduce consumption in the present and potentially increase consumption in the future, the key decision involves how to best allocate consumption across time. We will study this choice problem within the context of a two-period model. The basic insights to be gleaned from a simple two-period model continue to hold true in a more realistic model that features many periods. To focus on the intertemporal aspect of decision-making, we abstract from intratemporal decisions. In particular, the working assumption here is that intratemporal decisions are independent of intertemporal decisions. This assumption is made primarily for simplicity and can be relaxed once the basic ideas presented here are well understood.

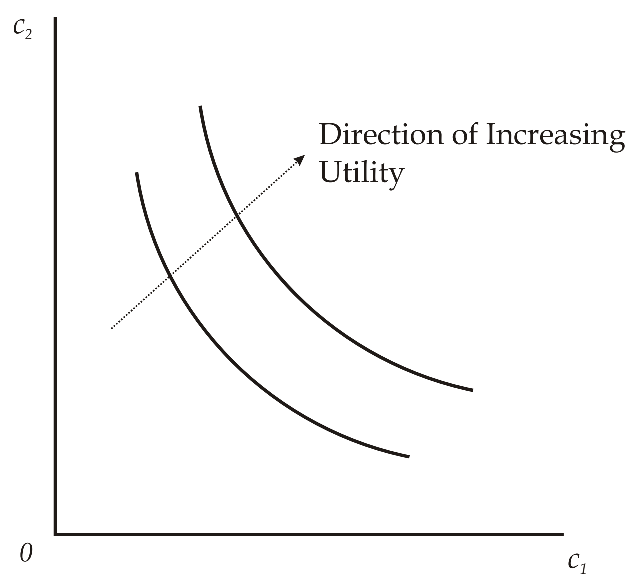
2. A Simple Dynamic Model

2.1 Preferences and Endowments

Our model economy consists of a representative household that lives for two periods (the entire duration of the economy). Each household has preferences defined over **time-dated** output (in the form of consumer goods and services). Let (c_1, c_2) denote an individual's lifetime consumption profile, where c_1 denotes 'current' consumption and c_2 denotes 'future' consumption. Note that consumption today is not the same as consumption tomorrow; **they are treated here as two distinct goods** (like apples and oranges, or output and leisure). The assumption that people have preferences for time-dated consumption simply reflects the plausible notion that people care not only for their material well-being today, but what they expect in terms of their material well-being in the future. In what follows, assume that there is no uncertainty over how the future evolves (this can be easily relaxed).

A lifetime consumption profile (c_1, c_2) can be thought of as a commodity bundle. The commodity space then is defined to be the space of non-negative commodity bundles, and can be represented with a two-dimensional graph. We make the usual assumptions about preferences; i.e., more is preferred to less, transitivity, and convexity. We will also make the reasonable assumption that consumption at different dates are normal goods and that preferences can be represented with a utility function $u(c_1, c_2)$. Figure 5.1 depicts a household's indifference curves in the commodity space.

FIGURE 5.1
Indifference Curves



Each household is endowed with an **exogenous** output profile (y_1, y_2), which constitutes a point in the commodity space. Since output is exogenous, the model constitutes an example of what is called an **endowment economy**. That is to say, output (per capita GDP) is not produced; it is simply endowed to households by Nature.[35] Assume further that output is **nonstorable**. In other

[35] Alternatively, one can think of each household as being endowed with one unit of time

words, one cannot store output across time (say, in the form of inventory). This assumption turns out to be important, and will be relaxed in a later chapter.

2.2 Robinson Crusoe

Let's pause here for a moment and talk about English literature. In the year 1719, Daniel Dafoe published his famous novel *Robinson Crusoe*, which is about a sailor shipwrecked on a deserted island. Dafoe's novel has inspired numerous books and films, including *Gilligan's Island* and *Castaway* (maybe you've seen this one–it stars Tom Hanks).

Enough with pop culture. I mention Robinson Crusoe only because his name occasionally appears in economic models as a **metaphor** for an environment that leaves a person with no opportunity to trade with others.[36] The model, as I have described it so far, has this feature. That is, assume that Robinson Crusoe (a representative household) has preferences for coconuts today and coconuts tomorrow. Coconuts are nonstorable (so that if they are not eaten, they spoil). Nature has endowed Crusoe with a coconut tree that yields y_1 coconuts today and (is expected to yield) y_2 coconuts tomorrow. Mathematically, Crusoe's choice problem can be stated very simply as:

$$\boxed{\text{Choose } (c_1, c_2) \text{ to maximize } u(c_1, c_2) \text{ subject to: } c_1 \leq y_1 \text{ and } c_2 \leq y_2.}$$

The solution to this choice problem is easy: Choose $c_1^D = y_1$ and $c_2^D = y_2$. In other words, the best that Crusoe can do is to consume his entire income in each period (with the level of income dictated by Nature). This solution is depicted diagrammatically in Figure 5.2.

Why is this interesting? Recall that a Robinson Crusoe economy is one in which the inhabitants of an island (country) have no opportunity to trade with others (in particular, households occupying other islands–or countries). Among other things, what this means is that Robinson Crusoe cannot undertake financial transactions with foreigners; i.e., the economy is **closed** to international trade in goods and assets. Such a restriction may be imposed by Nature (as in the case of Robinson Crusoe), or may be imposed by government policy (as in the case of Albania, under the rule of Enver Hoxha, 1944-85). Whatever the reason, our model makes clear that an economy that is closed in this manner **must live within its means on a period-by-period basis**. In other words,

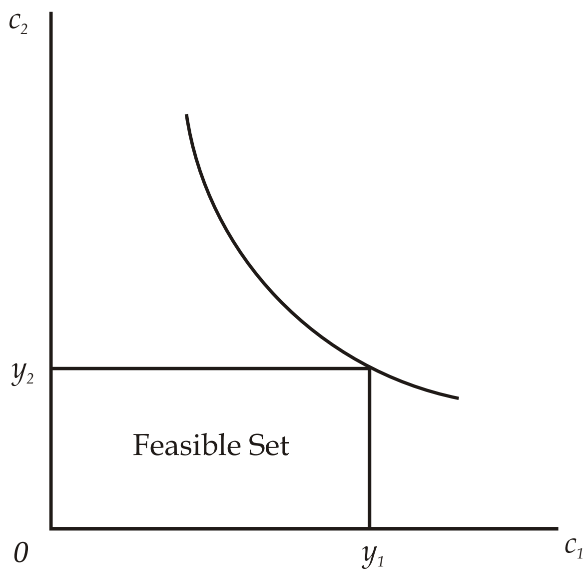
---

and a production technology $y = zn$. If households do not value leisure, then $n^* = 1$ and $y^* = z$. In other words, output is simply given by the (exogenous) level of productivity. In what follows then, we can think of shocks to y as being the result of shocks to productivity. Finally, note that since z is also the marginal product of labor, y will also correspond to labor earnings.

[36] If you do not know what a metaphor is, look it up at www.dictionary.com. A favorite of mine involves a Russian hockey scout commenting on a young Russian hockey prospect: *"Ah yes...Boris...he is strong like bull...and smart like refrigerator."*

consumption cannot exceed income at any point in time (imagine imposing such a restriction on university students).

FIGURE 5.2
Robinson Crusoe



Exercise 5.1: The marginal propensity to consume (out of current income) is defined as $MPC \equiv \Delta c_1^D / \Delta y_1$ (the change in desired current consumer spending as a result of a 'small' increase in current income). Compute the MPC for Robinson Crusoe. How does c_1^D depend on y_2 (the expected level of future income)?

2.3 An International Bond Market

In this section, I describe how Robinson Crusoe's circumstance is altered by the presence of an international financial market. The financial market is very simple: it features only one type of financial instrument–a risk-free private debt instrument called a **bond**. A bond constitutes a **promise** (made by the issuer) to deliver something of value at some future date (to the bond-holder). Since I want to abstract from bond risk, assume that bond-issuers always keep their promises (this allows us to ignore the complications that arise when default is a possibility).

We have to be clear here about the exact nature of the bond being considered. In the real-world, most bonds constitute promises to deliver money. For example, suppose that you issue a bond for \$10,000 (a student loan), which you promise to redeem (pay back) one year from now to the bond-holder (a bank). This is an example of what economists call a **nominal bond**. The \$10,000 constitutes the **principal** amount of the bond. If your debt contract with the bank requires you to pay back \$11,000, then the extra \$1000 constitutes the **nominal interest** paid on the principal. The **nominal interest rate** in this case is 10% per annum.

In our simple model, there is no role for money. People are assumed to care only about time-dated consumption (and you can't eat money–at least, it doesn't taste very good or have much nutritional content). Since money is not valued, there is obviously no role for a nominal bond. However, there is a potential role for what economists call a **real bond**. A real bond constitutes a promise to deliver **output** (i.e., in the form of future consumption). For example, suppose that you are off to a party one night and ask your roommate to lend you a case of beer (current consumption). You promise to pay your roommate back the next day with a new case of beer (future consumption). The principal amount of this loan is 12 bottles of beer. The **real interest rate** in this case is 0%. Notice that no money has changed hands in this financial transaction (note too that beer is more liquid than money).[37]

In Chapters 2 and 3, we considered a static model with two goods: output and leisure. These two goods exchanged in one market (a market in which households exchanged leisure for output). When only two goods can be exchanged, there can only be one price. In Chapters 2 and 3, this price was the real wage (the exchange rate between output and leisure). In the model studied here, there are only two goods: current consumption and future consumption. We want to think of these two goods as being exchanged in one market (a market in which households exchange current consumption for future consumption–i.e., a financial market). The relevant price in this case will be the (gross) real rate of interest (the exchange rate between current and future consumption). Let me now show this more formally.

Imagine that our representative household has access to an international bond market. In this case, the domestic economy is said to be **open** to trade (in financial instruments). Let R denote the (gross) real rate of interest prevailing in the bond market.[38] For now, we view R as an exogenous variable; i.e., the collective trading behavior of our domestic residents is assumed to have no effect on the interest rate prevailing in world financial markets. In this case, we are dealing with what economists call a **small open economy**. Assume that there is free trade in the sense that households are free to buy and sell bonds at the prevailing market interest rate.

The opportunity to buy and sell bonds implies an opportunity to save or

[37] That last part was just a joke. But in case you are questioning the empirical relevance of this example, note that many governments in fact do issue real bonds (nominal bonds indexed to the price-level).

[38] The **gross** interest rate R is related to the **net** interest rate r by the equation $R = (1+r)$.

borrow. In general terms, **saving** is defined as **current income net of expenditure on current needs**. Keep this definition in mind, since it will be useful whenever you want to think about calculating saving.

Assuming that our representative household begins time with no outstanding assets (or debt), current period income is given by the GDP. Further, since there is no government sector, expenditure on current needs is given solely by consumer spending. Thus, in the context of this model economy, we can define **net domestic saving** as:

$$s \equiv y_1 - c_1. \tag{25}$$

Note that in a closed economy (with no investment opportunities), it must necessarily be the case that $s = 0$. In an open economy, however, it is possible to have $s > 0$ or $s < 0$. In the former case, we say that the country is a **net lender**; in the latter, we say that the country is a **net borrower**. Again, note that the act of borrowing or lending current consumption in a private debt market is equivalent to selling or buying claims to future consumption (recall that a real bond represents a claim to future output).

As you all likely know from personal experience, the act of lending or borrowing has implications for your future opportunities. The same is true for a country as a whole. If a country lends s units of output today, it expects to receive the principal and interest on this loan in the future. This implies Rs units of future output in addition to whatever is produced domestically. In mathematical terms:

$$c_2 = y_2 + Rs. \tag{26}$$

In case you missed it, note that the gross interest rate R is related to the net interest rate r by the equation $R = 1 + r$.

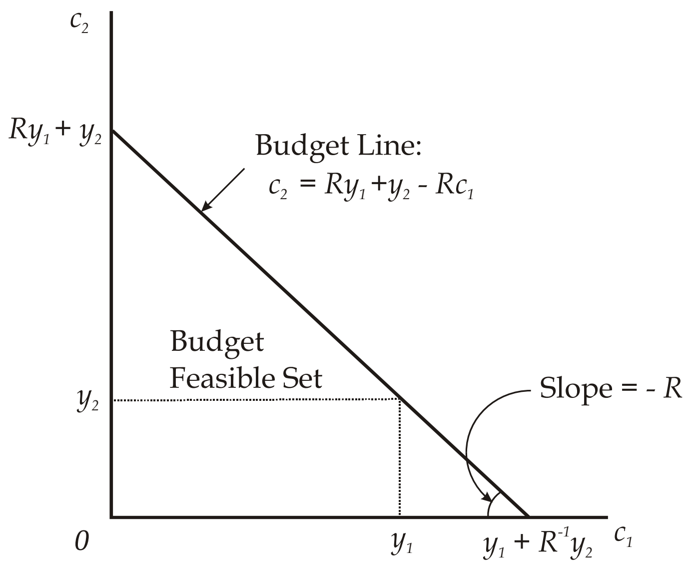
Keep in mind that equation (26) holds whether saving is positive or negative. In the case of a net borrower ($s < 0$), the quantity Rs represents the principal and interest that is repaid on accumulated debt. Thus, the act of borrowing today has the effect of reducing future consumption.

Before moving on, I would like to mention one more thing. In case $s \neq 0$, future income has two components. First, there is income that is produced domestically; this is the GDP y_2. Second, there is the income that is earned on (or paid to) financial assets (debt); this is given by rs (**net foreign interest income**). If $s > 0$, then rs represents net interest income received from foreigners; if $s < 0$, then rs represents net interest income paid to foreigners. What this implies is that in an open economy, the GDP need not equal the GNP (see Chapter 1). In particular, note that the future GNP in this economy is given by $y_2 + rs$. The s here represents nationally-owned assets 'employed' on foreign soil (a negative s implies foreign-owned assets 'employed' on domestic soil).[39]

[39] Note that since these assets are purely financial in nature (as opposed to consisting of

Alright, let's move on. Take the definition of saving in equation (25) and insert it into equation (26) to derive: $c_2 = y_2 + R(y_1 - c_1)$. This equation constitutes a budget line referred to as the **intertemporal budget constraint** (IBC). Using a diagram depicting the commodity space, this budget line has a y-intercept equal to $Ry_1 + y_2$ and has a slope equal to $-R$; i.e., see Figure 5.3.

FIGURE 5.3
Intertemporal Budget Constraint



Let me note several things of interest that are evident from Figure 5.3. The first thing to note is that the availability of an international bond market greatly expands Robinson Crusoe's intertemporal consumption opportunities (compare Figure 5.3 with Figure 5.2 in terms of what is feasible). Second, we know from basic theory that the slope of a budget line constitutes a relative price. Evidently, the gross real rate of interest R represents the **intertemporal price of consumption**. In particular, R represents the price of current consumption relative to future consumption. An increase in R makes current consumption more expensive (makes future consumption less expensive). Alternatively, think of R as the return to saving. Since saving is used to finance future consumption, an increase in R makes future consumption easier to obtain. Third, note that the IBC passes through the endowment point (y_1, y_2). Note that this must always be the case, since choosing $s = 0$ is always an option.

physical capital), they contribute nothing in the way of production (which is why they do not affect the GDP).

The IBC can be rearranged in the following useful way:

$$c_1 + \frac{c_2}{R} = y_1 + \frac{y_2}{R}. \tag{27}$$

The right-hand-side of the equation above is the **present value** of the economy's GDP flow. This is just a measure of **wealth** measured in units of current consumption (and is represented as the x-intercept in Figure 5.3). We can also measure wealth in units of future consumption; i.e., $Ry_1 + y_2$. This is called the **future value** of an economy's GDP flow (and is depicted by the y-intercept in Figure 5.3).

> *Do not confuse income (GDP) with wealth. Income is a **flow** (a sequence of numbers); wealth is a **stock** (a single number that measures the value of the income flow).*

Now, if the right-hand-side of (27) measures the present value of the GDP flow, you can pretty well guess (correctly) that the left-hand-side represents the present value of the economy's consumption flow. (Keep in that the principles involved here apply equally to an individual as they do to the economy as a whole).

The IBC imposes a restriction on behavior. It tells us that the consumption flow (c_1, c_2) must be such that its present value does not exceed the economy's wealth. Note that this restriction is much **weaker** than the one facing Robinson Crusoe in the absence of a bond market. Robinson Crusoe (an economy closed to international trade in financial assets) is constrained to live within his means on a period-by-period basis; i.e., $c_1 = y_1$ and $c_2 = y_2$. In contrast, with access to a bond market, an economy can potentially detach its consumption profile from its income profile so that $c_1 < y_1$ and $c_2 > y_2$ or $c_1 > y_1$ and $c_2 < y_2$ are possibilities. In other words, with access to an international financial market, an economy is no longer restricted to live within its means on a period-by-period basis. Instead, the economy is restricted to live within its means only in an intertemporal sense (i.e., on a **lifetime** basis).

Exercise 5.2: Consider an individual with an endowment of beer given by $(y_1, y_2) = (0, 12)$. That is, the individual has no beer today, but is expecting a shipment of beer tomorrow. If the (overnight) real rate of interest is $R = 1.20$ (a 20% net interest rate), what is the maximum amount of beer that this person can borrow today?

The exercise above demonstrates that just because one can borrow (live beyond one's current means), this does not mean that one can borrow infinite quantities. Creditors will only lend you (or countries) resources if they expect you to have the means to pay back the loan in the future. If the (gross) interest rate on a loan is R, then the **maximum** you can borrow is y_2/R. In some sense, y_2 is serving as **collateral** for the loan. **No collateral = No loan**. Or, to put things another way: **Loan + No Collateral = Charity**.

Exercise 5.3: What form of collateral do students (at least, implicitly) offer creditors (e.g., a bank or the government) when they take out a student loan? If students could default on their loans with impunity (a policy often advocated by student unions and other nut cases), what do you think would happen to the supply of student loans? Explain.

Before concluding this section, let me introduce some more terminology. Recall the income-expenditure identity $Y \equiv C + I + G + X - M$ (see Chapter 1). The value of exports X net of the value of imports M is called the **trade balance** (or net exports; let $TB \equiv X - M$. If G takes the form of consumption (expenditure on current needs), then domestic saving is defined as $S \equiv Y - C - G$. Combining these identities, we see that $S \equiv I + TB$. As there is no investment in our model economy, we have $S = TB$. In other words, net saving is only possible by running a **trade balance surplus**; conversely, net borrowing is possible only by running a **trade balance deficit**.

Exercise 5.4: Use the IBC in equation (27) together with the definition of the trade balance to show that: $TB_1 + \frac{TB_2}{R} = 0$. If a country is currently running a trade balance deficit $(TB_1 < 0)$, what does this imply about the future trade balance? Explain. Can one reasonably claim that a student who takes out a loan is like a small open economy that runs a trade balance deficit?

Related to the trade balance is the concept of the **current account position**. An economy's current account (CA) is defined as $CA \equiv GNP - C - G - I$. The only difference between these two concepts is that the trade balance is defined using GDP (Y), while the current account is defined using GNP. In our model economy, the current account (for each period) is given by:

$$\begin{aligned} CA_1 &= y_1 + 0 - c_1; \\ CA_2 &= y_2 + rs - c_2. \end{aligned}$$

Thus, it is possible, in principle at least, for a country to simultaneously be running a current account surplus and a trade balance deficit (e.g., $CA_2 > 0$ and $TB_2 < 0$). This would be the case, for example, for a net creditor nation whose citizens held a large quantity of foreign financial assets. Of course, the converse is also true.

The key thing to take away from this last bit is that the **trade balance** and the **current account** are **both inextricably linked to the consumption-saving decisions of domestic households**. Of course, it is important to keep in mind that the trade balance will also be determined by domestic investment choices as well (which is something we will deal with in a future chapter).

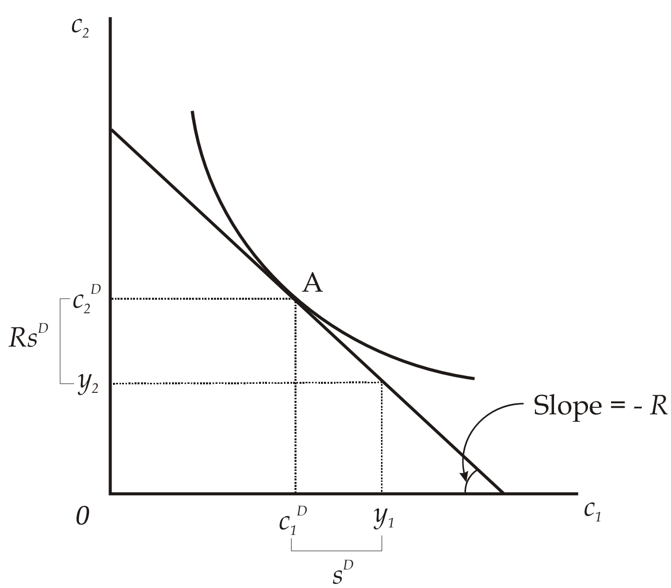
2.4 Consumption-Saving Behavior

Now that we have described the representative household's preferences and constraints (IBC), we can derive the household's optimal consumption-saving plan, conditional on a set of parameters (y_1, y_2, R). In mathematical terms, the decision-problem can be stated as follows:

$$\boxed{\text{Choose } (c_1, c_2) \text{ to maximize } u(c_1, c_2) \text{ subject to: } c_1 + \frac{c_2}{R} = y_1 + \frac{y_2}{R}.}$$

The solution to this choice problem is a pair of demand functions (c_1^D, c_2^D) that depend on the parameters (y_1, y_2, R). Once c_1^D is known, one can calculate the household's (economy's) desired saving function s^D from the definition of saving; i.e., $s^D = y_1 - c_1^D$. The solution is depicted graphically in Figure 5.4 as point A.

FIGURE 5.4
Consumption - Saving Choice



There are two mathematical conditions that describe point A in Figure 5.4. First, observe that at point A, the slope of the indifference curve is equal to the slope of the budget line. Second, observe that point A lies on the budget line. In other words,

$$\begin{aligned} MRS(c_1^D, c_2^D) &= R; \\ c_1^D + \frac{c_2^D}{R} &= y_1 + \frac{y_2}{R}. \end{aligned} \tag{28}$$

Exercise 5.5: Suppose that the utility function takes the following form: $u(c_1, c_2) = \ln(c_1) + \beta \ln(c_2)$, where $\beta \geq 0$ is a preference parameter. Explain how the parameter β can be interpreted as a 'patience' parameter. In particular, what would β be equal to for an individual who 'doesn't care' about the future?

Exercise 5.6: For the utility function in the previous exercise, one can derive $MRS(c_1, c_2) = c_2/(\beta c_1)$. Use the two equations in (28) to derive the consumer demand functions c_1^D, c_2^D and the desired saving function s^D. How is c_1^D predicted to depend on β? Explain whether or not this makes sense to you.

Let me help you get started on Exercise 5.6. Using the first equation in (28), we know that $c_2^D/(\beta c_1^D) = R$. or $c_2^D = R\beta c_1^D$. Now plug this c_2^D into the second equation in (28) to derive:

$$c_1^D = \left(\frac{1}{1+\beta} \right) \left[y_1 + \frac{y_2}{R} \right].$$

You should now be able to calculate the remaining unknowns, c_2^D and s^D.

Exercise 5.7: Using a diagram similar to Figure 5.4, depict a case in which households have a low current income, but expect a higher future income. Explain how this theory can also be used to explain why students go into debt. Is it a good thing that students are allowed to go into debt? (Use this theory to show how welfare is affected when students are prevented from taking on debt).

Figure 5.4 depicts a situation in which this economy is running a trade balance surplus in the current period (and a trade balance deficit in the future period). Whether a country runs a surplus or deficit depends on the configuration of parameters. In Figure 5.4, I have placed the endowment point in a position that implies that current GDP is significantly higher than the expected future GDP. It is not surprising then, that given the forecasted decline in GDP (falling income for domestic households), households wish to **save** in order to **smooth their consumption spending across time**. The way in which domestic households save is by purchasing foreign bonds (claims to future consumption). The way households pay for these bonds is by exporting current output–which is what accounts for the trade balance surplus in the current period.

Exercise 5.8: Consider a parameter configuration such that a small open economy is currently running a trade balance deficit. Many governments have in the past implemented **capital controls**, for example, a legal restriction that prevent foreign agents from purchasing domestic financial assets. Presumably, the idea here is that a trade balance deficit is a 'bad thing,' and that the capital controls will serve to prevent an 'excessive' deficit from occurring. Use the theory developed here to show that while capital controls can successfully reduce (or eliminate) a trade deficit, this is likely to come at the expense of lower economic welfare for domestic agents.

3. Small Open Economy Response to Shocks

The model developed above constitutes a theory of household consumer demand (and saving). Alternatively, in the context of a small open economy, we can aggregate (sum up) across households to develop a theory that explains the determination of aggregate consumer spending and the trade balance (or current account).[40] This theory takes the following form:

$$(c_1^D, c_2^D, s^D) = f(y_1, y_2, R, u).$$

Note that one of the benefits of being explicit about the intertemporal aspects of decision-making is that we can make a precise distinction between the effects of **transitory**, **anticipated**, and **permanent** changes in GDP (or productivity). In particular, note that our theory asserts that current consumer demand should depend not only on current income, but also on the level of income that is expected in the future. It follows therefore that, to an outside observers, an economy's trade balance may move for no apparent reason (when it is, in fact, responding to new **information** concerning the likely path of future GDP).

3.1 A Transitory Increase in Current GDP

In Chapter 2, we learned how a positive productivity shock could lead to an increase in GDP. Imagine here that this productivity shock is transitory (temporary) so that $\Delta y_1 > 0$ and $\Delta y_2 = 0$. Because this is a small economy, this productivity shock has no effect on R. How do the people living in our model economy react to such a development, and how does this behavior manifest itself in terms of the trade balance (or current account)?

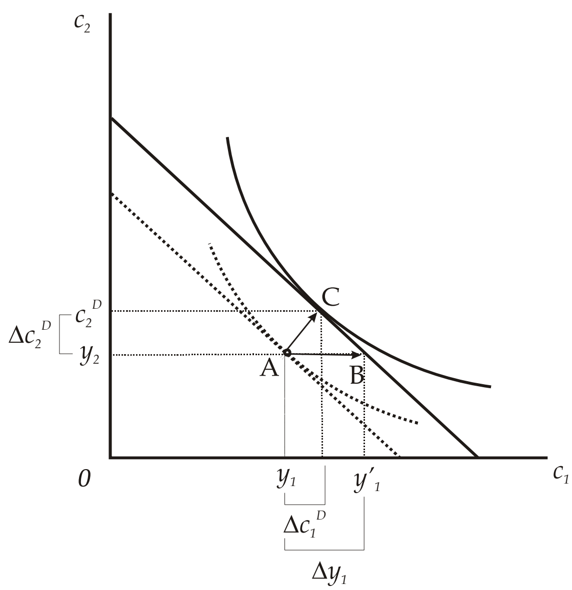
We can answer this question with the aid of a diagram. The first step is to depict the pattern of desired consumption and saving just prior to the shock; this

[40] Of course, this aggregation exercise is easy for an economy populated by a representative household. But keep in mind that one can easily extend this analysis to incorporate heterogeneous households. Doing so would not significantly alter any of the conclusions highlighted in this chapter (but would render welfare analysis a little trickier).

situation is depicted as point A in Figure 5.5. I have drawn point A such that the country is initially running a zero trade balance, but nothing important that I say below depends on this (feel free to begin with either a positive or negative trade balance). Now, beginning from this position, suppose that $\Delta y_1 > 0$. Since $\Delta y_2 = 0$, we can depict this shift as a rightward shift of the endowment (A $\rightarrow$ B). Since the interest rate is unaffected, this implies a rightward shift of the intertemporal budget constraint. Note that the shock has made domestic residents **wealthier**.

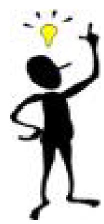
The question now is where to place the new indifference curve. If we make the reasonable assumption that time-dated consumptions are **normal goods**, then the increase in wealth results in an increase in consumer demand in **both** periods; i.e., $\Delta c_1^D > 0$ and $\Delta c_2^D > 0$. We can depict such a response by placing the new indifference curve at a point northeast of the original position; e.g., point C in Figure 5.5.

FIGURE 5.5
A Transitory Increase in GDP



We see that a **transitory** productivity shock results in a relatively **mild** but **prolonged** 'consumption boom.' Notice that the increase in current consumer demand is less than the increase in current GDP; i.e., $\Delta c_1^D < \Delta y_1$. Recall from a previous exercise that the ratio $\Delta c_1^D / \Delta y_1$ (for a 'small' Δy_1) is called the marginal propensity to consume out of current income. Since $\Delta c_1^D / \Delta y_1 < 1$, we see that a one unit increase in income results in a **less than** one unit increase in current consumer demand when the income shock is transitory. The extra income that is not consumed is saved. In this model, the extra saving takes the form of purchases of foreign bonds (hence, the country moves to a current account surplus). This foreign bond purchase is used to finance the higher consumption level that is desired in the future.

The assumption that consumption at each date is a normal good can be interpreted as a preference for **consumption-smoothing**. That is, any increase in wealth will be spread over all periods in the form of higher consumption at every date. The availability of a financial market allows households (and hence the economy as a whole) to smooth their consumption over time in response to transitory changes in their income (contrast this with how Robinson Crusoe would have to react to a similar shock). As such, one can think of financial markets as supplying a type of **shock-absorber** against transitory income shocks. That is, by saving (or borrowing) internationally, households in an economy can use the financial market to absorb the impact of transitory income shocks and in this way keep their lifetime consumption patterns relatively stable.



Access to an international financial market allows a small open economy to smooth consumption (living standards) in the face of transitory fluctuations in GDP. In this sense, the financial market serves as a type of 'shock absorber.'

Exercise 5.9: Imagine extending our 2-period model to allow for many time periods T. Consider a shock to current GDP of a given size. Explain why the response of current consumer spending is likely to grow smaller as the time horizon T is made larger.

Exercise 5.10: In a recent article, John Bluedorn[41] investigates how the current account position of small Caribbean and Central American economies react to 'hurricane shocks.' Hurricanes are not infrequent events in these parts of the world. When they hit, they invariably lead to a transitory decline in real per capita GDP (at least, controlling for several other factors). The author finds that the current account position of these economies first falls and later increases in response to a hurricane shock. Is this feature of the data consistent with our theory? Explain.

[41] www.economics.ox.ac.uk/faculty/EconDetails.asp?Detailno=173

3.2 Good News/Bad News

A 'news shock' refers to an exogenous shock to **information** concerning the future, leading households to alter their **expectations** concerning the likely path of future events. The particular news shock I want to consider here involves the arrival of new information that leads households to revise upward their **forecast** of future earnings. We can think of this as 'good news.' Everything I say below holds in reverse for 'bad news.'

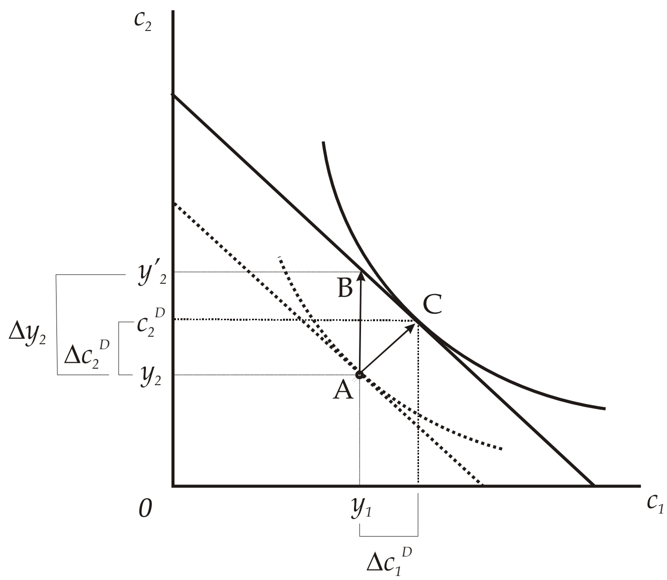
A shock to news about future productivity does not affect current productivity; we can model this by setting $\Delta y_1 = 0$. Good news here can be modeled as $\Delta y_2 > 0$. An example of such news could be what typically seems to happen just before an economy emerges from recession (households become optimistic of an impending recovery). Alternatively, one could imagine the arrival of a new technology that is expected to improve GDP in the near future. How do our model households react to such information?

Again, we can answer this question with the aid of a diagram. The first step is to depict the pattern of desired consumption and saving just prior to the shock; this situation is depicted as point A in Figure 5.6. Once again, I have drawn point A such that the country is initially running a zero trade balance (again, feel free to begin with either a positive or negative trade balance). Now, suppose that $\Delta y_2 > 0$. Since $\Delta y_1 = 0$, we can depict this shift as an upward shift of the endowment (A $\rightarrow$ B). Since the interest rate is unaffected, this implies an upward shift of the intertemporal budget constraint. Note that while the shock leaves current GDP unchanged, it nevertheless makes domestic residents **wealthier**. This example makes it clear why it is important to distinguish between **income** and **wealth**.

The question now is where to place the new indifference curve. Assuming that consumption at each date is a normal good, then the increase in wealth results in an increase in consumer demand in **both** periods; i.e., $\Delta c_1^D > 0$ and $\Delta c_2^D > 0$. This is just the **consumption-smoothing** motive at work again. We can depict such a response by placing the new indifference curve at a point northeast of the original position; e.g., point C in Figure 5.6.

Here, we see that the anticipated increase in future GDP also results in a 'consumption boom' that begins in the **current** period. The intuition for this is the same as before: the shock results in a higher level of wealth so that the consumption-smoothing motive implies that desired consumer spending in all periods rises. However, note that while consumption responds in a manner similar to when the economy is hit by a transitory shock, the behavior of savings is quite different. In particular, this shock causes a **decline** in domestic saving (so that the trade balance moves to a deficit position). Anticipating their higher future earnings, domestics increase their current consumption by borrowing from (selling bonds to) foreigners. Once again, observe how the availability of a financial market serves as a type of shock-absorber for consumption.

FIGURE 5.6
An Anticipated Increase in GDP



The example portrayed in Figure 5.6 also reveals another important point. Notice that while the trade balance (and current account) of this economy has **deteriorated** (to use language that is common in the financial pages of newspapers), the welfare of domestic residents is **higher** than before. An important lesson to be drawn here is that we must be careful in drawing any immediate link between a country's current account position and the welfare of its residents. In particular, note that a trade balance deficit may be the product of a bad shock (e.g., a recession) or of a good shock (optimism over future prospects).

A decline in the trade balance may be the product of either good or bad shocks. There is no theoretical justification for utilizing the trade balance as a measure of economic welfare.

Exercise 5.11: You receive some 'good' news that your Aunt has just passed away and left you with a huge inheritance. Unfortunately, you are able to collect on this inheritance only once you graduate (in the near future, hopefully). Since you are currently a student, your current income is rather low, you live in a cardboard box and you subsisting largely on Kraft macaroni and cheese dinners. Explain what action you could take to increase your current spending. Assume that the fact of your future inheritance is perfectly verifiable (by a bank manager, for example). Hint: collateral.

An interesting feature of real economies emerging from recession is that consumer spending often recovers before GDP does. This empirical observation is often interpreted as evidence that an increase in consumer spending 'causes' economic growth. According to our theory, however, the direction of causality actually works in **reverse**. That is, the increase in consumer spending today is caused by the arrival of information that leads households to revise upward their forecast of future income. For example, laid off workers may receive information that their former employers are planning to rehire in the immediate future. To the extent that individuals are on average correct in their forecasts, the increase in consumer spending will preceed the actual rise in aggregate income (GDP).

The example above warns us to be careful in trying to infer causality simply by looking at correlations in the data. Correlations by themselves are nothing more than measurements (descriptions) of the data; they do not constitute theory. Any particular intertemporal correlation may in fact be generated by what econometricians call **reverse causality**. To better understand the concept of reverse causality, think about the behavior of consumers during the Christmas season. It is an empirical fact that Christmas shopping preceeds Christmas. However, it would be wrong to conclude on the basis of this correlation that Christmas shopping causes Christmas. The direction of causality is obviously reversed.

> *Do not confuse correlation with causality. In particular, the fact that consumer spending is correlated with future GDP does not imply that consumer spending stimulates the economy.*

Exercise 5.12: In 1993, U.S. consumption spending increased by 2% while disposable income actually fell slightly (leading to an increase in the personal saving rate). At the time, many commentators predicted that the increase in consumption could not be sustained and that, as a result, future GDP growth was likely to slow. In contrast, Peter Ireland[42] argued that the decline in the saving rate was a sign that GDP growth was likely to remain strong. Use the theory developed above to justify Ireland's prediction (which turned out to be true, by the

[42] Ireland, Peter (1995). "Using the Permanent-Income Hypothesis for Forecasting," Federal Reserve Bank of Richmond *Economic Quarterly*, 81(1): 49–63.

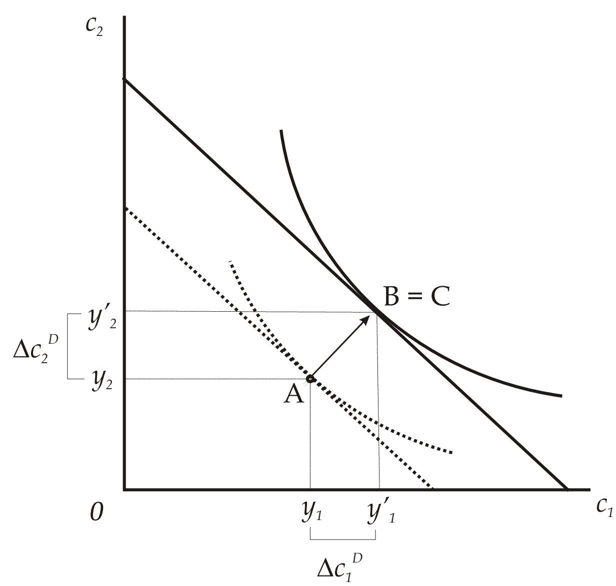
way). Hint: let preferences be such that $MRS = c_2/(\beta c_1)$ and solve for the desired saving rate s^D/y_1 as a function of (y_2/y_1).

3.3 A Permanent Increase in GDP

Imagine now that the economy experiences a productivity shock that is expected to be permanent. A permanent productivity shock can be modeled here as $\Delta y_1 = \Delta y_2 = \Delta y > 0$. Notice that a permanent shock to GDP is a combination of the two shocks studied above.

Again, we can answer this question with the aid of a diagram. The first step is to depict the pattern of desired consumption and saving just prior to the shock; this situation is depicted as point A in Figure 5.7. Once again, I have drawn point A such that the country is initially running a zero trade balance (again, feel free to begin with either a positive or negative trade balance). Now, since $\Delta y_1 = \Delta y_2 = \Delta y > 0$, we can depict this change as a 45^0 shift of the endowment ($A \rightarrow B$). Since the interest rate is unaffected, this implies an outward shift of the intertemporal budget constraint. Once again, the shock makes individuals wealthier. Note that the increase in wealth is **larger** than the case in which the shock to GDP was transitory.

FIGURE 5.7
A Permanent Increase in GDP



The question now is where to place the new indifference curve. Assuming that consumption at each date is a normal good, then the increase in wealth results in an increase in consumer demand in both periods; i.e., $\Delta c_1^D > 0$ and $\Delta c_2^D > 0$. Notice that the shift in the consumption pattern is similar to the shift in the endowment pattern. While this shift need not be precisely identical, for simplicity assume that it is. In this case, $\Delta c_1^D = \Delta y$ and $\Delta c_2^D = \Delta y$. We can depict such a response by placing the new indifference curve at a point northeast of the original position; e.g., point C in Figure 5.7.

Once again, the consumption response is similar to the other two experiments. Note, however, that the **size** of the increase in consumer spending is much larger here, compared to when the income shock was transitory. In particular, our theory predicts that the marginal propensity to consume out of current income, when the income shock is perceived to be permanent, is (approximately) equal to $\Delta c_1^D / \Delta y_1 = 1.0$. In other words, theory suggests that the marginal propensity to consume out of current income depends critically on whether shocks to income are perceived to be **transitory** or **permanent**.

Exercise 5.13: For preferences such that $MRS = c_2/(\beta c_1)$, derive the current period consumer demand function $c_1^D(y_1, y_2, R)$. Demonstrate that $\Delta c_1^D / \Delta y_1 < \Delta c_1^D / \Delta y$, where $\Delta y = \Delta y_1 = \Delta y_2$. Does this theoretical prediction make sense to you? To answer this, think of your how you are likely to behave under the following two circumstances. In scenario one, you arrive to class and spot a \$100 bill on the seat. In scenario two, you know (or expect) that there will be a \$100 bill on your seat throughout the entire semester. On the first day of class, your income is the same under both of these scenarios. I would venture to guess, however, that your spending pattern is likely to differ.

3.4 A Real Interest Rate Shock

How is an exogenous change in the **real** rate of interest predicted to influence behavior? As mentioned earlier, the real interest rate is an intertemporal price; i.e., it measures the relative price of output across different time periods. Any change in the market interest rate will have implications for the **ability** of households to substitute consumption intertemporally. As incentives for intertemporal substitution change, desired consumption and saving patterns are likely to change as well, with the corresponding implications for the trade balance (and economic welfare).

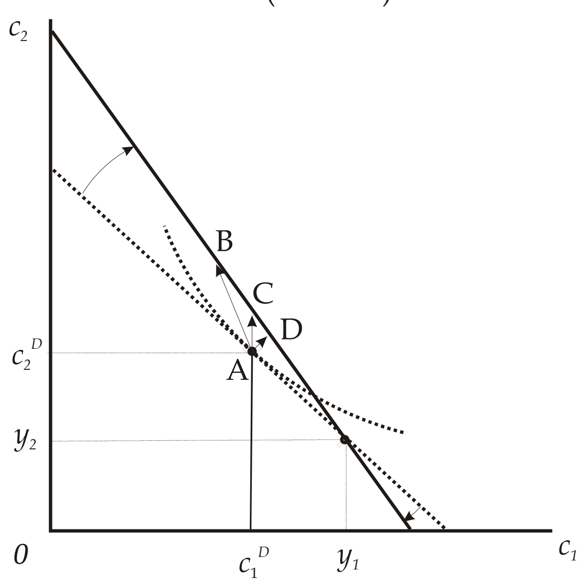
We know from basic theory that a change in prices will generally have both substitution and wealth effects. In terms of saving behavior, these two effects happen to work in the same direction for borrowers, but in opposite directions for lenders. Below, I consider each case in turn.

In our model, lenders are likely to be characterized by households with high current income and low future income (i.e., $y_1 > y_2$). A classic example here would be individuals in their peak earning years who are expecting to retire in the not-too-distant future. Alternatively, we can think of a small open economy that is currently experiencing a transitory boom in GDP.

Point A in Figure 5.8 depicts the case of a lender. If the interest rate rises, then current consumption becomes more expensive than future consumption. The substitution effect implies that people would want to substitute out of c_1 and into c_2. This applies to both borrowers and lenders. What will differ between the two cases is the wealth effect.

Observe that the effect of an increase in the interest rate on wealth depends on how wealth is measured. That is, wealth measured in present value declines, but wealth measured in future value rises. For a lender, it is appropriate to think of wealth as increasing with the interest rate. The intuition for this is that when R rises, the value of current output rises and lenders are those people who are relatively well endowed in current output. Consequently, the wealth effect for a lender implies that both c_1 and c_2 increase. Notice that while the substitution and wealth effects operate in the same direction for c_2, we can conclude that c_2^D unambiguously rises. However, the substitution and wealth effects on c_1 operate in opposite directions. Thus, c_1^D may either rise or fall, depending on the relative strengths of these two effects. Nevertheless, we can conclude that an increase in the interest rate leads to an unambiguous increase in welfare for lenders.

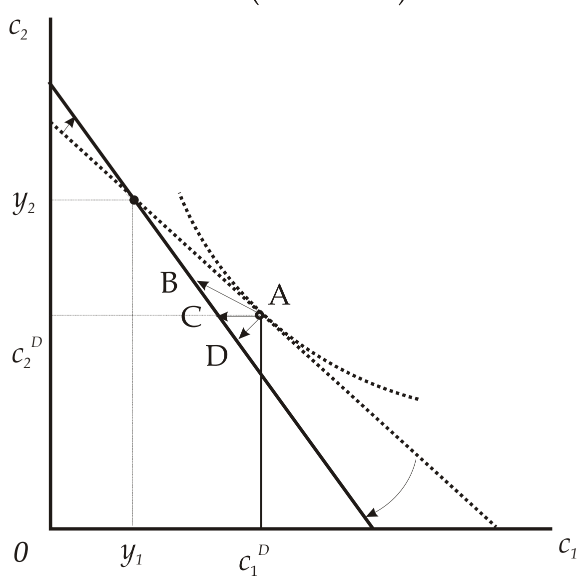
FIGURE 5.8
An Increase in the Interest Rate
(Lenders)



In our model, borrowers are likely to be characterized by households with low current income and high future income (i.e., $y_1 < y_2$). A classic example here would be young people with their peak earning years approaching in the not-too-distant future. Alternatively, we can think of a small open economy that is currently experiencing a transitory recession.

Point A in Figure 5.9 depicts the case of a borrower. The substitution effect associated with an increase in the interest rate works in the same way as before: Households desire to substitute out of the more expensive good (c_1) into the cheaper good (c_2). The difference here, relative to the case of a lender, is in the wealth effect. For a borrower, an increase in the interest rate lowers the value of the good that borrowers are relatively well endowed with (future income). Consequently, they are made less wealthy. This reduction in wealth leads to a decline in both c_1 and c_2.

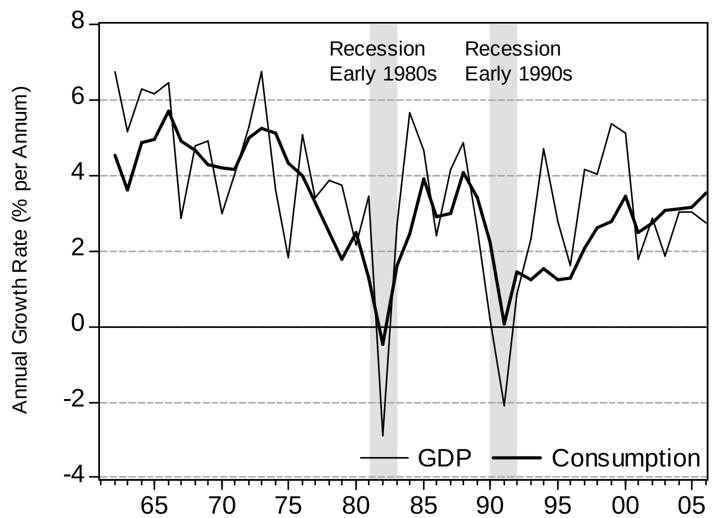
FIGURE 5.9
An Increase in the Interest Rate
(Borrowers)



Note that the substitution and wealth effect now operate in the same direction with respect to c_1 (and hence saving). Consequently, we can conclude that an increase in the interest rate leads those who are planning to borrow to scale back on their borrowing (i.e., increase their saving), so that c_1^D unambiguously declines. On the other hand, the substitution and wealth effects operate in opposite directions with respect to c_2. Therefore, c_2^D may either rise or fall depending on the relative strength of these two effects. In any case, it is clear that borrowers are made worse off (they are on a lower indifference curve) if the interest rate rises.

4. A Look at Some Evidence

FIGURE 5.10
Growth in Real GDP and Consumption
Canada (1961 - 2006)



Data from Cansim. GDP series: v3860085; Consumption series: v3860062 - v3860063 + v3860067 (excludes consumer durables and includes government purchases).

Two of the central predictions of the theory developed in this chapter are:

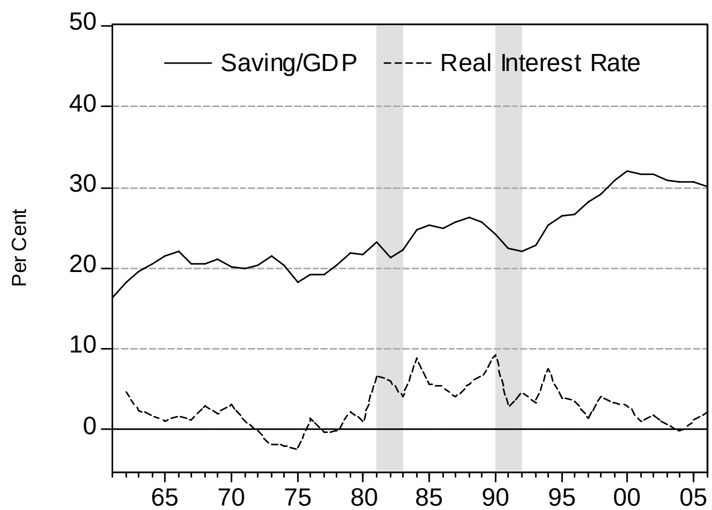
1. Consumption should grow on average at the same rate as income (GDP); and
2. Spending on consumption is not likely to vary as much as total spending (GDP).

The first prediction follows from the fact that in a growing economy, the real GDP grows to permanently higher levels. Our theory predicts that consumer spending should essentially react one-to-one with changes in GDP that are perceived to be permanent. The second prediction follows to the extent that there is a transitory (short-run) component in the level of GDP as it grows over time. To the extent that a recession is perceived to be persistent (long-lasting), consumption may fall along with income. But to the extent that a recession is not perceived to be permanent, consumption is not likely to fall as much as income. In other words, desired saving should fall during a recession (and rise during a boom).

Both of these predictions are broadly consistent with the data for Canada (see Figure 5.10) and for virtually every other country I've looked at; although, one should always keep in mind that alternative interpretations may be possible.

Our theory also predicts that the domestic saving ratio (savings as a ratio of GDP) should bear some relationship to the prevailing real rate of interest. Under a reasonable restriction, our theory predicts that–holding other factors constant (in particular, GDP)–one would expect the desired saving ratio to move in the same direction as the real interest rate.

FIGURE 5.11
Saving Ratio and the Real Interest Rate
Canada (1961 - 2006)



Real interest rate is computed as the average nominal yield on 6-month Canadian Treasuries (June of each year), minus the realized CPI inflation rate.

Real interest rates around the world were unusually low (even negative) during the decade of the 1970s and unusually high during the decade of the 1980s. There was a particularly sharp rise in the real interest in the early 1980s, that many attribute primarily to U.S. monetary policy at the time (attaching some blame to this policy for the ensuing recession). However, real interest rates remained high throughout the decade, and began to fall only in the 1990s.

From Figure 5.11, no clear-cut relationship is evident between the saving ratio and the real interest rate in Canada. While there is some evidence that the saving rate appears to be relatively low in the early part of the sample (when the interest rate is relatively low) and appears to be relatively high during the 1980s (when the interest is relatively high), these two variables appear to drift in opposite directions late in the sample period. Of course, any discrepancy might be explained by the fact that other factors (besides the interest rate) have important effects on the saving rate. On the other hand, one might view this evidence as a rejection of the theory. I am inclined to view such evidence as reflecting the likelihood that the interest rate plays only a secondary role in determining desired saving (with the expected growth rate in income playing a primary role).

5. Summary

Many, if not most, decisions involve an intertemporal dimension. Actions today can have implications for the future. Any act of saving is necessarily dynamic in nature. By saving more today, a household (country) can consume more tomorrow. Since saving more today implies less consumption today (for a given stream of income), the saving decision is related to the choice of how to allocate consumption over time. In other words, the concept of consumer demand should be thought of as the solution to a dynamic choice problem.

With the availability of financial markets, households (small open economies) are no longer constrained to live within their means on a period-by-period basis. Instead, they are constrained to live within their means on a lifetime basis. As such, financial markets provide a type of 'shock absorber' for individuals; allowing them to smooth their consumption in the face of shocks to their income. As a corollary, it follows that desired consumer spending at any point in time is better thought of as depending on the wealth of the household sector, rather than on income. Shocks to income may influence consumer spending, but only to the extent that such shocks affect wealth. From this perspective, it also follows that the impact of income shocks on consumer demand can depend on whether such shocks are perceived to be transitory or persistent.

From the perspective of an open economy, the level of the trade balance is related to domestic saving decisions. A trade balance surplus corresponds to the lending of output to foreigners; while the converse holds true for a trade balance deficit. Whether a country is in a surplus or deficit position reveals nothing about the welfare of domestic residents. A large increase in the trade balance deficit may, for example, be the result of either a domestic recession (lower welfare) or 'good news' concerning the future growth prospects of the domestic economy (higher welfare).

Appendix 5.1
Borrowing Constraints

The theoretical analysis in this chapter has assumed that households (small open economies) are free to borrow or lend at the market interest rate. In reality, however, this may not always be the case. Some economists argue that households are generally free to save, they often have trouble borrowing so that, in addition to their intertemporal budget constraint, some households (small open economies) are subject to a **borrowing constraint**.

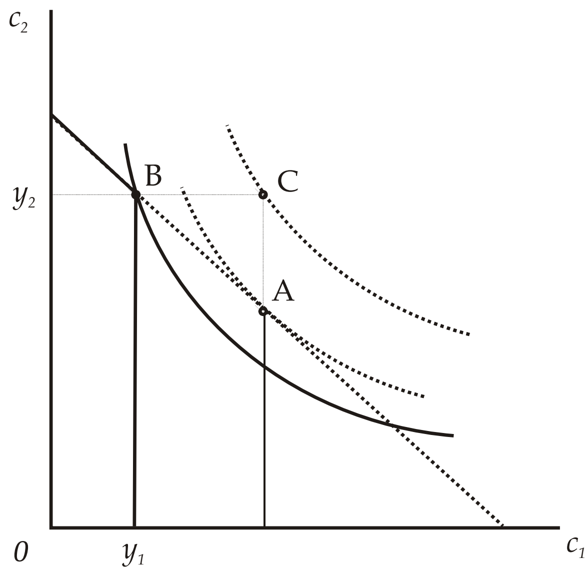
Before going further, let me be clear about the distinction between the intertemporal budget constraint (IBC) and a borrowing constraint. I have already remarked above that the IBC implies a constraint on borrowing. In particular, for a given interest rate R and future income y_2, the IBC implies that a household can borrow no more than y_2/R (the present value of its future income flow). Extending credit beyond this point is not an act of lending; but rather, it is an act of charity (since the household would not have the means to pay back the principal and interest on the debt).

A borrowing constraint refers to a situation in which a household has the **ability** to make good on a debt, but where creditors nevertheless refuse to extend a desired amount of credit **below** this level. One can model this by assuming that creditors are willing to use as collateral for a loan only some fraction $0 \leq \theta \leq 1$ of future income, so that while a household has the ability to repay a loan as large as y_2/R, creditors will only allow a loan size as large as $\theta y_2/R$. In the case where $\theta = 0$, households cannot borrow at all (even if their IBC implies that they have the ability to repay the loan).

To demonstrate how a borrowing constraint can influence behavior, consider Figure 5.10. Imagine that the household has an endowment given by point B. In the absence of a borrowing constraint, this household would like to borrow (i.e., choose point A). Clearly, point A does not violate the household's IBC. On the other hand, suppose (for whatever reason) that $\theta = 0$, so that creditors refuse to lend any resources at all to this household. Then the best that this household can do is to choose point B. In this case, we say that the borrowing constraint is **binding**.

Exercise 5.14: Consider a household that is facing a binding borrowing constraint. Explain why the marginal propensity to consume out of current income is larger for this household relative to some other household that is not similarly constrained.

FIGURE 5.12
Borrowing Constraint



Do borrowing constraints exist in reality? The answer to this question is not as straightforward as one might imagine. As an empirical matter, it is often difficult to ascertain whether a household that claims to have trouble borrowing is simply running up against its IBC or whether it is truly borrowing constrained. One problem here is that it is difficult for creditors to observe the household's future income (y_2). In particular, a household (or country, for that matter), may **claim** to have a high y_2, when in fact it does not. The world is, unfortunately, full of people wanting to 'borrow' with little intention of ever repaying their debt.

If a borrowing constraint exists, it is likely there for an important reason: a **lack of commitment** on the part of the borrower. The IBC assumes that debtors can keep their promises; either willingly, or because of a legal system that **enforces** contractual terms. It is often the case, however, that governments impose laws that prevent creditors from enforcing their claims. A law that allows individuals to declare **personal bankruptcy** constitutes one such example. These laws often allow debtors to discharge (get rid of) their debt to private creditors with virtual impunity.[43] If a potential debtor has the option of declaring bankruptcy, a creditor may question the level of commitment and

[43] It is interesting to note that one may not typically discharge debt owing to the government in a bankruptcy proceeding. The government only allows you to discharge debts owing to private sector agents. Clever, aren't they?

scale back the size of the loan accordingly.[44] For some reason, this is often referred to as a 'financial market imperfection' when, in fact, it represents the logical outcome of a government legal restriction.

The lack of commitment power is, if anything, even more severe in an international context. If your neighbor doesn't want to pay you back for a loan, you can take him to court and have a sheriff seize his assets on your behalf. If a foreign country does not want to make good on its debt, there is very little you can do to collect your money (sending in a sheriff to seize property would be tantamount to a declaration of war). Given this fact, one might be forgiven for wondering how international financial markets operate at all. Why don't debtor nations simply default on their foreign debt obligations?

The incentive for default can be seen plainly in Figure 5.10. That is, imagine that a small open economy initially sells bonds to foreigners, allowing it to import goods and attain point A. Presumably, foreigners only extended this credit in the first place because they expected to be repaid. But imagine now that our small open economy 'surprises' creditors (in the second period) by refusing to make good on its obligations. In doing so, future consumption jumps from point A to point C. This looks like a good deal for domestic agents. But if they get away with it once, do you think that they could get away with it again? Probably not (at least, not for a while). Would you be willing to lend to (purchase the bonds of) agents of a country with a track-record of default? We see then that there is some incentive to pay back your debt (even if you could default) so as to protect your **reputation**. A good reputation (good credit history) is valuable because it will likely grant you access to the loan market some time in the future when you might really need it. The same principle holds true for individuals, businesses, and governments. To the extent that this reputation mechanism works well, borrowing constraints are unlikely to be a quantitatively significant problem.

[44] In effect, the bankruptcy law legally prevents a debtor from using some or all of future earnings as collateral for a loan. As I remarked earlier: no collateral = no loan.

Appendix 5.2
Milton Friedman Meets John Maynard Keynes

Many of you have likely already encountered a theory of consumption in your introductory macroeconomics class called the **Keynesian consumption function**. The Keynesian consumption function is often specified as a relationship that takes the following form:

$$C = a + bY,$$

where $a > 0$ is a parameter that denotes 'autonomous' (exogenous) consumer spending, and $0 < b < 1$ is a parameter called the marginal propensity to consume. This consumption function embeds the common sense notion that desired consumer spending is an increasing function of income, but that a one dollar increase in income generally results in a less than one dollar increase in consumer demand. Note that this theory makes no distinction between income changes that are perceived to be temporary or permanent.

In a debate that occurred decades ago, Milton Friedman (1957) argued that consumer demand should depend on **wealth**, not income. According to Friedman, the consumption function should be specified as:

$$C = \alpha W,$$

where $\alpha > 0$ is a parameter and W denotes wealth. Thus, according to Friedman, consumer demand should be proportional to wealth and should only depend on income to the extent that income influences wealth.

We can understand both views by appealing to our theory (which builds on the early work of Irving Fisher). In particular, suppose that preferences are such that $MRS = c_2/(\beta c_1)$. Then our theory implies a consumption function of the following form:

$$c_1^D = \left(\frac{1}{1+\beta}\right)\left[y_1 + \frac{y_2}{R}\right].$$

If we let $\alpha = 1/(1+\beta)$, then we see that our theory is consistent with Friedman's hypothesis, since $c_1^D = \alpha W$, where $W = y_1 + \frac{y_2}{R}$.

On the other hand, we can rearrange our consumption function in the following way:

$$c_1^D = \left(\frac{1}{1+\beta}\right)\left(\frac{y_2}{R}\right) + \left(\frac{1}{1+\beta}\right) y_1.$$

If we define $a = \left(\frac{1}{1+\beta}\right)\left(\frac{y_2}{R}\right)$ and $b = \left(\frac{1}{1+\beta}\right)$, then we see that our consumption function also agrees with Keynes; i.e., $c_1^D = a + by_1$.

While the two theories look similar, they can in fact have very different implications for consumer behavior. For example, consider two individuals that have the same level of wealth but different lifetime income patterns. The Friedman consumption function implies that these two individuals should have the same level of consumption, while the Keynesian consumption function implies that the person with the higher current income should have higher (current) consumer demand.

Our theory is consistent with Friedman's hypothesis when households are not debt constrained. But if households are debt constrained, then our theory supports Keynes' hypothesis. In any case, our theory is to be preferred over either because it makes explicit where the parameters a, b and α come from, as well as stating the conditions under which either hypothesis may be expected to hold.

Exercise 5.15: Friedman's theory of consumption is sometimes called the "permanent income hypothesis." To understand how permanent income relates to wealth, consider a person with an income flow (y_1, y_2) that generates wealth $W = y_1 + y_2/R$. Permanent income is defined to be some hypothetical constant lifetime income flow y that generates wealth W; i.e., $y + y/R = W = y_1 + y_2/R$. Solve for this person's permanent income y and show that it is proportional to wealth W. Show on a diagram how permanent income can be identified as the intersection of a 45° line through the intertemporal budget constraint.

Problems

1. Canada entered a recession in the early 1990s. A Toronto newspaper article explained events this way: "Canada has just entered a recession. Adding to our problems is the ballooning current account deficit." Explain why the 'ballooning' current account deficit was likely a 'good' thing. Hint: model the recession as an exogenous $\Delta y_1 < 0$ and evaluate economic welfare under two scenarios: one in which the current account moves into deficit; and one in which the government prevents domestic residents from selling bonds to foreigners (so that the current account position remains in balance).

2. Consider the following quote from a recent commentary by James Arnold (BBC News): "Consumer spending is certainly the foundation of many economies. The long boom of the mid to late 1990s was built on buoyant spending - especially in the US and UK, where service industries have long replaced manufacturing as the main economic motor. Similarly, the predicted slump in consumer spending is seen as the main threat now, as the US attacks (9/11) crunched into an already-vulnerable global economy." (Note: the predicted slump in consumer spending did not materialize). The quote seems to suggest that economic growth is driven by (presumably exogenous) consumer spending. Offer a critique of this perspective.

3. Consider a small open economy populated by N people with preferences $u(c_{1i}, c_{2i}) = \ln(c_{1i}) + \beta \ln(c_{2i})$ for $i = 1, 2, ..., N$. Each person has a potentially different earnings stream (y_{1i}, y_{2i}).

   (a) Calculated the current period demand function for each person i, $c_1^D(R, y_{1i}, y_{2i})$.

   (b) Calculate the aggregate demand for current period consumption and demonstrate that it only depends on the time path of GDP (and not on how the GDP is distributed across people).

   (c) Calculate the desired saving function for each person, s_i^D. Now, assume that a fraction λ of the population have an income stream $(y_1, 0)$; and that the remaining fraction $(1-\lambda)$ have an income stream $(0, y_2)$. In a closed economy version of this model, aggregate net saving must sum to zero. Use this condition to derive an expression for an equilibrium real interest rate R^* that will equate the supply and demand for credit. How does R^* depend on parameters? Explain.

CHAPTER 6

Fiscal Policy

1. Introduction

The term "fiscal policy" refers to those aspects of government that are broadly concerned with the collection and disbursement of money. These are just fancy words for "taxing" and "spending." It is important to keep in mind that almost every act of government taxation and spending has a redistributive component to it. This is most obvious in the case of transfer programs (e.g., social assistance). But it also holds true to some extent for other expenditures (e.g., hospitals or bridges that service a local community). Because of this, politics usually plays an important role in shaping the nature of fiscal policy.

Fiscal policies are in reality multidimensional objects. Moreover, they are typically implemented at all levels of government; i.e., federal, provincial or state; and municipal. In spite of all this heterogeneity, the common principle remains taxation (the acquisition of resources) and spending (the disbursement of resources); and this shall constitute the focus of this chapter.

This is not a chapter designed to explain why people collectively erect institutions that are called "governments." To address this difficult question would take an entire textbook. Instead, I focus on the more modest question concerning the likely economic effects of different fiscal policies, without explaining why these fiscal policies are put in place to begin with. In short, I am going to treat government behavior as exogenous. Needless to say, such a treatment is not entirely satisfactory, but it seems like a good place to start.

2. Accounting and Data

Almost all of government **income** is collected in the form of **taxes**. Government **spending** can be broadly classified into two components: **purchases** of output; and **transfers**. It is important to understand the difference between government purchases and transfers. A government purchase G involves an expenditure on output (goods and services). A government transfer A involves no expenditure on output; it simply constitutes a transfer of existing output from one set of agents to another. For this reason, only government purchases are counted as a component of the economy's total expenditure; recall, from Chapter 1, the income-expenditure identity:

$$Y \equiv C + I + G + NX.$$

Hence, total government spending can be denoted as $E \equiv G + A$ (the sum of purchases and transfers). If we let T denote total government income (tax revenue), then government sector saving is defined as:[45]

$$S_G \equiv T - A - G.$$

Here, T denotes **gross** tax revenue and $(T - A)$ denotes **net** tax revenue.

Private sector saving is defined as:

$$S_P \equiv Y + A - T - C.$$

Hence, the domestic saving identity is given by $S \equiv S_P + S_G$; or

$$S \equiv Y - C - G;$$

or, equivalently,

$$S \equiv I + NX.$$

The key point to note here is that while $(T - A)$ constitutes net income for the government, it constitutes a net expense for the private sector. In the aggregate, these two quantities cancel each other out; which is why taxes and transfers do not appear in the aggregate income-expenditure identity.

Table 6.1 depicts the Canadian government's (consolidated across federal, provincial, and municipal levels) income statement for the year 2005.

Table 6.1 Consolidated Government Income Statement
Canada 2005 (millions of dollars)

Income	
Direct taxes from persons	164,979
Direct taxes from corporations	49,492
Direct taxes from non-residents	5,478
Contributions to social insurance	65,340
Taxes on production and imports	173,081
Other taxes	10,442
Investment Income	48,446
Total Revenue	**517,258**
Spending	
Goods and services	262,650
Transfers to persons	134,766
Transfers to businesses	16,900
Transfers to non-residents	4,700
Interest on Debt	62,765
Total Expenditure	**481,791**
Surplus (Deficit)	**35,467**

Source: CANSIM II 3840004.

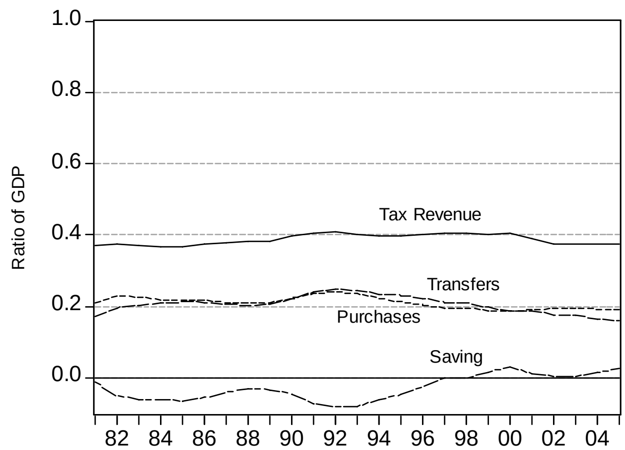
[45] Technically, government income should include investment income net of interest payments on outstanding debt. I also assume that to a first approximation, all of G constitutes spending on consumer goods and services.

To place the numbers in Table 6.1 in context, note that Canada's GDP for 2005 was 1,375,080 million dollars. Hence, we have the following rough breakdown:

$\left(\frac{T}{Y}\right)$	$\left(\frac{517{,}258}{1{,}375{,}080}\right)$	0.38
$\left(\frac{G}{Y}\right)$	$\left(\frac{262{,}650}{1{,}375{,}080}\right)$	0.19
$\left(\frac{A}{Y}\right)$	$\left(\frac{219{,}141}{1{,}375{,}080}\right)$	0.16
$\left(\frac{S_G}{Y}\right)$	$\left(\frac{35{,}467}{1{,}375{,}080}\right)$	0.03

Thus, almost 40% of all income generated in Canada is taxed away. Out of this tax revenue, roughly 50% is used to purchase output (the bulk of these purchases are provided without charge to the public) and with the remainder distributed as transfers (there was also a small surplus). Figure 6.1 shows how these ratios have evolved over time since 1981.

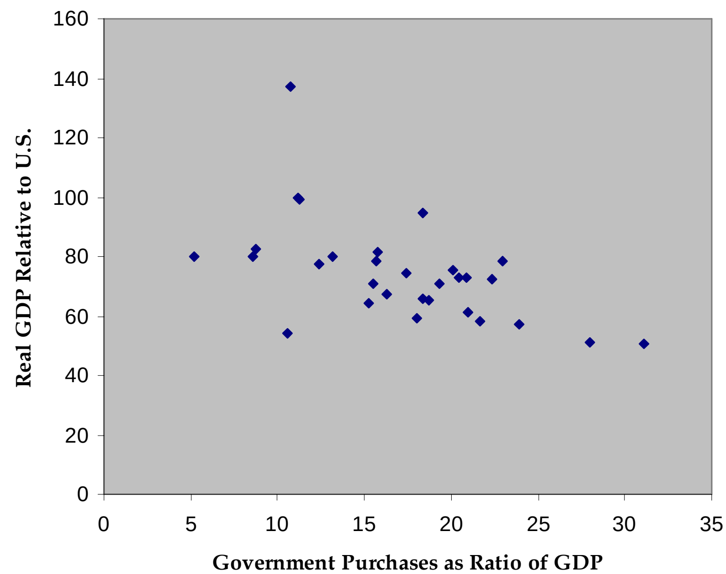
FIGURE 6.1
Government Revenue and Expenditure
Canada 1981 - 2005



As you can see, there is really nothing very exciting happening in Canada with respect to fiscal policy over the last couple of decades. I suppose that this is a good thing–assuming that you like boring countries. There is some evidence of a modest contraction in the relative size of the public sector (as measured by the sum of purchases and transfers) beginning some time in the mid 1990s. More recently, we have seen a modest decline in the (gross) tax rate and a move from a deficit to surplus position. A part of this might be explained by a general political climate amenable to the idea that cutting taxes and paying down government debt is likely to stimulate economic activity.

Perhaps a better idea of the relationship between economic activity and the size of government can be gleaned by examining a cross section of countries. For this purpose, I make reference to the Penn World Tables, which provides various data for a large set of countries.[46] In Figure 6.2, I plot for several countries in 2005 the country's real GDP (measured relative to U.S. GDP) and government purchases as a ratio of GDP.[47] In this figure, I limit attention to the set of countries that have a GDP that is at least 50% that of the U.S.

FIGURE 6.2



The relationship in Figure 6.2 is not perfect, but the data does suggest a negative relationship between the level of GDP and the size of government (the country at 140% of U.S. GDP is Luxembourg). In the next section, I develop a simple economic model that might be used to interpret this relationship.

[46] Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

[47] For Canada, the ratio of government purchases to GDP is recorded as 0.13, so I presume that this data refers to purchases at the federal level only.

3. Government Spending and Taxation

The model I have in mind here is an extension of the simple static model developed in Chapter 2. Since there is no saving in this model, $S_P = S_G = 0$.

Let us begin by assuming that the government sector 'demands' g units of output, where g is exogenous (and takes the form of consumer goods and services). There are two ways of viewing the production of output destined for the government sector. The first way is to suppose that all output y is produced by the private sector and that the government sector purchases the output it desires g from the private sector. The second way is to suppose that the government produces the output it needs by employing workers (public sector workers). If the government has access to the same production technology as the private sector, then either approach will yield identical results.

Since government purchases (or production) of output are typically distributed in some manner across households, we need to address the question of how individuals in the household sector value g. The answer to this question will depend on the nature of the purchases made by the government. In reality, the government allocates g to a wide variety of uses, ranging from outright waste to the delivery of an assortment of valuable goods and services. In what follows, I am going to assume that household preferences can be represented by a utility function of the following form:

$$u(c, l) + v(g), \tag{29}$$

where v is an increasing and concave function.

The question that concerns us here is the following: How does an exogenous increase in government spending (an **expansionary fiscal policy**) affect output and employment? As it turns out, the answer to this question depends critically on the nature of the tax instrument used to finance the expansion in government purchases. In what follows, I consider two such instruments. The first is a **lump-sum tax** (also referred to as a head-tax). Because lump-sum taxes are rarely used in practice, this case is primarily hypothetical; nevertheless, it serves as a useful benchmark since lump-sum taxes work exclusively through their effect on household wealth (a force that is likely to be present in any type of tax). The second is an **income-tax**. Studying this type of tax is interesting because it is realistic and has additional interesting effects on incentives.

3.1 Expansionary Fiscal Policy: Lump-Sum Tax

Imagine that government spending is financed with a **lump-sum** tax τ on households. The key feature of a lump-sum tax is that the amount of tax paid by the household in no way depends on household behavior. What this means is that the household has no incentive to change its behavior in an attempt to escape its tax burden. Since a lump-sum tax does not distort incentives, it is called **non-distortionary**.

The key restriction on government behavior is given by the **government budget constraint** (GBC). That is, if the government desires g units of output, it must levy a net tax τ on the household sector sufficient to cover its desired expenditure. Mathematically, what this implies is:

$$\tau = g. \tag{30}$$

Now, let's consider a representative household. The household takes as given the prevailing wage rate w, non-labor income d, and the level of taxes τ. The household's budget constraint is given by:

$$c = w(1-l) + d - \tau;$$

where labor supply is given by $n = 1 - l$. Hence, the choice problem facing the representative household can be stated formally as:

$$\boxed{\begin{array}{c}\text{Choose } (c,l) \text{ to maximize } u(c,l)+v(g)\\ \text{subject to: } c=(1-l)w+d-\tau.\end{array}}$$

The solution to this choice problem is a pair of demand functions: $c^D(w,d,\tau)$ and $l^D(w,d,\tau)$. Once the demand for leisure is known, one can compute the supply of labor $n^S(w,d,\tau) = 1 - l^D(w,d,\tau)$.

Mathematically, the choice (c^D, l^D) is the solution to:

$$\begin{aligned} MRS(c^D, l^D) &= w; \\ c^D &= (1-l^D)w + d - \tau. \end{aligned}$$

This should look very familiar to you.

Exercise 6.1: Depict the solution (c^D, l^D, n^S) on a diagram and show how n^S is predicted to respond to an increase in τ.

Having characterized optimal behavior on the part of households, the next step involves examining the behavior of the business sector. As it turns out, there is not much work to do here since the choice problem for a representative firm remains exactly the same as in Chapter 2. In fact, one can use the reasoning developed there to conclude–in this simple model, at least–that this fiscal policy will have no effect on the equilibrium real wage or profits (verify this fact as an exercise).[48] Accordingly, we can conclude that $(w^*, d^*) = (z, 0)$.

[48] This will not be the case for the more general model studied in Appendix 2.1. Nevertheless, the general conclusions arrived at here (regarding the effects of government spending on output and employment) continue to hold in even the more general model.

The final step is to invoke the government budget constraint, which in this case is simply $\tau = g$. Now let us combine all of these restrictions to determine the equilibrium allocation (c^*, l^*) :

$$\begin{aligned} MRS(c^*, l^*) &= z; \\ c^* + g &= z(1 - l^*). \end{aligned} \tag{31}$$

The equilibrium level of employment can then be calculated as $n^* = 1 - l^*$; with the equilibrium level of GDP given by $c^* + g = y^* = zn^*$. Notice that the model studied in Chapter 2 is just the special case for which $g = 0$.

Exercise 6.2: Given preferences such that $MRS(c, l) = c/l$, use (31) to solve for (y^*, n^*) as a function of g.

Since g is a parameter (an exogenous variable), the equilibrium level of output and employment (y^*, n^*) will, in general, depend on g. One way to investigate this dependence is through the use of algebra, as in Exercise 6.2. This approach requires us to specify an explicit mathematical form for the MRS. But we can also investigate this dependence by way of a diagram (exploiting the fact that both consumption and leisure are normal goods). To see how this can be done, consider the following.

Imagine that initially, $g = 0$. Then point A in Figure 6.3 depicts the equilibrium allocation. Notice that, as in Chapter 2, the **equilibrium** budget line in this case corresponds to the economy's production possibilities frontier (PPF).[49] Suppose now that the government embarks on an expansionary fiscal policy by increasing spending to some positive level $g > 0$. Because the government spending program requires a tax on households, the budget constraint moves downward in a **parallel** manner (leaving the PPF in its original position).[50] If consumption and leisure are normal goods, then the new equilibrium is given by point B in Figure 6.3.

Thus, according to this theory, an expansionary fiscal policy **financed by a lump-sum tax** induces an increase in output and employment. However, note that the policy also induces a decline in consumer spending. From the income-expenditure identity, we know that $c^* + g \equiv y^*$ in this model economy (there is no investment and no foreign sector). While an increase in g results in an increase in y^*, it also appears to partially **crowd out** private sector spending (so that c^* falls, but by less than the increase in public sector spending, g).

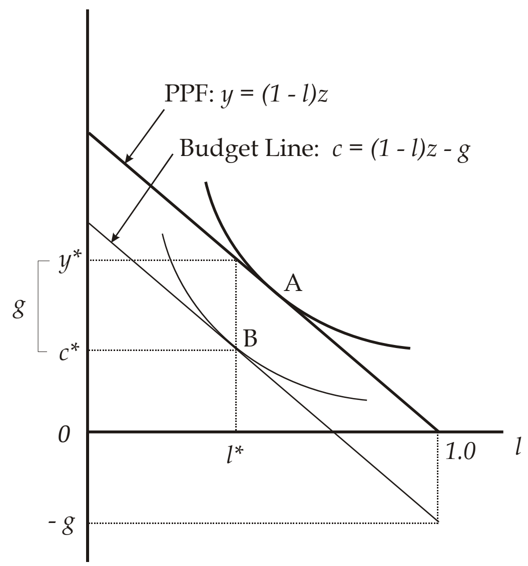
The basic force at work here is a **pure wealth effect** (there is no substitution effect here because the real wage remains unchanged). In particular, because the **after-tax wealth** of households declines, they naturally demand

[49] Again, note that this will generally not be the case; i.e., see Appendix 2.1.

[50] The PPF remains in its original position because government spending does not affect the technology available for producing output.

less consumption and leisure (assuming that these are both normal goods). To put things another way, the tax on households compels them to work harder so that they might mitigate (but not eliminate) the impending decline in disposable income. This increase in the supply of labor is what leads to the expansion in GDP.

FIGURE 6.3
Expansionary Fiscal Policy
Financed with Lump-Sum Tax



Note that while this expansionary fiscal policy causes GDP and employment to rise, it is not immediately clear how welfare is affected. Recall that welfare here can be measured by $u(c^*, l^*)+v(g)$. On the one hand, the decline in personal consumption and leisure suggests that welfare will decline. On the other hand, to the extent that households value government purchases, welfare will increase. The net effect on welfare will depend on which of these two effects is stronger. In any case, it is once again useful to stress the following:

> *Do not confuse GDP with economic welfare!*

3.2 Expansionary Fiscal Policy: Income Tax

One might be inclined to think that the result described above could have been derived much more easily by simply examining the income-expenditure identity for this simple economy: $y \equiv c + g$. Does it not follow from this identity that an increase in g must lead to an increase in y? The answer is **no**.

To see this, let us now consider a more realistic case where government taxation is **distortionary**. An important example of a distortionary tax is an income tax. Consider an economy where the income tax rate is 20%. What this means is that for every dollar that is earned in income, 20 cents must be paid to the government. The more income you make, the more you pay (unlike a lump-sum tax). An income tax effectively reduces the **after-tax return** to market activity. In the present model, it **distorts** the relative return between consumption and leisure. These types of distortions alter economic incentives; and people can be expected to act accordingly.

Let us return to the model developed above. We can already guess that $(w^*, d^*) = (z, 0)$ as before, so let me impose this here. One important difference, however, is that $w^* = z$ now represents the **gross** (pre-tax) wage. If the government imposes an income tax rate $0 \leq \theta \leq 1$, then the household's budget constraint is given by:

$$c = (1-\theta)z(1-l).$$

Here, $(1-\theta)z$ denotes the **after-tax wage** (or return to labor). Note that, as far as the household is concerned, an increase in the tax rate θ will have much the same effect as a decrease in the real wage z.

Exercise 6.3: On a suitable diagram, depict the economy's PPF and the representative household's budget constraint. Verify that the slope of the budget constraint is equal to $-(1-\theta)z$. Demonstrate that maximizing utility implies that the indifference curve is tangent to the budget line (not the PPF).

From the exercise above, we know that the solution to the household's choice problem is now characterized by:

$$\begin{aligned} MRS(c^D, l^D) &= (1-\theta)z; \\ c^D &= (1-\theta)z(1-l^D). \end{aligned}$$

The revenue collected by the government is now given by $\tau = \theta z(1-l^D)$. Hence, the government budget constraint implies that:

$$\theta z(1-l^D) = g.$$

All three of these conditions must hold in equilibrium. Therefore, for a given (exogenous) g, we can think of these three equations as determining the three unknowns (c^*, l^*, θ^*); i.e.,

$$MRS(c^*, l^*) = (1-\theta^*)z; \tag{32}$$

and

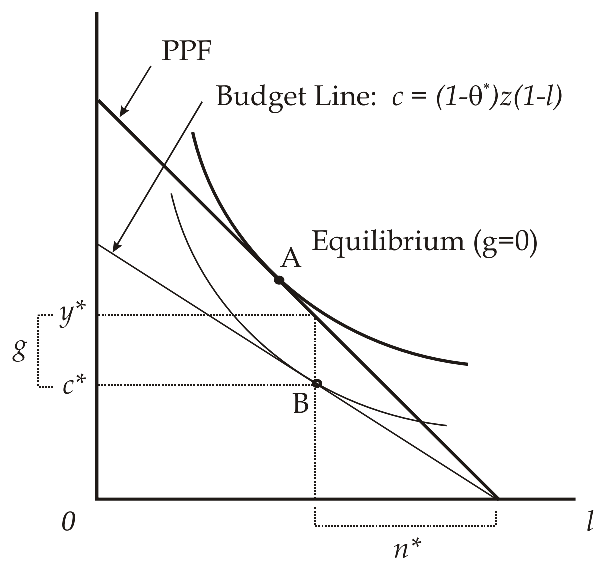
$$\begin{aligned} c^* &= (1-\theta^*)z(1-l^*); \quad (33) \\ \theta^* z(1-l^*) &= g. \quad (34) \end{aligned}$$

We can combine the last two equations to form:

$$c^* + g = z(1 - l^*) = y^*.$$

Hence, we have $\theta^* = g/y^*$. In terms of a diagram, the equilibrium is depicted as point B in Figure 6.4.

FIGURE 6.4
Expansionary Fiscal Policy
Financed with Income Tax



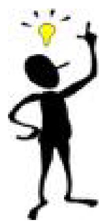
In Figure 6.4, an equilibrium is depicted for the case in which $g = 0$. The experiment is then to ask what happens when we increase g and finance the spending with an income tax. There are two effects at work here.

1. **Wealth Effect:** An increase in taxes reduces household wealth (as in the lump-sum tax). The demand for all normal goods falls. That is, private consumer demand falls and the demand for leisure falls (labor supply rises); and

2. **Substitution Effect:** An increase in the tax rate reduces the return on labor (makes leisure cheaper relative to consumption). The demand for consumption falls and the demand for leisure rises (labor supply falls).

We can predict that consumer demand declines unambiguously (as both the substitution and wealth effects operate in the same direction here). However, the response of the supply of labor is, in general, ambiguous (the substitution and wealth effects here operate in opposite directions). Point B in Figure 6.4 reflects the case in which the substitution effect dominates the wealth effect.

Observe that it is still true for this economy that $y \equiv c + g$. What has happened here, however, is that the increase in g is more than offset by the resulting decline in c. So while the income-expenditure identity holds true, this relation cannot be used to make predictive statements. This is because the income-expenditure identity is not a theory; it is simply an accounting relation.



Whether an exogenous increase in government purchases stimulates GDP and employment depends on how the added expenditure is financed.

Finally, note that the simple theory developed here provides one way to interpret the data in Figure 6.2. The basic idea is that countries with relatively large public sectors imply economies with relatively high distortionary tax rates; and that these high tax rates provide disincentives for market activity.

3.3 Government Transfers

We know that government transfers are not counted as part of the GDP. But does this mean that a government transfer policy has no effect on GDP? The answer is **no**.

To see this, consider the model described above. Let us consider as a benchmark the allocation (c^*, l^*), which would prevail in the absence of any government. Now, consider the following fiscal policy. The government makes no purchases (so that $g = 0$), but instead chooses to make a **lump-sum** transfer of resources $a > 0$ to the representative household. The government chooses to finance this subsidy with an income-tax.

In this case, the household's budget constraint is given by:

$$c = (1 - \theta)z(1 - l) + a.$$

Notice that the slope of the budget line is once again equal to $-(1-\theta)z$. Therefore, the following two conditions describe household behavior:

$$\begin{aligned} MRS(c^D, l^D) &= (1 - \theta)z; \\ c^D &= (1 - \theta)z(1 - l^D) + a. \end{aligned}$$

Since $g = 0$, the government's budget constraint in this case is given by:

$$\theta z(1 - l^D) = a.$$

Let me denote the equilibrium of this economy by (c^0, l^0, θ^0). For a given transfer policy a, this equilibrium allocation and tax rate can be deduced by combining the three equations above; i.e.,

$$MRS(c^0, l^0) = (1 - \theta^0)z;$$

and

$$\begin{aligned} c^0 &= (1 - \theta^0)z(1 - l^0) + a; \\ \theta^0 z(1 - l^0) &= a. \end{aligned}$$

Combining the last two equations, we derive:

$$c^0 = z(1 - l^0). \tag{35}$$

Hence, we can say the following three things about the equilibrium allocation in this case:

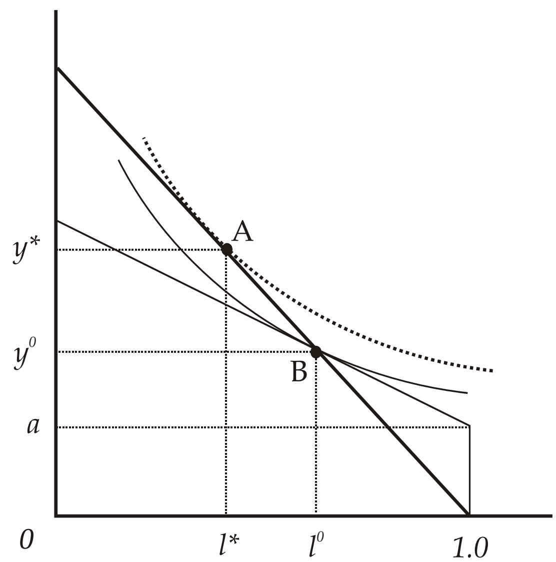
1. The allocation (c^0, l^0) lies on the PPF (see equation 35); and
2. The allocation (c^0, l^0) lies on the budget line, tangent to the indifference curve; and
3. The slope of the indifference curve at (c^0, l^0) is "flatter" than the slope of the PPF.

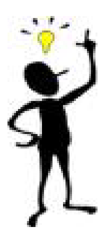
The three facts above imply that $l^0 > l^*$ (so that $n^0 < n^*$). In terms of a diagram, the equilibrium allocation can be depicted by a point like B in Figure 6.5.

Exercise 6.4: Prove that the equilibrium allocation under this transfer policy must lie on a point like B that is to the right of point A in Figure 6.5 (that is, demonstrate that the allocation cannot lie to the left of point A).

Once again, you can verify that the income-expenditure identity holds in this model; i.e., $y^0 = c^0$. The effect of this transfer policy (a) is in some ways similar to the effect of a purchase policy (g). In particular, both policies require an increase in the tax rate; and this tax rate lowers the return to labor, leading to a substitution away from market activity. The only difference between these two policies is that the transfer policy induces a positive wealth effect (the transfer adds to household income). But the result of this added wealth effect is to further reduce the demand for leisure. In other words, as far as the supply of labor is concerned, the negative wealth effect stemming from the reduction in after-tax wages is approximately offset by the positive wealth effect induced by the transfer. In this case, the substitution effect dominates, so that employment and output declines. This is yet another way to interpret the data in Figure 6.2.

FIGURE 6.5
Lump-Sum Transfer Financed
with an Income Tax





The fact that government spending on transfers is not counted as a part of GDP does not imply that a government transfer policy has no effect on GDP!

In the context of our example above, which features a representative household, we see that such a government policy also leads to a reduction in welfare. However, in a more realistic model that featured different types of households (e.g., high-skilled versus low-skilled), we would find that some households would benefit while others would lose. The model's predictions concerning the effect on GDP and aggregate employment, however, would remain intact.

4. Government Deficits

The analysis thus far has abstracted from government saving and borrowing. To study the effects of government budget surpluses and deficits, we have to consider a dynamic model. For this purpose, we can draw on the theory developed in Chapter 4.

As with households, the government faces a budget constraint. We saw in Chapter 4 that if households are free to save and borrow, then they are no longer constrained to live within their means on a period-by-period basis; instead, they must live within their means on a lifetime basis. The same general principle applies to the government sector.

4.1 Government Intertemporal Budget Constraint

Consider then a government with an exogenous spending flow (g_1, g_2). The government's income flow is given by the sum of net tax revenues that it collects over time (τ_1, τ_2). Government saving in the current period is therefore given by:

$$s_G \equiv \tau_1 - g_1. \tag{36}$$

In the static model developed in the previous section, no saving was possible, so that $s_G = 0$ (and $\tau_1 = g_1$). But assume now that the government, like households, can save or borrow at an exogenous interest rate $R \equiv (1 + r)$. In this case, future government spending is constrained to obey:

$$g_2 \leq \tau_2 + R s_G. \tag{37}$$

If $s_G < 0$, then the government is borrowing (running a deficit). The implication of this is that the government will at some point in the future have to run a surplus; i.e., $g_2 < \tau_2$. This future surplus is necessary to pay off the principal and interest on its maturing debt, Rs_G. The converse holds true if instead the government is saving (running a surplus).

Assume that condition (37) holds with equality. Then, by combining equations (36) and (37), we can derive the **government's intertemporal budget constraint** (or GBC, for short):

$$g_1 + \frac{g_2}{R} = \tau_1 + \frac{\tau_2}{R}. \tag{38}$$

The left-hand-side of this equation is the present value of the government's expenditure program. The right-hand-side of this equation is the present value of the government's net tax revenue. Clearly, this constraint allows for the possibility that $g_1 \neq \tau_1$ and $g_2 \neq \tau_2$. However, the constraint implies that if $g_1 > \tau_1$, then $g_2 < \tau_2$ (and vice-versa).

Thus, for a **given** expenditure program (g_1, g_2), we see that the fiscal authority is more or less free to determine the **timing of taxes** (τ_1, τ_2) That is, a finance minister may choose to lower taxes today (thereby increasing the deficit), but without any planned changes in the government's expenditure program, such a policy must imply higher taxes at some point in the future (assuming that the government does not default on its debt obligations).

4.2 The Ricardian Equivalence Theorem

Is a government budget deficit any cause for concern? The answer to this question is not as obvious as one might think.

To understand why, let us consider a simple dynamic model, with a representative household that has an exogenous earnings flow (y_1, y_2) and faces a sequence of tax obligations given by (τ_1, τ_2). In this case, private sector saving is given by:

$$s_P \equiv y_1 - \tau_1 - c_1,$$

where $(y_1 - \tau_1)$ denotes the current period **disposable** (after-tax) income. In the future period, private expenditures must obey:

$$c_2 = y_2 - \tau_2 + R s_P.$$

Combining these two equations, we can derive the household sector's intertemporal budget constraint:

$$c_1 + \frac{c_2}{R} = (y_1 - \tau_1) + \frac{(y_2 - \tau_2)}{R}.$$

The only difference here, relative to Chapter 4, is that the right-hand-side now represents **disposable** (after-tax) wealth. It will be convenient to rewrite the household's intertemporal budget constraint in the following manner:

$$c_1 + \frac{c_2}{R} = \left[y_1 + \frac{y_2}{R}\right] - \left[\tau_1 + \frac{\tau_2}{R}\right].$$

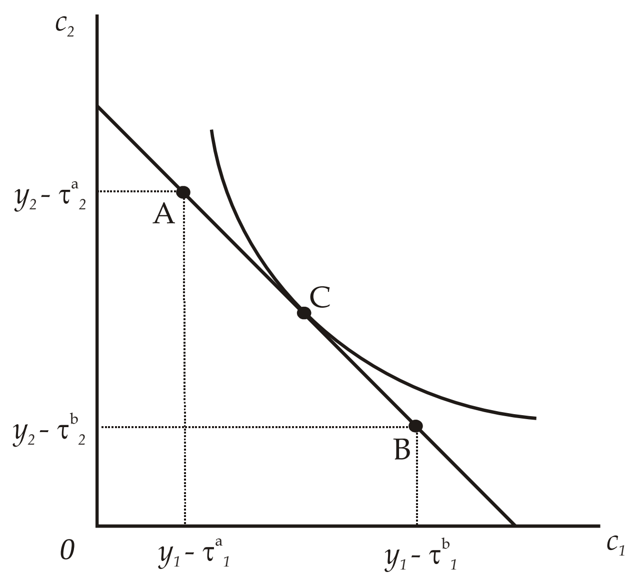
Now, consider a **given** government expenditure policy (g_1, g_2) and consider two possible ways in which this program can be financed; i.e.,

$$g_1 + \frac{g_2}{R} = \tau_1^a + \frac{\tau_2^a}{R} = \tau_1^b + \frac{\tau_2^b}{R};$$

where $\tau_1^a > \tau_1^b$. That is, under program a, the government taxes a lot today (and taxes little in the future); while under program b, the government taxes little today (and taxes a lot in the future).

The important point to note here is that while both programs have very **different** implications for the **timing of taxes**, both programs entail exactly the **same** present value tax obligation as far as the representative household is concerned. It therefore follows that the timing of taxes in no way affects the household's after-tax wealth position. Since after-tax wealth remains unchanged, either program in no way affects consumer demand (or welfare); see Figure 6.6.

FIGURE 6.6
A Change in the Timing of Taxes



Let us examine Figure 6.6 and assume that we are initially at point A (with the household choosing point C). Now, imagine that the government decides to cut taxes today, without changing its program spending. You can think of this program change as a **deficit-financed tax cut**. Then the government budget constraint implies that future taxes must rise. In short, a change in the finance program from policy a to policy b moves the household's after-tax disposable **income** flow from point A to point B. But as the policy change leaves disposable **wealth** unchanged, it has no effect on consumer demand. In short, it appears that the resulting government budget deficit **does not matter**.

While this result may appear counterintuitive, it follows from the logic embedded in our model. The intuition is as follows. The government is offering to cut taxes today. Such a policy might be motivated by the desire on the part of the government to stimulate consumer spending. The basic idea (conventional wisdom) is that an increase in disposable income should stimulate consumer spending. However, our model agents are forward-looking and they understand the government's budget constraint. They deduce that a cut in taxes today (without any change in program spending) must imply a higher tax burden in the future. Therefore, a standard consumption-smoothing argument implies that households will be motivated to increase their personal saving s_P^D for two reasons:

1. s_P^D will rise as households save a part of the increase in their current disposable income; and

2. s_P^D will rise as households save to offset the decrease in their future disposable income.

As it turns out, households end up **saving** the **entire tax cut** (as c_1^D remains unchanged). They are motivated to save their tax cut in order to help pay off what they know will be a higher future tax burden (taxes must rise in the future to pay off both the principal and interest on the maturing government debt). The rise in private sector saving matches–dollar for dollar–the decline in public sector saving, leaving desired domestic saving unchanged.

The irrelevance of government budget deficits (or surpluses) is sometimes known as the **Ricardian Equivalence Theorem**. The theorem states that (under certain conditions, to be discussed further below), the timing of taxes does not matter in the sense that different timing structures have no effect on consumer demand, domestic saving, or economic welfare (the timing of taxes does effect the composition of domestic saving). In short, deficits and taxes are **equivalent** ways to finance a given government expenditure program.

The Ricardian Equivalence Theorem is a proposition that holds in environments that are much more complicated than the one considered here (for example, including uncertainty does not alter the result). Nevertheless, the proposition does rest on some key assumptions, which I now make explicit.

[A1] Households are not debt-constrained.

You might want to refer briefly to Appendix 5.1 on borrowing constraints. Consider the finance policy a and imagine that households cannot borrow. Then the indifference curve in Figure 6.6 will pass through point A. At this point, households would like to borrow, but cannot. A deficit-financed tax cut (moving the after-tax income flow from point A to point B) in this case will stimulate consumer spending. In other words, households will not increase their saving by the full amount of the tax cut. In effect, the government is now borrowing on behalf of households who could not. It is doubtful, however, whether this is a serious problem for the theorem, as modern financial markets appear more than willing to extend credit to households.

[A2] Taxes are not distortionary.

In the model described above, taxes have impose no distortionary incentives. If we extended the model, as in the previous section, and assumed that governments must collect taxes via distortionary instruments, then the result of the theorem will no longer hold. Compare, for example, the following two policies: a 100% income tax today (and a 0% income tax tomorrow) versus a 0% income tax today (and a 100% income tax tomorrow). The timing of taxes will clearly matter here, because of the distortions they impose. This is potentially a big problem for the theorem, as taxes are distortionary in reality.

[A3] Households are forward-looking.

In the model developed above, households are forward-looking and they understand the government budget constraint. Neither of these assumptions appear to do great violation to reality. The fact that households save clearly indicates that they must be forward-looking. The fact that much political debate surrounds government budget deficits clearly suggests that the population generally understands the government budget constraint.

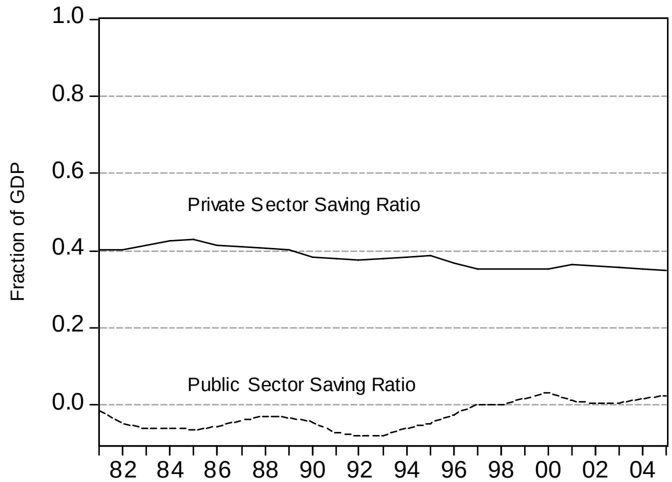
[A4] Households must live as long as the government.

In the model developed above, both households and the government "live" for two periods (the entire duration of the economy). In reality, people die, while the government stays in place for longer periods of time. Cutting taxes today for people who will not be around to pay higher taxes in the future implies that the tax-cut will increase the disposable wealth of short-lived people (and hence, affect their consumer spending). The government is, in effect, shifting the tax burden from a current generation to a future generation (who will suffer the consequences). On the other hand, while people do not live forever, it is not so unreasonable to assume that households do. If parents care for their children, then they will save the tax-cut and pass on the savings to their children (who will then use the bequest to pay off the higher future taxes).

Given the assumptions that underlie the Ricardian Equivalence Theorem, one may well wonder of what use it is. You may be surprised to learn however, that economists who try to reject the theorems predictions frequently find it hard to do so conclusively. Even a casual look at the data suggests that the theorem's predictions are not all that crazy; see, for example, Figure 6.7.

In reality, of course, the theorem cannot hold exactly (all of the assumptions it makes are literally violated in reality). Nevertheless, the theorem suggests an important caveat for policy makers contemplating a deficit-financed tax cut (a policy frequently considered during periods of recession). The basic lesson is that, while it may turn out to be the case that cutting taxes will stimulate consumer spending by increasing disposable income, the **quantitative** impact on consumer spending is not likely to be as strong as one might think (if, for example, one was inclined to make predictions based on conventional wisdom).

FIGURE 6.7
Private and Public Sector Saving Ratios
Canada 1981 - 2005



Private sector saving ratio = (GDP - Taxes + Transfers - Consumption)/GDP
Public sector saving ratio = (Taxes - Transfers - Purchases)/GDP

One thing to bear in mind here is that the Ricardian theorem is a statement about the effects of financing a **given** government expenditure program in different ways. The theorem asserts that for a **given** expenditure program, how the government goes about financing it really does not matter that much. The theorem does **not**, however, suggest that government expenditure programs do not matter. From the household's intertemporal budget constraint above, it should be clear that any increase in government program spending (g_1, g_2) is going to reduce the disposable wealth of the household sector; and therefore affect consumer spending.

> *The Ricardian Equivalence Theorem asserts the irrelevance of government budget deficits;* **not** *the irrelevance of government spending policy.*

4.3 Financing a Transitory Increase in Government Purchases

A classic example of a "transitory" government spending shock is when a nation goes to war. During a war, a government must temporarily increase its claim on the economy's output–diverting these resources for the purpose of pursuing the war effort. During a major war, like World War 2, the impact of an increase in government purchases appears to be an increase in employment and GDP, and a decrease in private sector consumption (severe wars are frequently associated with the rationing of many consumer goods).

In a world where the Ricardian Equivalence Theorem holds, the manner in which a transitory increase in government purchases is financed will have no major consequences. Of course, the increase in government purchases themselves will have some major consequences, whether or not the theorem holds. We can already deduce that because household wealth will decline, the demand for all normal goods (c_1, l_1, c_2, l_2) will all decline as well. The effect of a **transitory** increase in government purchases will therefore result in a **persistent** increase in employment and GDP (that is, output and employment will remain higher even after the war ends). In what follows, I assume that this wealth effect is relatively weak.

Of course, the persistence generated by the negative wealth effect is only one force that might operate on the economy (it would be the only force, if Ricardian Equivalence holds). If taxes are distortionary, there may be other consequences as well. Let's see whether we can disentangle the possible effects.

Consider a representative household with preferences $u(c_1, l_1, c_2, l_2)$. Let $(w_1, w_2) = (z_1, z_2)$ and $(d_1, d_2) = (0, 0)$; i.e., if the production technology is linear in labor, we already know that this will describe the equilibrium gross wages and non-labor income.

Next, consider a government spending program $(g_1, g_2) = (g, 0)$. Here, I let $g > 0$ denote the transitory increase in government purchases. The government's intertemporal budget constraint is given by:

$$g = \theta_1 z_1(1 - l_1) + \frac{\theta_2 z_2(1 - l_2)}{R}.$$

That is, even though the spending g is transitory, the government has many different ways to collect the resources it needs. Let me consider two extreme cases:

[P1] The government finances g entirely out of current taxes (and hence, runs no deficit). This implies $\theta_1 > 0$ and $\theta_2 = 0$.

[P2] The government finances g entirely by issuing debt. This implies $\theta_1 = 0$ and $\theta_2 > 0$.

Finally, let us write down the household's intertemporal budget constraint:

$$c_1 + \frac{c_2}{R} = (1 - \theta_1)z_1(1 - l_1) + \frac{(1 - \theta_2)z_2(1 - l_2)}{R}.$$

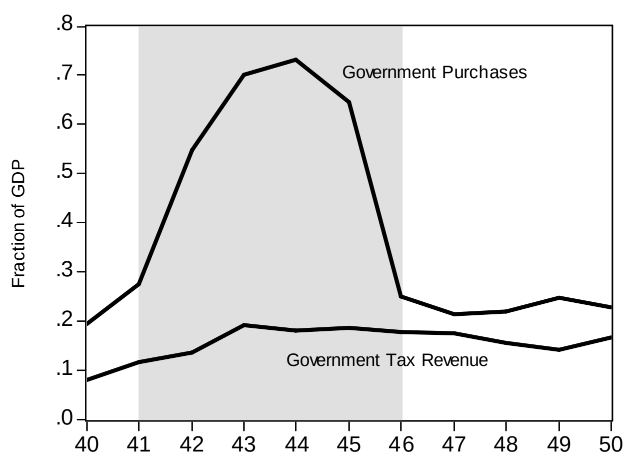
Since taxes are distortionary here, the Ricardian theorem will fail to hold; the two finance policies **[P1]** and **[P2]** are likely to have very different effects. In particular, they will induce different **intertemporal substitution effects**.

[E1] Under finance policy **[P1]**, the current tax on labor is high and the future tax on labor is low. It will make sense here for households to reduce current labor supply and increase future labor supply (i.e., substitute leisure across time). Hence, output and employment are likely to decline during the war, and expand when the war ends.

[E2] Under finance policy **[P2]**, the current tax on labor is low and the future tax on labor is high. It will make sense here for households to increase current labor supply and decrease future labor supply (i.e., substitute leisure across time, but this time in the opposite direction). Hence, output and employment are likely to expand during the war, and contract when the war ends.

Thus, we see here that our theory is consistent with the conventional wisdom that, in the **short-run**, a bond-financed increase in government purchases is likely to be more expansionary than a tax-financed increase in government purchases. The reason for why this is so in our model, however, differs from the conventional wisdom. That is, the conventional wisdom suggests that the expansion will occur because lower taxes today increases disposable income and therefore stimulates consumer demand. In contrast, our theory suggests that the expansion will occur because the timing of taxes induces households to substitute leisure today for leisure in the future.

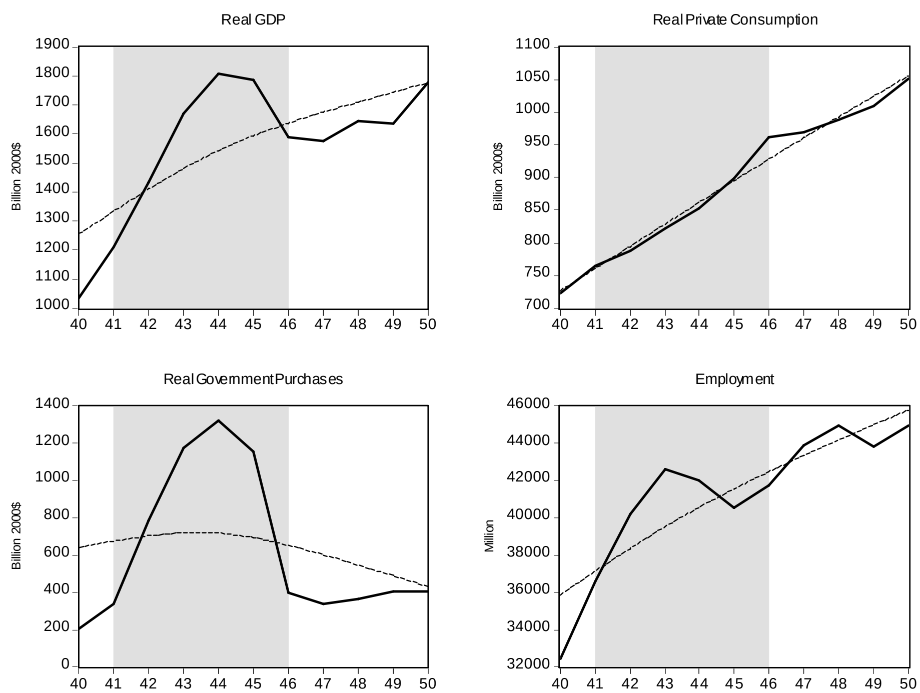
FIGURE 6.8
Government Purchases and Taxes Relative to GDP
U.S. 1940 -1950



Let's take a look at some data. We know that the United States entered World War 2 in December of 1941 (and was building up military hardware in anticipation of this event). Figure 6.8 reveals an enormous expansion in government purchases during the war years (1941-1946). The same figure also reveals that this expansion was financed primarily by issuing government debt (war bonds).

This episode does not quite fit our story as we see that the tax rate did not rise following the end of the war (the debt-to-GDP ratio fell in subsequent years, but this owed more to economic growth, rather than an increase in the tax rate).[51] Nevertheless, let's take a look at how this "fiscal shock" impacted the U.S. economy; see Figure 6.9.

FIGURE 6.9
Output and Employment
U.S. 1940 - 1950



In Figure 6.9, the "dashed" lines measure a rough statistical trend (to correct for productivity and population growth) for each of the series and the shaded

[51]The rise in the tax rate during the war years is also understated here, as there was a significant inflation-tax over this period. We will address monetary phenomena in later chapters.

regions correspond to the war years. The trend for government purchases peaks during the war years, and subsequently shows a downward trend. Given that the trend in GDP is rising throughout the sample, the trend here is a long-run decline in the size of government purchases relative to GDP (and perhaps we can take this as a general expectation of a decline in the future tax burden).

The data clearly shows a transitory increase in GDP and employment (relative to trend) during the war years; behavior that is generally consistent with our theory. There is also a slight decline in private sector consumption (relative to trend), but consumption is much smoother than the time-path of income. This consumption-smoothing behavior is also consistent with our theory.

4.4 Barro's Tax-Smoothing Argument

Suppose that the government's expenditure program (g_1, g_2) is fixed in place. When taxes are lump-sum, the government's finance department faces a trivial decision: choose any (τ_1, τ_2) that satisfies the government's intertemporal budget constraint. However, when taxes are distortionary, Robert Barro has pointed out that it would be **optimal** for the government to **smooth taxes** over time.[52] That is, the government should choose a tax rate policy that balances not only the government's intertemporal budget constraint, but balances government spending and revenue on **average** throughout time. This implies a relatively stable tax rate and a budget deficit/surplus that fluctuates over time (but balance out over the long-run) along with any transitory fluctuations in government purchases. By smoothing taxes in this manner, the government is in effect smoothing out (and therefore minimizing) the distortions that its taxes create over time.

Since taxes are distortionary, the timing of taxes does matter. If government expenditures fluctuate over time, an optimal tax policy entails a relatively stable tax rate, with surpluses and deficits to absorb transitory fluctuations in expenditure.

[52] Barro, Robert J. (1979). "On the Determination of the Public Debt," *Journal of Political Economy*, 64: 93–110.

5. Summary

Government fiscal policy can have important effects on the macroeconomy. Government purchases and transfers require the acquisition of resources from the private sector. One way or another, these resources are acquired by taxing agents in the private sector. While a government may borrow resources to finance expenditures in the short-run, such borrowing merely entails a postponement of taxes into the future.

Taxes have two important effects on the private sector. First, a net tax burden reduces disposable wealth in the private sector. Second, even if net taxes are low (say, because most tax revenue is simply transferred), taxes typically distort private sector incentives.

To the extent that Ricardian equivalence serves as a good approximation for reality, the timing of taxes is not likely to have a significant impact on consumer demand by influencing the timing of disposable income; private sector saving can largely be expected to offset any change in public sector saving. On the other hand, as taxes are largely distortionary, the timing of taxes can be expected to affect the pattern of economic activity over time. There is a good theoretical rationale for keeping tax rates roughly constant (sufficient to balance the government budget on average) if the demand for government purchases varies over time. Such a policy would entail running budget surpluses when government spending is temporarily below average, and budget deficits when government spending is temporarily above average.

Problems

1. Consider an economy populated by a representative household with preferences $u(c,l) = 2c^{1/2} + 2l^{1/2}$. For these preferences, $MRS(c,l) = (c/l)^{1/2}$. Assume that $(w^*, d^*) = (z, 0)$ and that the income tax rate is set exogenously at some number $0 \leq \theta \leq 1$. Assume that taxes are used to finance government purchases $g^* = \theta z n^*$. Demonstrate that the equilibrium level of employment in this economy is given by:

$$n^*(\theta) = \left[\frac{(1-\theta)z}{1 + (1-\theta)z} \right].$$

   How does $n^*(\theta)$ depend on θ? Explain.

2. In the question above, a given tax rate θ generates government revenue (and expenditure) $g^* = \theta z n^*(\theta)$. Notice that total revenue is the product of the **tax rate** θ and the **tax base** $zn^*(\theta)$. Thus, any change in θ will affect revenues directly through the tax rate and indirectly through the tax base. Demonstrate that in this economy, there is a **Laffer curve** that places a limit on how much revenue the government can extract from the economy. [See: http://en.wikipedia.org/wiki/Laffer_curve].

3. You work hard during the year to produce one apple. This contributes to the annual GDP by the value of the apple. Suppose that the government takes this apple away from you and redistributes it to someone else. Explain why this government policy will not affect this year's GDP (given that you have already produced the apple). On the other hand, explain why such a policy, if it is to remain in place, is likely to reduce future GDP by one apple.

4. A transfer policy frequently advocated by economists is the so-called **Negative Income Tax** (NIT); see: http:// en.wikipedia.org/wiki/ Negative_income_tax. In the context of our model, a NIT would entail a lump-sum transfer (a) to all people, financed by a flat income tax (θ). Now, imagine that people in the economy differ in skill (i.e., so that they have different real wage rates). Explain why a NIT is likely to reduce the labor supply among low-skill people; while high-skill people may even end up working harder. Hint: consider the wealth and substitution effects of a NIT and how they are like to differ across skill groups.

5. Explain why the Ricardian Equivalence Theorem is unlikely to hold in an economy that experiences net immigration or emigration flows.

6. Consider a world in which the Ricardian Equivalence Theorem holds. Imagine now that the government announces that it plans to increases its purchases in the future (i.e., $\Delta g_1 = 0$ and $\Delta g_2 > 0$). Explain why this announcement is likely to depress private consumer spending **today** (and

in the future). Explain too why the method of financing this added expenditure does not matter. That is, **either** finance policy ($\Delta\tau_1 = \Delta g_2/R$, $\Delta\tau_2 = 0$) **or** ($\Delta\tau_1 = 0$, $\Delta\tau_2 = \Delta g_2$) will have exactly the same effect.

7. You are employed as a policy advisor in the Department of Finance. Recent data shows that the economy is weakening and quite possibly in recession. The finance minister has a plan that he/she thinks will help "stimulate" the economy. The plan involves a significant reduction in the income tax rate (and a corresponding increase in the government budget deficit). The basic idea, you are told, is to increase the disposable income for households, thereby stimulating consumer demand. You are asked to provide an assessment and critique of this plan. (Do not use math or diagrams).

CHAPTER 7

Capital and Investment

1. Introduction

To this point, we have limited our analysis to economies that produce output in the form of consumption. Consumer goods and services can be loosely defined as output that cannot be stored across time; or output that in no way affects future production possibilities. While a good deal of what an economy produces may be broadly classified as consumption, not all of production takes this form. An important branch of the production flow instead takes the form of goods and services that can be stored over time, or in some way influence future production possibilities. Economists call this type of output **investment** (or new capital goods and services).

Some examples of investment include: the construction of a new residential building, the manufacture of machinery and equipment, additions to inventory, and research and development. Resources allocated to investment activities constitute an economy-wide act of saving (a sacrifice of current consumption). That is, the construction worker, factory worker, and scientist could all potentially be employed instead as taxi drivers, farmers, or accountants. More factors employed in the investment sector imply less factors employed in the consumption sector. But if people value consumption, then why allocated to activities that yield no immediate benefit?

The answer, of course, is that while investment has no immediate benefit, it may have a future benefit. The new home, the new piece of equipment, and the new idea might all serve to increase the economy's ability to produce future consumption. In short, investment is a flow of new capital goods and services that augments the economy's future stock of productive capital.

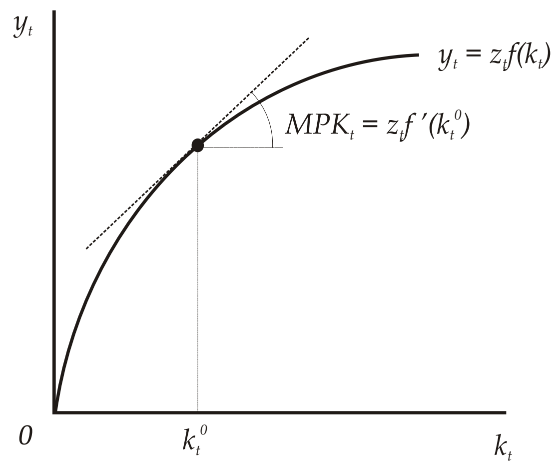
2. The Demand for Investment

Consider a simple two-period economy. Assume that employment is fixed at some level. Production at date $t = 1, 2$ is determined by:

$$y_t = z_t f(k_t);$$

where k_t denotes the stock of capital available for production at date t; and where z_t is a productivity parameter. More capital means that more can be produced; hence, we assume $f'(k) > 0$. Since labor is fixed, it also seems reasonable to suppose that there are diminishing returns to increasing the capital stock; i.e., $f''(k) < 0$. In other words, assume that f is an increasing and strictly concave function; see Figure 7.1.

FIGURE 7.1
Production Technology



The slope of this production function represents the **marginal product of capital**; or $MPK_t \equiv z_t f'(k_t)$. The MPK represents the additional output that can be produced by increasing the capital input by one (small) unit. In other words, it represents the **return** (measured in units of additional output) that you would get by expanding the capital stock a little bit. Notice that this return is diminishing in the size of the capital stock (the slope gets flatter as you increase k_t).

Assume that the capital used in production at date t depreciates in value at some exogenous rate $0 \leq \delta \leq 1$. That is, δk_t represents the value of capital that is "consumed" in the act of production. Capital consumption (depreciation) at date t will reduce the amount of capital available at date $t+1$. But the stock of future capital can also be replenished by investing more of today's output toward the production (or purchase) of new capital goods. Let x_t denote domestic investment at date t. In this case, the domestic stock of capital evolves in the following way:

$$k_{t+1} = k_t - \delta k_t + x_t.$$

Let me simplify here by assuming $\delta = 1$ (100% depreciation rate), so that $k_{t+1} = x_t$. That is, investment today corresponds exactly to the amount of capital available tomorrow for production. The question we want to answer is: what determines the level of domestic investment demand? Let us try to answer this question in the context of a simple two-period economy.

Consider a representative firm that can borrow resources at a given real (gross) rate of interest R. If the firm borrows x units of output for the purpose of domestic investment, it will then owe to its creditors Rx units of output in the future (this represents both the principal and net interest owed on its debt). The purchase and installation of x units of new capital goods today implies a stock of productive capital tomorrow equal to x (recall that capital is assumed to depreciate fully). This future capital stock produces $y_2 = z_2 f(x)$ units of future output. Hence, the investment x yields a future profit:

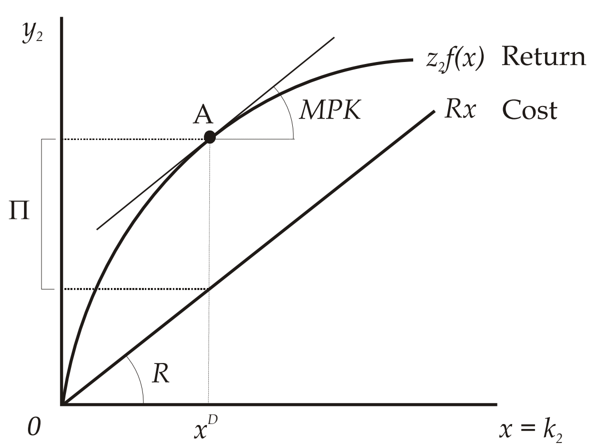
$$\Pi(x) \equiv -Rx + z_2 f(x).$$

The choice problem of our representative firm is to choose the x that maximizes $\Pi(x)$. In making this choice, the firm must consider the prevailing interest rate R; and it must make a forecast of future productivity z_2. Given (R, z_2), the demand for investment x^D is determined by the condition:

$$z_2 f'(x^D) = R; \tag{39}$$

see Figure 7.2.

FIGURE 7.2
Domestic Investment Demand



Condition (39) has a very natural economic interpretation. The left-hand-side represents the expected marginal product of future capital; this is, it represents the expected marginal return to current investment spending. The right-hand-side represents the marginal cost of investment. The optimal investment level x^D equates the marginal benefit and marginal cost of investment.

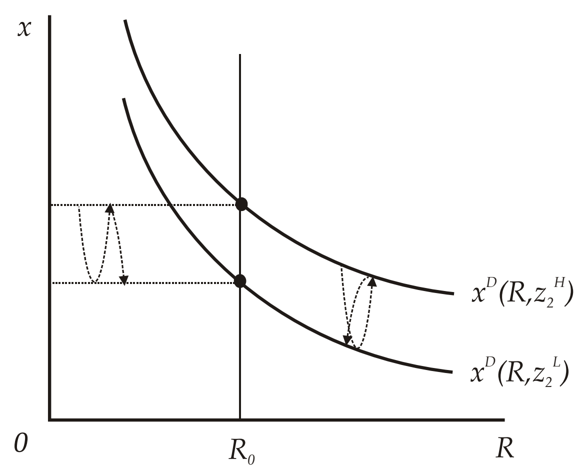
Notice that the demand for domestic investment depends on both the "cost of capital" R and the expected productivity of future capital z_2. This dependence can be stressed by writing the investment demand function as $x^D(R, z_2)$. Not surprisingly, our theory predicts that desired domestic investment spending depends negatively on the interest rate R and positively on expected productivity z_2.

Exercise 7.1: Using Figure 7.2, show how an exogenous increase in the interest rate leads to a reduction in investment, profits, and future GDP.

Exercise 7.2: Let $MPK = \alpha z_2 x^{\alpha-1}$ where $0 < \alpha < 1$. Solve for the investment demand function $x^D(R, z_2)$ and show that it is increasing in z_2.

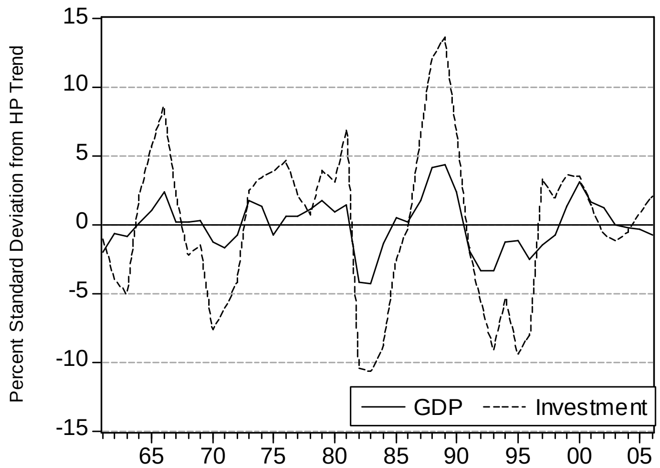
Your answers to the two exercises above can be summarized with the aid of a diagram; see Figure 7.3.

FIGURE 7.3
Investment Demand Function



Empirically, investment spending is the most volatile expenditure component of GDP; see Figure 7.4. This fact is consistent with even casual empiricism. For example, it is well-known that employment in the construction sector is very volatile; say, compared to employment in the service sector. Why investment demand fluctuates is a matter of debate. As in our discussion of what might cause the demand for labor to fluctuate (Chapters 2 and 3), one might take two broad views. The neoclassical view contends that expectations of future productivity fluctuate with the exogenous arrival of information that leads private sector actors to revise their expectations. The conventional view contends that expectations fluctuate exogenously owing to the psychology of the market (animal spirits). The reality probably lies somewhere between.

FIGURE 7.4
GDP and Investment
Canada 1961 - 2006



Source: CANSIM II. GDP v3860085; Investment is the sum of Business Gross Capital Formation (v3860070) and Consumer Durable Spending (v3860063).

3. A Small Open Economy with Saving and Investment

In Chapter 5, we modeled a small open economy with saving, but no investment. You can think of that chapter as describing a theory of domestic saving. In the previous section, we developed a theory of domestic investment. In this section, I bring these two theories together.

It will be helpful to quickly review the material presented in Chapter 5. There, we considered a two-period model populated by a representative household with preferences for time-dated consumption, $u(c_1, c_2)$. We considered the household's choice problem, assuming an exogenous stream of GDP (y_1, y_2) and a given interest rate R. The solution to this choice problem can be summarized by a desired level of domestic saving $s^D(R, y_1, y_2)$. Remember that this saving function is just the flip side of the consumer demand function $c_1^D(R, y_1, y_2)$.

Exercise 7.3: Show that desired saving s^D is an increasing function of R (assuming SE > WE); an increasing function of y_1; and a decreasing function of y_2. Explain.

The effect of introducing investment into that simple model is to endogenize y_2. That is, since $y_2 = z_2 f(x)$, it now follows that the future GDP will be a function of current period investment. We can still treat y_1 as exogenous here since $y_1 = z_1 f(k_1)$, where $k_1 > 0$ is a predetermined level of "old" capital that is already in place.

Now, for a given (R, z_2), the demand for investment is given by $x^D(R, z_2)$; and this generates a future GDP $y_2^S(R, z_2) \equiv z_2 f(x^D(R, z_2))$. Clearly, the future GDP is decreasing in R (why?) and increasing in z_2 (why?).

In an open economy, there are two uses of domestic saving s^D. Savings can either be used to finance domestic investment spending x^D; or they can be used to purchase foreign bonds b^D (which earn a rate of return equal to R). Note that b^D may be either positive or negative. If $b^D > 0$, then domestic residents are net purchasers of bonds (the trade balance is positive). If $b^D < 0$, then domestic residents are net sellers of bonds. (the trade balance is negative). Hence, we have the relation:

$$b^D \equiv s^D(R, y_1, y_2) - x^D(R, z_2); \tag{40}$$

where $b^D \lesseqgtr 0$ denotes the value of the trade balance (net exports).

Let y_2^* denote the level of future GDP expected to prevail in the future. I am going to assume that households have rational expectations, in the sense that:

$$y_2^* = y^S(R, z_2). \tag{41}$$

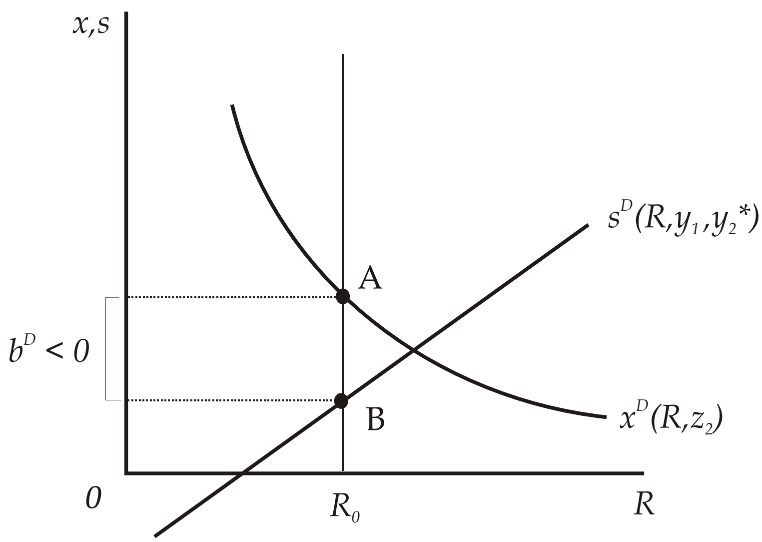
Hence, for a given interest rate R and a given expected productivity z_2, the equilibrium trade balance will be determined by combining (41) and (40); see Figure 7.5.

Figure 7.5 depicts the equilibrium trade balance for some arbitrary interest rate R_0. Point A determines the level of domestic investment spending and point B determines the level of domestic saving. Since $s^D < x^D$, the level of domestic saving is not sufficient to finance the entire level of desired domestic investment spending. In this case, the gap $x^D - s^D$ must be financed with net imports of capital goods and services. These net imports are purchased by issuing bonds and selling them to foreigners (in exchange for their exports to domestic residents). As a consequence, the trade balance is in deficit; $b^D < 0$.

Exercise 7.4: Using Figure 7.4, explain how the economy's trade balance will react to an exogenous increase in the real interest rate.

Using Figure 7.4, explain how the trade balance is expected to react to a transitory decline in productivity (a transitory decline in y_1). Why does a transitory decline in productivity not affect domestic investment spending here?

FIGURE 7.4
Equilibrium Trade Balance



3.1 An Interpretation of Trade Balance Behavior

Enrique Mendoza documents the following facts concerning small open economies:[53]

[F1] The correlation between domestic saving and investment is positive;

[F2] Domestic saving does not fluctuate as much as investment;

[F3] The trade balance is countercyclical (tends to move in the opposite direction of GDP over the cycle).

Medoza also examines the predictions of a model that is very similar to the one developed in this chapter. In his paper, he assumes the existence of "persistent productivity shocks." In the context of our model, we can model a persistent productivity shock as $\Delta z_1 = \Delta z_2 > 0$. Let's see if we can work out the implications.

A persistent productivity shock will have two effects. First, it will have a direct positive impact on current GDP; i.e., $\Delta y_1 > 0$. This first effect will not influence investment spending (see Exercise 7.4), but will increase desired saving

[53] Mendoza, Enrique G. (1991). "Real Business Cycles in a Small Open Economy," *American Economic Review*, 81(4): 797–818.

(consumption smoothing motive). Second, it will have the effect of increasing the private sector's forecast of future productivity z_2. This second effect will lead to an increase in investment spending (see Exercise 7.2). The future GDP is therefore expected to increase for two reasons: the increase in future productivity and the increase in future capital. Households must therefore expect $\Delta y_2^* > 0$. But this expected increase in future income has exactly the opposite effect on desired saving than $\Delta y_1 > 0$. That is, the expectation of higher future income leads households to save less (or borrow more); again, this is just a straightforward consumption-smoothing argument.

If we consider all these effects together, then we can conclude that savings will not vary as much as investment (consistent with [**F2**]) and that the trade balance will tend to move in a countercyclical manner (consistent with [**F3**]). If, in addition, desired savings rise by a little bit, then our model's predictions are also consistent with [**F1**]. You can understand this in terms of Figure 7.4 by shifting the investment demand curve up by a lot and shifting the desired saving curve up by a little bit. A negative productivity shock will have the opposite effect.

4. Closed Economy Analysis

To this point, we have simply assumed that the real interest rate R was given to us by nature (exogenous). This assumption appears to be innocuous enough when we are dealing with a small open economy. It is unlikely, for example, that the domestic saving and investment choices made by Canadians have any significant impact on the interest rate that prevails in world financial markets.

World interest rates (I am talking here about real interest rates) do appear to move around. We know, for example, that the real interest rate appeared to be very low during the 1970s and very high during the 1980s. How are we to understand these types of movements?

One way to understand the determination of the interest rate is to note that while the saving and investment choices made in any small open economy may have no significant effect on R, this is unlikely to be the case when we consider the behavior of all small open economies together. Let us imagine a world made up of N small open economies. We can think of our earlier analysis as determining the saving and investment behavior of each small open economy or a given interest rate R and a given set of domestic conditions $\{z_{1i}, z_{2i}\}$ for $i = 1, 2, ..., N$. For economy i, we have:

$$b_i^D \equiv s^D(R, y_{1i}, y_{2i}^*) - x^D(R, z_{2i}).$$

Now, if we add up all these choices across countries, we derive:

$$\sum_{i=1}^{N} b_i^D \equiv \sum_{i=1}^{N} s^D(R, y_{1i}, y_{2i}^*) - \sum_{i=1}^{N} x^D(R, z_{2i}).$$

The expression $\sum_{i=1}^{N} b_i^D$ represents the **world's** trade balance position. Now, let's think about this for a second. What must the world trade balance position sum up to (assuming that there is no interplanetary trade)? Clearly, it must add up to **zero**. That is, for every country with a positive trade balance, there must be some other country in the world with a negative trade balance.

For an arbitrary interest rate R, the world trade balance is not likely to be equal to zero. For example, if we interpret s^D and x^D in Figure 7.4 as the **world** supply and demand for loanable funds, then the world would desire to import goods from other planets. The distance between point A and point B represents an **excess demand for loanable funds** in the world financial market. In such a circumstances, it seems plausible to suppose that a competition for loanable funds will put upward pressure on the real interest rate. One might expect such pressure to move the interest rate to a point R^* at which the world supply and demand for loanable funds is equated; i.e.,

$$\sum_{i=1}^{N} s^D(R^*, y_{1i}, y_{2i}^*) = \sum_{i=1}^{N} x^D(R^*, z_{2i}). \tag{42}$$

That is, in the context of Figure 7.4, the equilibrium interest rate would be determined by the intersection of the two curves s^D and x^D.

What this analysis suggests is that the real interest is likely to be influenced by how rapidly the world economy is expected to grow. We know, for example, that the 1970s was a decade characterized by a slowdown in world economic growth. In the context of our model, we might think of this as a general decline in z_{2i} across most (but not necessarily all) countries. The effect of such a "growth shock" would be to depress the world demand for investment and increase the world supply of saving. Both effects would serve to put downward pressure on the interest rate. Of course, the converse would be true in the event of a positive "growth shock."

5. The IS Curve

In the previous section, I described a theory that is distinctly "neoclassical" in flavor. For economists trained in the "conventional wisdom," such a theory can be thought of as, at best, describing how economies function in "the long-run" (whatever this means). In the "short-run," things appear to work very differently. Let me try to explain.

It is a neoclassical view that prices serve as "equilibrating variables." An example of what is meant by this phrase is to consider the discussion surrounding equation (42). The idea there was that if $s^D \neq x^D$, then the interest rate (a price) would move (by the magic of "market forces") to equilibrate the supply and demand for loanable funds. Likewise, in Chapter 2, the view there was that if $n^S \neq n^D$, then the real wage (a price) would move to equilibrate the supply and demand for labor.

It was an insight of Keynes (and touched upon by Malthus) that an economy's equilibrating variables may not be prices; rather they may be quantities. To give you an example of what is meant by this, consider a closed economy with:

$$s^D(R, y_1, y_2) = x^D(R, z_2). \tag{43}$$

The neoclassical view is to see this equality holding by an appropriate adjustment in R. But what if, instead of movements in R, the equilibrium between saving and investment occurs through an adjustment in y_1 (a quantity variable)?

According to this interpretation then, we should view R and z_2 as exogenous variables (y_2 is frequently viewed as exogenous as well, even though we have demonstrated above the dependence of y_2 on R and z_2). If this is the case, then equation (43) can only be brought to hold with equality by adjustments in y_1. How might this adjustment process be thought to occur?

One way to think of the adjustment process is as follows. In a closed economy (without government purchases), the aggregate demand for output today is given by:

$$AD \equiv c_1^D(R, y_1, y_2) + x^D(R, z_2).$$

This aggregate demand function is decreasing in R and increasing in y_1. Next, we have the aggregate supply of output given by:

$$AS \equiv y_1.$$

So far, all of this is consistent with the neoclassical model. The point of departure is in the next step.

The neoclassical model assumes that the AS is determined through price adjustment in the current period factor market; in this case, $y_1 = z_1 f(k_1)$. The equation $AS = AD$ then determines R^*; i.e.,

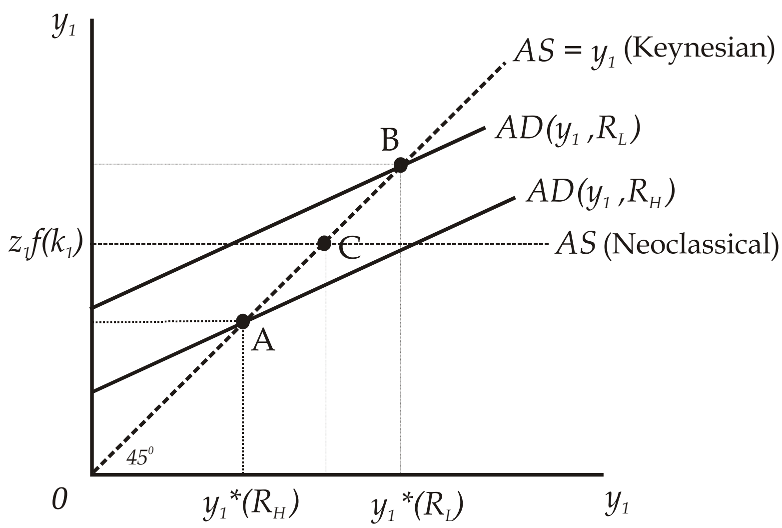
$$z_1 f(k_1) = c_1^D(R^*, y_1, y_2) + x^D(R^*, z_2).$$

In contrast, the "Keynesian" model assumes that the (short-run) AS is independent of price adjustments in the current period factor market (imagine, for example, that factor prices are "sticky" for some unexplained reason). In this case, the "Keynesian" model simply assumes that factors are supplied to meet the demands of producers; and that producers simply supply output to meet the aggregate demand for output; i.e.,

$$y_1^* = c_1^D(R, y_1^*, y_2) + x^D(R, z_2).$$

Figure 7.5 depicts these two different ways in viewing the AS relationships.

FIGURE 7.5
The Keynesian Cross

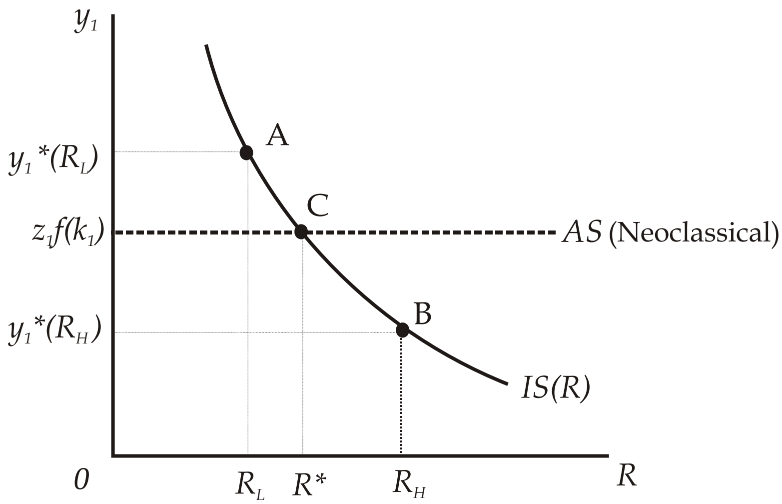


In Figure 7.5, point C depicts the neoclassical equilibrium. While I have not drawn it, you can imagine an AD curve passing through point C with interest rate R^*. Figure 7.5 also depicts two AD relationships associated with two interest rates $R_H > R^* > R_L$. That is, a high interest rate depresses the aggregate demand for output; while a low interest rate stimulates the demand for output. The supply of output is imagined to response passively to whatever the demand for output turns out to be. Point A depicts the Keynesian equilibrium for a high interest rate; and point B depicts the Keynesian equilibrium for a low interest rate. Notice that the equilibrium level of output is a decreasing function of the interest rate. This negative relationship $y_1^*(R)$ is called the **IS Curve**; see Figure 7.6.[54]

While there is reason to doubt whether this "Keynesian" analysis accurately reflects Keynes' own interpretation (he never derived an IS curve; this is something that Hicks invented much later), there can be little doubt that the IS curve reflects "conventional wisdom." The basic idea is that the level of GDP–in the "short-run" at least–is largely determined by the demand for output (with supply reacting passively to fulfil this demand). You can see clearly how this view has permeated everyday language. For example, both Statistics Canada and the Bank of Canada regularly refer to **measured GDP** as **aggregate demand**.

[54] IS stands for "Investment-Saving." Note that the IS curve can alternatively be derived from equation (43). That is, the expression AS = AD is equivalent to equation (43).

FIGURE 7.6
The IS Curve



In Figure 7.6, points A and B are said to reflect a "short-run disequilibrium." In an ideal world, prices (in this case, the real interest rate) would move the economy quickly to its "long-run" neoclassical equilibrium at point C. But since prices are not reliable equilibrating variables, this process may take a long time. In the meantime, the actual level of GDP (demand) may deviate from its long-run "potential" (neoclassical) level. The difference between actual GDP (demand) and potential GDP (long-run supply) is commonly referred to as the **output gap**. Since natural economic forces cannot be trusted to close the output gap quickly, there appears to be a role for government policy to help things work more smoothly.

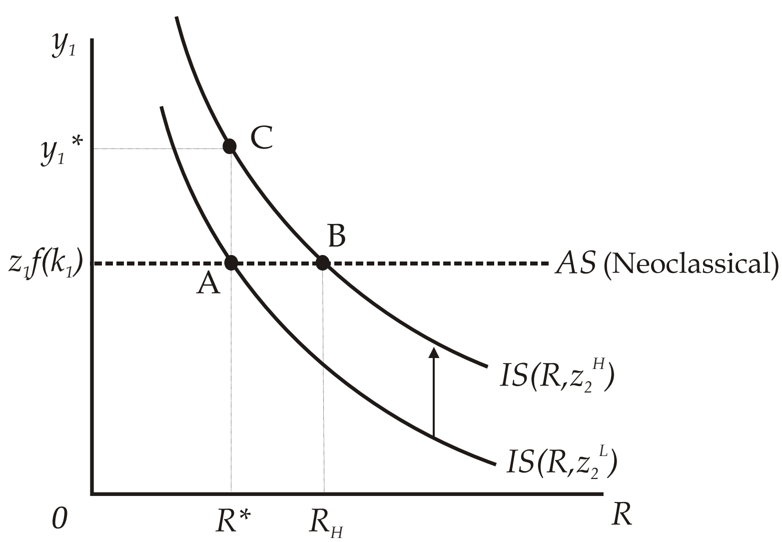
6. Policy Implications (Conventional Wisdom)

The conventional view is that the business cycle is ultimately caused by wild fluctuations in aggregate demand, stemming primarily from the demand for investment. These fluctuations are typically thought of as stemming from psychological factors (animal spirits) that lead to volatile movements in private sector expectations. But whether this is true or not need not concern us here. We may, in particular, assume that expectations change for "good" (fundamental) reasons; for example, an exogenous change in z_2 (the expected productivity of future capital).

In the neoclassical model, an exogenous increase in z_2 generates an increase in investment demand. In a closed economy, the effect of this increase in demand for output leads to an increase in the interest rate (leaving current GDP unchanged).[55] The higher interest rate suppresses current consumer demand (which is what allows more of the current GDP to be diverted toward the construction and purchase of new capital goods).

In the Keynesian model, an exogenous increase in z_2 also generates an increase in investment demand. However, the increase in aggregate demand does not result in (an immediate) increase in the interest rate. Instead, the increase in demand is simply met by an increase in current production. That is, this positive "demand shock" results in a positive "output gap." The economy appears to be "overheating;" see Figure 7.7.

FIGURE 7.7
Aggregate Demand Shock



In Figure 7.7, the initial "long-run" equilibrium is depicted by point A. The effect of the "aggregate demand shock" ($z_2^H > z_2^L$) is to shift the IS curve up. In the neoclassical model, the increase in demand results in an increase in the

[55] The neoclassical model can be extended, for example, by endogenizing capacity utilization. In this case, the increase in investment demand will lead firms to increase the supply of current GDP by utilizing capital more intensively. Greater utilization also implies an increase in the demand for labor and hence, and increase in the real wage.

real interest rate; so that the economy moves from point A to point B. In the Keynesian model, the increase in demand results in an expansion in current period GDP; so that the economy moves from point A to point C. It is natural to suppose that if the market cannot be relied upon to "equilibrate" the market, then perhaps the government can.

In fact, this latter view is exactly the one adopted by central banks around the world. The basic idea is very simple. An exogenous increase in aggregate demand results in a positive "output gap." A positive output gap implies that the current demand for output is outpacing the current "long-run" supply of output. To a central bank, the primary concern associated with a positive output gap is that since "demand exceeds supply," the result is likely to be inflationary pressure. To combat the inflationary pressures that result from an "overheating" economy, the central bank must try (by whatever means at its disposal) to increase the real rate of interest. By acting in this manner, the central bank can "cool off" the "excess" aggregate demand (i.e., move the economy from point A to point B, instead of point C). Of course, the opposite would be true in the event of a negative "aggregate demand shock." In this case, the central bank would be motivated to cut the real rate of interest. Failure to do so would result in a negative output gap and deflationary pressure.

7. Summary

Capital is a durable asset that augments the economy's productive capacity. While capital depreciates in value as it is used in production, it may also be augmented with investment. Since current investment translates into future capital, investment constitutes a form of domestic saving. The demand for domestic investment depends primarily on the opportunity cost of investment (the real rate of interest) and the expected productivity of future capital. Changes in either of these factors is likely to influence the level of domestic investment demand.

An economy's trade balance is determined by the difference between domestic saving and investment. When desired domestic saving exceeds domestic investment demand, the excess saving is exported (in exchange for foreign bonds). When desired domestic saving falls short of domestic investment demand, the difference is made up by imports (in exchange for domestic bonds). A negative trade balance is not necessarily the sign of a weak economy. It may, for example, reflect high expectations over the future return to investing domestically.

In a closed economy, well-functioning financial markets ensure that the real rate of interest adjusts to equate the supply and demand for loanable funds. An alternative view, however, argues that price variables do not adjust quickly enough to bring balance in the supply and demand for loanable funds. If this is the case, then quantity variables may serve as the economy's equilibrating variables. The effect of various types of shocks would then imply the inefficient movement in quantities (like GDP), rather than the efficient movement in prices (like the real interest rate). This latter view appears to describe the conventional wisdom. In particular, central banks around the world appear to view their role as governors of the interest rate.

Problems

1. Consider a firm that must borrow at interest rate R to finance a new capital expenditure x. The expected return associated with this capital expenditure is $z_2 f(x)$. As explained in the text, the optimal level of investment is given by an x^D that satisfies $z_2 f'(x^D) = R$.

   (a) Demonstrate that x^D also maximizes the net present value (NPV) of the firm; i.e.,
   $$V \equiv -x^D + \frac{z_2 f(x^D)}{R}.$$

   (b) Consider an economy with an investment technology as described in the question above. In this case, $c_2 = z_2 f(x)$ and $x = y_1 - c_1$. If we combine these two relations, we can derive the economy's production possibilities frontier (PPF); i.e., $c_2 = z_2 f(y_1 - c_1)$. Plot this function on a diagram in $c_2 - c_1$ space and demonstrate that the slope of the PPF is the (negative of the) expected marginal product of capital.

   (c) In the economy described above, absent any investment, the economy's GDP flow is given by $(y_1, 0)$. Assume, for the moment, that a representative household is simply endowed with $(y_1, 0)$, but can borrow or lend at an interest rate R. Derive the household's intertemporal budget constraint and plot it in the diagram you used to answer the previous question.

   (d) From part c, you can see that the household's wealth (measured in present value) is given by y_1. But if this household owns the economy's investment technology, then its wealth will be equal to $W = y_1 + V$ (where V was determined in question 1). Assuming that the household now owns the investment technology, derive its intertemporal budget constraint and depict it on the same diagram you used to answer part c. Hint: the budget line will be tangent to the PPF; so that $MPK = R$. On the same diagram, depict the level of domestic investment demand.

   (e) On the same diagram, depict the representative household's desired consumption plan. Compare the desired level of saving s^D and the desired level of investment x^D. Can you draw this diagram in a manner that is consistent with Figure 7.4?

   (f) Finally, assume now that the economy is closed to trade. Depict (on a similar diagram) the equilibrium level of saving and investment, and the equilibrium interest rate.

2. In what sense is it incorrect to refer to measured GDP as "aggregate demand?" Does the fact that measured GDP fluctuates necessarily constitute evidence that the economy is subject to "aggregate demand" shocks?

3. Write a short essay (no math or diagrams) explaining why the government (or central bank) should try to manipulate the interest rate over the business cycle.

4. Write a short essay (no math or diagrams) explaining why the government (or central bank) should not try to manipulate the interest rate over the business cycle.

CHAPTER 8

Money and Inflation

1. Introduction

It may appear curious to most of you that in discussing theories of the macroeconomy, we have to this point avoided mentioning money. In Chapter 2, for example, we thought of the labor market as a place where households exchange their labor for goods. In reality, households exchange their labor for money first, and then exchange their money for goods. But does this intermediate step really matter? It is evident that people do not value money for its own sake (you cannot eat money, for example). When people exchange labor for money, they do so because they expect to use their money as a claim against output. To the extent this is true, we see that money is just a "veil" that conceals the true fundamentals that underlie the motivation for exchange. The earlier chapters in this text simply removed this veil.

On the other hand, the fact that money is used in exchange does imply that money must serve some role in the economy. We all have an intuitive sense that money facilitates the exchange process. That is, monetary exchange helps people engage in trades that they might not otherwise engage in without money to serve as a means of payment. In this chapter, I develop a simple model intended to capture this basic idea. I then use the model to help us understand the determination of a variety of nominal economic variables, for example, the price-level, inflation, the nominal interest rate, and the nominal exchange rate.

2. A Simple Monetary Model

Consider an economy where time is indexed by $t = 1, 2, ..., \infty$. At the initial date 1, there are N agents who live for one period only. Call these agents the "initial old." These agents are endowed with nothing but money. If we let M denote the initial money supply, then each initial old agent is endowed with (M/N) dollars. The initial old value consumption; which I denote c_0. If possible, they would like to trade their money for output.

But with whom might the initial old trade their money? Imagine that at date 1, there also exist N agents who are endowed with no money, but are instead endowed with one unit of time and an ability to produce (nonstorable) output with a technology $y = z(1 - l)$. Here, $z > 0$ denotes an exogenous productivity parameter, and l denotes leisure (time devoted to home production). Assume that these agents live for two periods; so that at date 1 they are "young" and at date 2 they are "old." Assume further that these agents do not value output when young; but that they do value leisure when young. These agents also value output when they become old. We can summarize their preferences with a utility function $u(l, c)$; where l denotes leisure when young and where c denotes consumption when old.

Let's summarize what we have so far. At date 1, we have N initial old agents who desire output, but have nothing to offer in exchange except possibly money. At date 1, we also have N initial young agents who might potentially produce output (at the cost of foregone leisure). The question here is whether they young might be willing to work for money.

The answer to this last question is not so obvious. After all, work requires effort (foregone leisure). And the young cannot eat money (paper tastes lousy). But seeing as how the young value output in the future when they are old (much like the current old value output today), they young might be willing to work for money **if** they expect money to have value in the future. To allow for such a possibility, let us assume that at each date $t \geq 2$ there is a new generation of young agents of size N who enter the economy. Assume that these new generations of young agents have identical preferences and are endowed in exactly the same way. Since everyone lives for exactly two periods, the population of this economy will forever be fixed at $2N$ (an equal number of young and old at each date).

Now, let us imagine that money is used to purchase output at each date on a competitive spot market. Let p_t denote the price of output measured in units of money at date t (this is the date t **price-level**). Let us take as given, for the moment, a sequence of prices (a **price-system**) $\mathbf{p} \equiv \{p_1, p_2, ..., p_\infty\}$. For money to have value, it must be the case that $p_t < \infty$ at every date t; so let us assume (for the moment) that this is the case (we will verify later that this will be the case).

Next, consider the choice problem facing a **representative** young agent at some date t. This agent has the option of producing output $y = z(1 - l)$. Since the young agent does not value this output, his best option is to sell it for money m_t. If the price of output is p_t, then:

$$m_t = p_t z(1 - l). \tag{44}$$

The money accumulated in this manner can then be carried over into the next period, where it can be used to purchase output. This implies:

$$p_{t+1} c = m. \tag{45}$$

Combining equations (44) and (45), we have:

$$c = \frac{p_t}{p_{t+1}} z(1 - l). \tag{46}$$

Equation (46) describes how a young agent can trade off current leisure for future consumption, given a price-system $\mathbf{p}$. That is, equation (46) is the agent's **intertemporal budget constraint**.

Let $\Pi_{t+1} \equiv p_{t+1}/p_t$. That is, Π_{t+1} represents the expected (gross) rate of **inflation** (the rate of change in the price-level). In what follows, I am going to restrict attention to stationary equilibria. That is, let us assume that $\Pi_{t+1} = \Pi$. In this case, we can state the representative young agent's choice problem formally by the statement:

$$\boxed{\text{Choose } (l, c) \text{ to maximize } u(l, c) \text{ subject to: } c = \Pi^{-1} z(1-l).} \tag{47}$$

The solution (l^D, c^D) can be depicted in the usual way by using a familiar diagram; see Figure 8.1.

FIGURE 8.1
The Demand for Real Money Balances

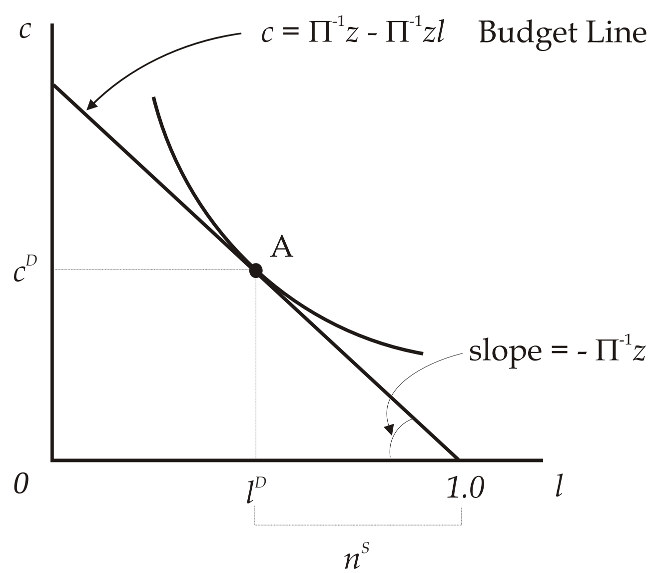


Figure 8.1 is labeled "the demand for real money balances." This label is motivated by equation (44), $m_t = p_t z(1-l)$. That is, a young agent who works $(1-l)$ hours produces $z(1-l)$ units of output, which he exchanges for m_t dollars. Hence, $m_t^D \equiv p_t z(1-l^D)$ denotes the demand for **nominal** money balances. The purchasing power of m_t dollars is given by m_t/p_t. Therefore, the demand for **real** money balances is given by:

$$\left(\frac{m_t^D}{p_t}\right) = z(1-l^D).$$

For notational convenience, let me define $q_t \equiv m_t/p_t$. In this case, we can write the demand for real money balances as:

$$q^D = z(1 - l^D). \tag{48}$$

Note that the demand for real money balances depends on the productivity parameter z. It will also depend on z and Π via the effect that these variables have on labor supply $n^S = (1 - l^D)$.

2.1 Properties of the Money Demand Function

According to this theory, the demand for real money balances should depend on both z and Π; a fact that we can stress by writing $q^D(z, \Pi)$. Keep in mind that Π should be interpreted here as the **expected** inflation rate.

From Figure 8.1, it is clear that an increase in Π will serve to make the budget constraint "flatter." That is, future consumption (which can only be acquired with accumulating cash today) becomes more expensive relative to leisure. Alternatively, you can think of $\Pi^{-1}z$ as the expected return to labor. As with any relative price change, there will be substitution and wealth effects to consider. By equation (48), it is clear that the reaction of q^D will depend on the reaction of l^D (n^S). In what follows, I will assume that the substitution effect dominates the wealth effect on labor supply when the return to labor changes. In this case, q^D will be a decreasing function of Π.

From Figure 8.1, it is clear that an increase in z will serve to make the budget constraint "steeper." Since I have assumed that the substitution effect dominates the wealth effect, this increase in the return to labor will increase n^S (decrease l^D). This effect alone will serve to increase the real demand for money. But there is an added effect as well. Since an increase in z increases wealth, the demand for future output c^D also increases owing to a positive wealth effect. As future consumption can only be acquired by first accumulating cash balances, the demand for money increases for this reason (even if n^S was to remain unchanged).

Exercise 8.1: Is there any logic in labeling c a "cash good" and l a "non-cash good?" Explain how an expected inflation affects the relative price of cash goods vis-à-vis non-cash goods.

Exercise 8.2: Suppose that the preferences of a representative young agent are described by the utility function $u(l, c) = \ln(l) + \beta c$, where $\beta \geq 0$ is a preference parameter. For these preferences, $MRS(l, c) = 1/(\beta l)$. Demonstrate that the demand for real money balances is in this case given by:

$$q^D(z, \Pi) = z - \frac{\Pi}{\beta}.$$

Note that for these preferences, if the expected inflation rate is "too high," then the demand for money will be zero. What is the upper bound on inflation here?

2.2 Monetary Equilibrium

To this point, I have simply assumed a given price-system $\mathbf{p}$. Since I have also assumed a constant Π, I can summarize $\mathbf{p}$ by the pair (p_1, Π). That is, if I know p_1 and Π, then I know $p_2 = \Pi p_1$, $p_3 = \Pi p_2$, and so on. In this way, I can recover the entire price-system $\mathbf{p}$. The question here is how to determine the equilibrium (p_1^*, Π^*)?

To answer this question, we need to look at the market-clearing conditions that must hold in this economy at every date. Remember that each date, there is only one market. On this market, the young produce and sell their output for money. Hence, we are dealing with a sequence of "money-output" markets, with p_t denoting the price of output relative to money at date t.

For a given Π, each young agent demands $q^D(z, \Pi)$ in real money balances. Since there are N young agents, the aggregate demand for real money balances at each date is given by $Nq^D(z, \Pi)$. Likewise, at each date, there are N old agents holding the entire money supply M. Hence, the aggregate supply of real money balances at date is given by M/p_t. Equilibrium in the money-output market at date t therefore requires:

$$\frac{M}{p_t} = Nq^D(z, \Pi). \tag{49}$$

Note that this condition must hold at every date $t \geq 1$. Hence, it must also hold at date $t+1$; i.e.,

$$\frac{M}{p_{t+1}} = Nq^D(z, \Pi).$$

Now, divide the former equation by the latter to derive:

$$\frac{p_{t+1}}{p_t}\frac{M}{M} = \frac{N}{N}\frac{q^D(z, \Pi)}{q^D(z, \Pi)}. \tag{50}$$

Cancelling terms, we see that this equation implies that $p_{t+1} = p_t$; that is, market-clearing at each date implies that the equilibrium price-level must be constant over time. Of course, this must imply that $\Pi^* = 1$ (zero inflation).

Now that we know $\Pi^* = 1$, we are in a position to determine p_1^*. How can we do this? This can easily be done by noting that condition (49) must also hold at date 1; i.e.,

$$p_1^* = \frac{M}{Nq^D(z, 1)}. \tag{51}$$

Now that we have determined the equilibrium price-system $\mathbf{p}^* = (p_1^*, \Pi^*)$, we can determine the equilibrium allocation. Since the initial old each start off with (M/N) dollars, their consumption is given (from their budget constraint) by:

$$c_0^* = \frac{M}{Np_1^*} = q^D(z, 1).$$

That is, their consumption is dictated by the amount of output the initial young are willing to hand over for money; this is just q^D.

Now, what about the young generations? Each young person works (in equilibrium) $n^* = 1 - l^D(z,1)$ hours. Hence, each young person produces $y^* = zn^*$ units of output, which they exchange for real money balances q^*. That is,

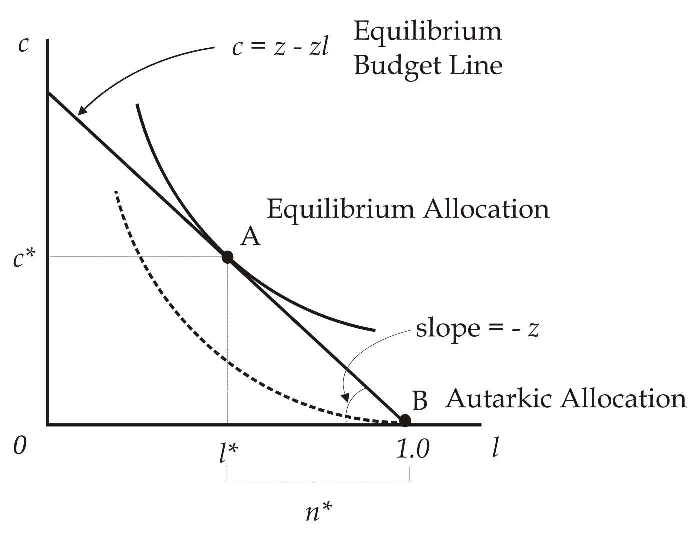
$$c^* = y^* = q^* = q^D(z,1).$$

Note that $c^* = c_0^*$. That is, in this monetary equilibrium, all old agents (including the initial old) end up consuming exactly the same amount of output (note too that the initial old were endowed with nothing but money).

2.3 The Welfare-Enhancing Role of Money

The monetary equilibrium described above is depicted in Figure 8.2, with the equilibrium allocation depicted by point A. One may well ask whether money is playing any essential role in this economy. To answer this question, let us try to imagine how this model economy might function if money did not exist.

FIGURE 8.2
Monetary Equilibrium vs. Autarky



The first thing that is clear is that if money did not exist, then the initial old would not be able to purchase any output. That is, they have nothing to offer the young in exchange for their labor services.

What about the young agents? In the monetary equilibrium, they worked for money at date t and they used this money to purchase output at date $t+1$. But if there is no money, how are they to acquire the resources necessary to consume in the future? One possibility is that they might work in exchange for a bond (a claim against output at date $t+1$). The only people in a position to deliver output at date $t+1$ are the new generation of young at date $t+1$. But these latter agents only value output at date $t+2$; they do not value the output produced by the current young at date t. In other words, this economy features a complete **lack of double coincidence of wants** (there are no gains to trade between any two individuals). In the absence of money then, all trade shuts down. This autarkic allocation is depicted by point B in Figure 8.2.

In autarky, the representative young person simply consumes their endowment (they consume zero output and all of their leisure). This autarkic allocation is clearly less preferable than the allocation they could achieve in an economy with money. Hence, we can conclude here that money plays an essential role in enhancing economic welfare. Everyone (including the initial old) are made better off in an economy with money. This is because money helps people trade when there is a lack of double coincidence of wants.

2.4 Money Neutrality

A classic question in monetary theory is whether changing the quantity of the money supply in the economy might have any real (rather than nominal) effects. The thought experiment runs as follows. Imagine that we all wake up one morning and see that our money balances have magically increased by a factor of ten. We all have ten times more money than before. Might this have any effect on the real level of economic activity? For example, might this lead to an increase in the real GDP?

We can answer this question easily within the context of the model developed above. In that model, the initial old are endowed with M/N dollars each, and everyone else has zero money balances to start off with. Imagine now that we increase everybody's initial money holdings by a factor of $\lambda > 1$. Now the initial old have $(\hat{M}/N)$ dollars each, where $\hat{M} = \lambda M$; and everyone else continues to have zero money balances.

According to equation (50), this **one-time** increase in the money supply is predicted to have **no effect** on the expected (and actual) equilibrium inflation rate. Because this is so, it can have no effect on the demand for **real** money balances q^D. Hence, there can be no effect on the level of employment or the level of **real** GDP. The only effect would appear, from equation (51) to increase the equilibrium price-level p_1^* to $\hat{p}_1^* = \lambda p_1^*$. That is, the only effect is to increase the price-level (a nominal variable) by the same factor as the increase in the money supply. After this initial jump in the price-level, the expected (and actual) inflation rate remains constant; i.e. $\Pi^* = \hat{p}_{t+1}/\hat{p}_t = 1$ for all $t \geq 1$.

Because the price-level is permanently higher, the value of all nominal variables is higher in the same proportion.

When a one-time increase in the money supply has no effect on any real economic variables, we say that **money is neutral**. The reason for why money is neutral in this experiment is because the added money simply leads the initial old to bid up the price of output. Since the inflation rate is expected to remain unchanged after the initial price jump, the young are not motivated to supply any more output (their demand for real money balances remains unchanged). Since there is more money chasing the same amount of output available for sale, the price-level must rise to clear the money-output market.

The thought experiment described above may sound strange, but it is of interest to note that similar policies are carried out in reality–usually in the opposite direction. In 1998, for example, the then President of the Russian Federation, Boris Yeltsin, announced that the Russian money supply was to be reduced by a factor of 1000 beginning January 1. That is, beginning on that date, anyone with "old" Russian roubles were asked to redeem them for "new" Russian roubles with three fewer zeros. This redenomination of the Russian rouble apparently went off without any major repercussions–consistent with the prediction of money neutrality.

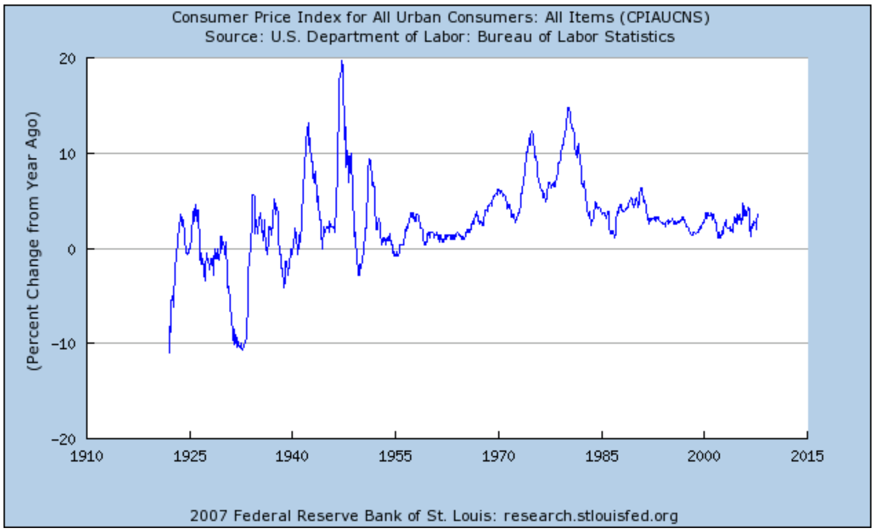
This is **not** to say that **any** one-time increase in the money supply will always be neutral. The key to understanding the likely effects of any one-time change in the money supply is to understand how the new money is to be injected into the economy. For example, instead of increasing the money supply in proportion to everybody's initial money balances, imagine that the new money is injected disproportionately. In the context of our model economy, we can think of the new money $(\lambda - 1)M$ being transferred to the initial young instead of the initial old. Injecting new money in this manner will have the effect of transferring purchasing power away from the initial old toward the initial young. That is, the new money leads to an increase in the price-level, as before; but as the initial old have the same amount of money, their purchasing power declines. This effect, however, only lasts for one period in our model, so that money ends up being non-neutral only in the short-run.

3. Inflation

3.1 Evidence

Consider Figure 8.3, which plots the rate of change of the U.S. price level since 1921. What are the striking properties of the inflation rate over this sample period?

FIGURE 8.3
U.S. Inflation 1921-2007

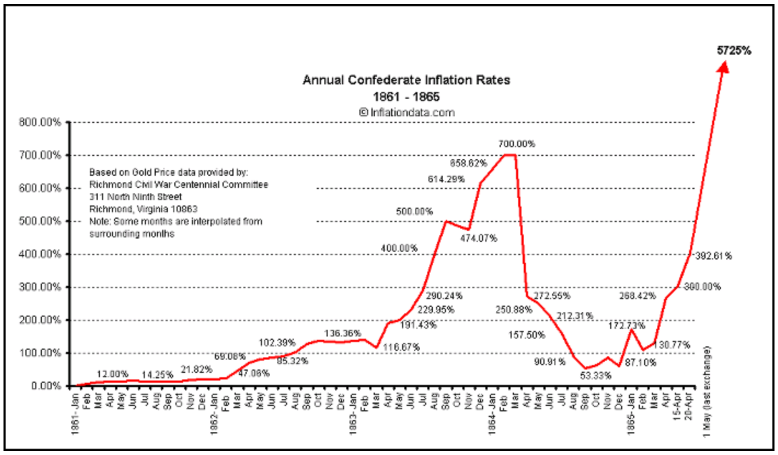


The first thing to note is that the sample can roughly be divided between pre-1955 and post-1955. In the early part of the sample, both very high and very low inflation rates appear to occur with equal frequency. In fact, the inflation rate is frequently negative–that is, there appear to be frequent periods of **deflation** (a declining price-level). Since 1995, the inflation rate has remained positive (so that the price-level is always rising, although at different rates of change).

The second thing to note is how the episodes of inflation and deflation correlate with the business cycle. On the whole, there appears to be no consistent pattern. The sharp deflation in the mid 1920s was associated with a boom. The sharp deflation in the early 1930s was associated with the Great Depression. The three sharp spikes of inflation during the 1940s and early 1950s were generally boom periods. The last two spikes, occurring in the 1970s, were associated with recession.

One thing that appears clear enough, not just from this data, is that high inflation is frequently associated with periods of war. Consider what happened to the inflation rate in the Confederate States during the U.S. Civil War; see Figure 8.4.

FIGURE 8.4
Inflation in the Southern Confederacy 1861-1865



Similar patterns are evident for every major war that I am aware of (going far back into history). If there is one consistent pattern, we know that it is this: **wars are associated with inflation**.

Why are major wars associated with inflation? There appear to be two main reasons:

1. During a major war, the government needs to acquire a large amount of resources (goods and manpower); and
2. Printing money seems like a relatively painless way for the government to acquire some of these resources.

It appears, however, that a major war is not the only "cause" of high inflation. For example, the high inflation experienced in the United States during the 1970s has been called "The Great Peacetime Inflation."[56] We also know that high inflations often plague underdeveloped economies even when they are at peace. The great common denominator in all these examples, however, appear to be governments that feel the need to acquire resources by printing money. For underdeveloped economies, this may be because they do not have extensive tax collection agencies in place (collecting taxes from remote villages may be difficult). For developed economies, some combination of higher taxes, deficits, and money creation may be an optimal way to finance a transitory increase in government spending.

[56] Perhaps this is a bit of a misnomer, however. In particular, Figure 8.2 shows how the U.S. inflation rate generally rose during the escalation of the Vietnam war (1965-75).

3.2 Seigniorage

We know that governments can acquire resources by taxing its citizens directly. We also know that a government may acquire the same resources instead by issuing bonds. A bond finance, however, is just another form of taxation (a postponement of taxes off into the future). But when the government is in control of the money supply, it has a third way to acquire resources; i.e., it can print money. This latter method of government finance is called **seigniorage**.

As we shall see, seigniorage turns out to be just another tax. In particular, the act of printing new money to acquire resources will generate an inflation. And because inflation reduces the return from holding money, it acts as a tax on all those who choose to hold money. The new money created by the government competes with the old money held by households. As the government prints more money, it can acquire more resources; but this comes at the expense of the broader population. In other words, seigniorage implies an **inflation tax**.

To see how this works, let me extend our model above in a straightforward way. First, let me assume that the only way the government acquires resources is by printing money. This will allow us to focus on the effects of seigniorage as a revenue device. As before, assume that the initial old are endowed with some given quantity of money M_0. Then, at each date $t \geq 1$, the government expands the money supply at a constant (gross) rate $\mu \geq 1$; so that:

$$M_t = \mu M_{t-1}.$$

The new money available for purchasing output at date t is therefore given by $M_t - M_{t-1}$; or, combining this with the equation above:

$$M_t - M_{t-1} = \left[1 - \frac{1}{\mu}\right] M_t.$$

To derive the purchasing power of this new money, we have to divide it by the price-level. Hence, the government's real seigniorage revenue is given by:

$$S_t \equiv \left[1 - \frac{1}{\mu}\right] \frac{M_t}{p_t}. \tag{52}$$

Let us now consider household behavior. Because the government (in this example) levies no direct taxes, the choice problem facing a representative young agent as described by (47) remains unchanged. The solution to their choice problem can be summarized by their demand for real money balances $q^D(z, \Pi)$. Recall that we are assuming that this demand function is decreasing in the expected inflation rate Π.

Next, consider the market-clearing condition:

$$\frac{M_t}{p_t} = Nq^D(z, \Pi). \tag{53}$$

Note that as the right-hand-side of this equation remains constant over time, so must the left-hand-side. Since M_t grows at rate μ, it follows that p_t must grow at the same rate. In other words, $\Pi^* = \mu$.

The final step is to combine (53) with (52) to derive:

$$S(\mu) = \left[1 - \frac{1}{\mu}\right] Nq^D(z, \mu). \tag{54}$$

This equation tells us how much real seigniorage revenue the government can extract from the economy in a monetary equilibrium when it expands the money supply at rate μ.

The amount $S(\mu)$ can be thought of real tax revenue (revenue extracted by way of an inflation tax). The term $[1-1/\mu]$ in (54) can be thought of as the **tax rate** and the term $Nq^D(z, \mu)$ can be thought of as the **tax base**. An increase in μ is seen to have two effects. First, it increases the tax rate. Second, it reduces the tax base (as households substitute out of cash goods into non-cash goods). How seigniorage revenue responds to an increase in the inflation tax depends on the balance of these two effects; see Figure 8.5.

FIGURE 8.5
The Limits to Seigniorage

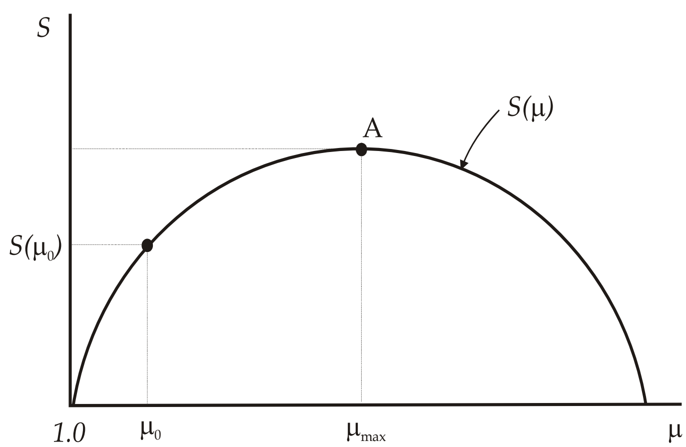
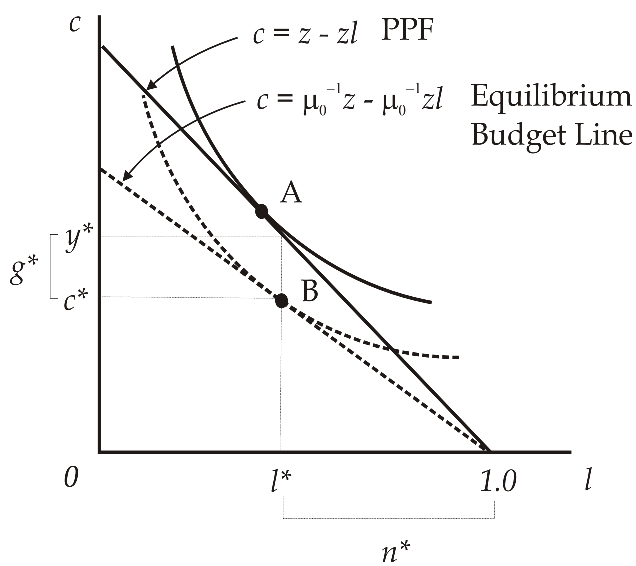


Figure 8.5 is reminiscent of a "Laffer curve." That is, for low tax rates, an increase in the tax rate increases tax revenue. But for high enough tax rates, further increases in the tax rate actually reduce tax revenue (as people substitute out of the object that is being taxed). We can conclude from this analysis that the power to print money does not give the government an unlimited ability to extract resources from the economy. Household behavior will place some discipline on the government's ability to tax by way of inflation. That is, at

very high inflation tax rates, the demand for real money balances will be very low; and may ultimately approach zero in a hyperinflation.

To see how inflation affects the economy, imagine that the government expands the money supply at rate μ_0 (see Figure 8.5) and uses the revenue to purchase (and consume) output. The monetary equilibrium allocation is depicted by point B in Figure 8.6 (point A represents the equilibrium with zero inflation).

FIGURE 8.6
Monetary Equilibrium with Inflation

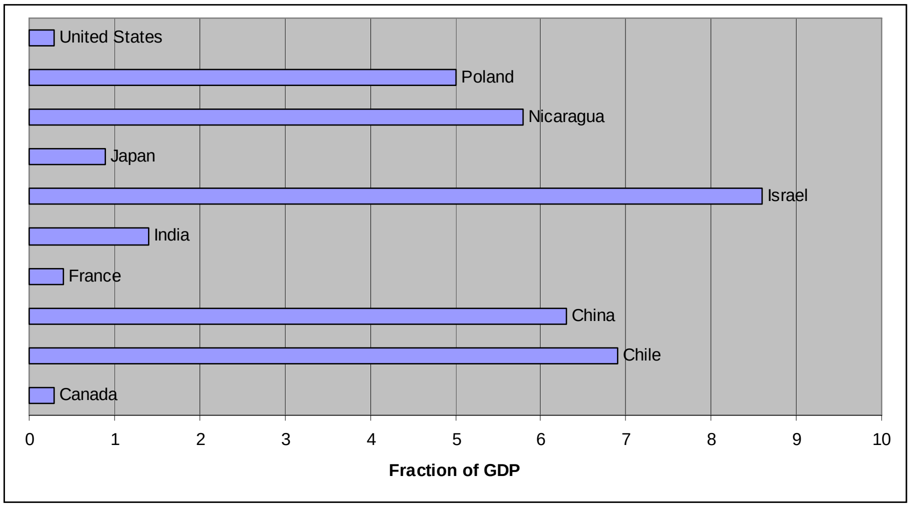


Exercise 8.3: Compare Figure 8.6 with Figure 6.4 and note any similarities.

As with the case of a distortionary income-tax (studied in Chapter 6), the effect of the inflation tax here is to reduce the return to labor. This is because labor is used to accumulate cash today, which must be carried over to the future to purchase consumption (which costs more owing to the rise in the price-level). The effect here is to reduce the equilibrium level of employment and GDP. The resulting seigniorage revenue is used to purchase $g^* = S(\mu_0)/N$ units of output. (Note that the income-expenditure identity holds here: $y^* = c^* + g^*$).

How important is seigniorage revenue in reality? The answer appears to vary across countries and across time. For most developed economies, seigniorage revenue accounts for a relatively small fraction of government revenue. However, for less developed economies, it appears that seigniorage revenue constitutes a significant source of government revenue; see Figure 8.7.

FIGURE 8.7
Seigniorage Revenue for Selected Countries (1960-1999)[57]



4. Money, Capital, and Banking

In the model economy studied above, the only source of money in the economy was assumed to be government money. While these days, governments do maintain monopoly control over small denomination paper notes, it is important to keep in mind that this form of money is not the only asset that is used in payments. In fact, the vast majority of an economy's money supply is created by the private sector; primarily in the form of debt instruments created by the banking sector. Historically, these private debt instruments were issued in paper form. In modern economies, they exist primarily as "electronic money credits" (electronically recorded book entry items in bank accounts). Whenever you make a payment with a debit card, you are using a form of this electronic money. I will henceforth refer to such money as **private money**.

The total money supply is therefore given by:

$$\text{Total Money Supply} = \text{Government Money} + \text{Private Money}.$$

Government money–in the form of paper notes–is referred to as **base money** (or the monetary base). In the context of our model, this is denoted by M_t. If

[57] Source: Aisen, Ari and Francisco José Viega (2005). "The Political Economy of Seigniorage," IMF WP/05/175.

we let D_t denote the supply of private money (demand deposit liabilities created by private banks), then we can define:

$$M1_t \equiv M_t + D_t; \tag{55}$$

where M1 denotes the total money supply. Another important concept here is that of the **money multiplier**; defined as $M1/M$. Using the relation above, we see that:

$$\frac{M1_t}{M_t} = \left[1 + \frac{D_t}{M_t}\right]. \tag{56}$$

That is, the money multiplier depends on the **deposit-to-money ratio** D_t/M_t.

Empirically, M1 appears to behave procyclically (the correlation between M1 and real GDP is positive). Moreover, M1 appears to lead real GDP over the cycle. That is, $M1_t$ is strongly correlated with future GDP y_{t+1}. Most of the variation in $M1_t$ is attributable to movements in D_t, rather than M_t. In other words, the money multiplier is highly variable.

4.1 A Simple Model

To help interpret these patterns, I want to modify our model somewhat. Let me first assume that young agents do not value leisure. As before, assume that agents only value consumption when they are old. This future consumption, however, comes in two types: A and B. Hence, let me denote preferences by $u(c_A, c_B)$ with $MRS(c_A, c_B)$.

A young agent has one unit of time. This time can be used (when young) to produce either of two goods: q (the type A good) or x (an investment good). Let n denote the fraction of time spent producing y_A. If we let $z_1 > 0$ denote the worker's productivity, then:

$$\begin{aligned} q &= z_1 n; \\ x &= z_1(1-n). \end{aligned}$$

Note that $q+x = z_1$. Moreover, note that the young agent does not value either of these goods (the c_A in his utility function represents future consumption).

Now, assume that the A good can only be purchased with cash (government money). Since the young agent values c_A in the future, this assumption ensures that he is motivated to produce q in exchange for cash (which will be purchased by the current old generation, who value the type A good today, and who happen to be holding cash). Hence, the sale and purchase of this "cash good" must obey the following budget constraint:

$$p_{t+1}c_A = p_t q.$$

That is, the young agent sells q units of type A good today at price p_t for money $p_t q$. He then carries this money forward in time where he then purchases c_A at price p_{t+1}. Note that we can rewrite this budget constraint as follows:

$$c_A = \Pi^{-1} q. \tag{57}$$

Next, let me assume the existence of a business sector in control of an investment technology. This technology is such that if x units of output (in the form of new capital goods) are invested today, the return to this investment is equal to $y_B = z_2 f(x)$ units of type B output tomorrow (I assume that capital depreciates fully after it is used in production). You can think of this business sector as consolidating the functions of a firm and bank. Assume that the bank is owned by some other agents, who simply consume all bank profits.

The bank has no resources of its own. In order to acquire new capital x, it must borrow the resources from the young. Let R denote the (gross) real rate of interest. Then the bank can "pay" workers for x with "banknotes" that promise to deliver Rx units of type B output tomorrow. Or, instead of issuing banknotes, you can think of the bank as simply opening an interest-bearing bank account for the worker. When the future rolls around, the worker can then just pay for type B output by drawing on this account (e.g., a debit card transaction). For this reason, I will refer to c_B as a "credit good."

Thus, the young agent is also subject to the following budget constraint:

$$c_B = Rx. \tag{58}$$

Since $q = z_1 - x$, I can rewrite (57) as $c_A = \Pi^{-1}(z_1 - x)$ or $x = z_1 - \Pi c_A$. If we plug this last expression into (58), we have:

$$c_B = Rz_1 - R\Pi c_A.$$

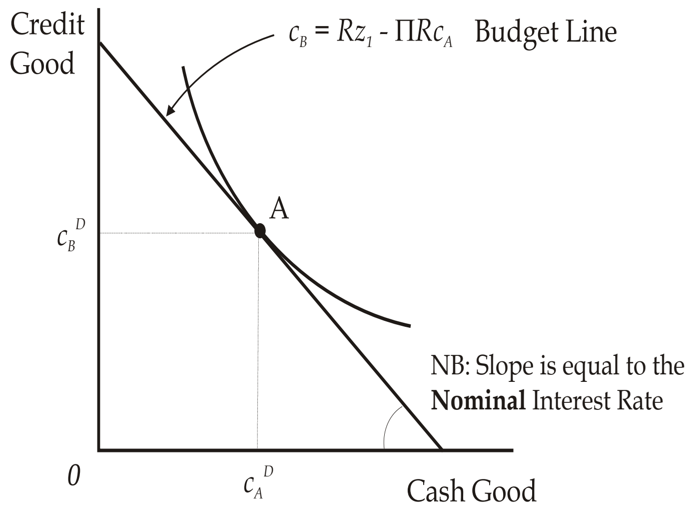
This equation describes how the young agent can trade off c_A and c_B for a given real interest rate R and inflation rate Π. Note that the slope of this budget line is $-R\Pi$. This product of these two objects is called the (gross) **nominal interest rate**; and I will denote it by $R^n \equiv R\Pi$.

For a given (R, Π), the solution to the young agent's choice problem is given by a pair of demand functions (c_A^D, c_B^D) that satisfy:

$$\begin{aligned} MRS(c_A^D, c_B^D) &= R\Pi; \\ c_B^D &= Rz_1 - R\Pi c_A^D. \end{aligned}$$

This solution is depicted in by point A in Figure 8.8.

FIGURE 8.8
The Demand for Cash and Credit Goods



Exercise 8.4: Explain how an increase in the nominal interest rate (say, brought about by an increase in the expected inflation rate) makes cash goods more expensive relative to credit goods. How would you expect a higher nominal interest rate to affect the demand for real money balances? Explain.

The solution to the young agent's choice problem depends on the parameters (R, Π, z_1). Once this solution is known, we can easily recover the demand for real money balances (the output supplied for cash) and the supply of investment goods from the budget constraints (57) and (58); i.e.,

$$\begin{aligned} q^D(R, \Pi, z_1) &= \Pi c_A^D(R, \Pi, z_1); \\ x^S(R, \Pi, z_1) &= R^{-1} c_B^D(R, \Pi, z_1). \end{aligned} \tag{59}$$

Let us now turn to the business sector. We can write the bank's profit function either in nominal or real terms. Measured in units of money, the bank's nominal profit is given by:

$$p_{t+1} z_2 f(x) - R^n p_t x. \tag{60}$$

That is, the bank "purchases" (borrows) x units of output at price p_t and promises to repay $R^n p_t x$ dollars worth of output in the future. Note that R^n

here represents the nominal interest rate. The bank then produces $z_2 f(x)$ units of output in the future, which is worth $p_{t+1} z_2 f(x)$ dollars.

Alternatively, we can write the bank's profit function in real terms. To do so, divide both sides of (60) by p_{t+1} to derive:

$$z_2 f(x) - R^n \Pi^{-1} x;$$

or,

$$z_2 f(x) - Rx;$$

as $R = R^n \Pi^{-1}$ (the real interest rate is the nominal interest rate divided by inflation). This problem should look familiar to you from Chapter 7. If the function f is increasing and strictly concave, the demand for investment x^D satisfies:

$$z_2 f'(x^D) = \frac{R^n}{\Pi}.$$

That is, the demand for investment depends negatively on the real interest rate and positively on the expected productivity of future capital.

Although this is not crucial to the analysis, let me make a simplifying assumption; i.e., that $f(x) = x$ (the investment technology is linear). This is similar to the simplifying assumption I made in Chapter 2. In this case, the bank's real profit is given by $(z_2 - R)x$. If we assume free-entry into the banking sector, then competition among banks will drive the real interest to a point where profits are zero. This allows us to determine the equilibrium **real** rate of interest as:

$$R = z_2.$$

That is, the real interest rate will be equal to the expected marginal product of capital (this is also consistent with what we learned in Chapter 7). In this case, the demand for investment is indeterminate, so that the equilibrium level of investment will be determined entirely by the supply of investment.

We now have to consider government policy. Here, I assume that the government simply expands the money supply at rate μ and uses the resulting seigniorage revenue to finance government purchases of the cash good. Therefore, from (52), we have:

$$S_t \equiv \left[1 - \frac{1}{\mu}\right] \frac{M_t}{p_t}.$$

And finally, we invoke the market-clearing condition (53):

$$\frac{M_t}{p_t} = N q^D(R, \Pi, z_1).$$

Since this condition must hold at every date, it follows that the equilibrium inflation rate is given by:

$$\Pi = \mu.$$

The equilibrium level of seigniorage revenue is (as before) given by:

$$S^* = \left[1 - \frac{1}{\mu}\right] Nq^D(R, \Pi, z_1);$$

where $R = z_2$ and $\Pi = \mu$.

4.2 Equilibrium

Let me summarize what we have derived in the previous subsection. We know that the equilibrium real rate of interest is given by $R = z_2$. We know that the equilibrium inflation rate is given by $\Pi = \mu$. Hence, it follows that the equilibrium consumption allocation can be described by point A in Figure 8.8 by setting $(R, \Pi) = (z_2, \mu)$; i.e.,

$$\begin{aligned} c_A^* &= c_A^D(R, \Pi, z_1); \\ c_B^* &= c_B^D(R, \Pi, z_1). \end{aligned}$$

From (59), we can then derive the equilibrium quantity of real money balances and investment:

$$\begin{aligned} q^* &= q^D(R, \Pi, z_1) = \Pi c_A^*; \\ x^* &= x^S(R, \Pi, z_1) = R^{-1} c_B^*. \end{aligned}$$

The equilibrium level of real GDP is given by:

$$Y^* = Nz_1 + Nz_2 x^*.$$

From the market-clearing condition, we can derive the equilibrium price-level path by:

$$p_t^* = \frac{M_t}{Nq^*}. \tag{61}$$

The equilibrium level of government spending is just given by S^*.

Now, recall the definition of M1 (55):

$$M1_t = M_t + D_t;$$

where D_t denotes the nominal value of private money. In this model, the nominal value of the liabilities created by the banking sector is given by:

$$D_t^* = p_t^* N x^*$$

Since we know, from the market-clearing condition, that $p_t^* = M_t/(Nq^*)$, we may alternatively write:

$$D_t^* = M_t \left(\frac{x^*}{q^*}\right).$$

Substituting this expression into the definition of M1, we have:

$$M1_t^* = \left[1 + \frac{x^*}{q^*}\right] M_t. \tag{62}$$

The term in the square brackets is the equilibrium money multiplier; see (56).

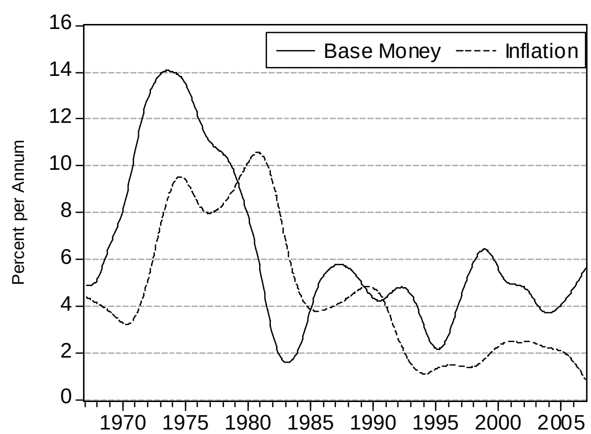
This model economy has four exogenous variables: (z_1, z_2, μ, M_t). We can use the model to see how this economy might react to an exogenous "shock" to any one of these parameters. Alternatively, we might consider two hypothetical economies that are identical in every respect except for one of these exogenous variables. In what follows, I consider two applications.

4.3 Different Money Growth Rates

Here, I consider two economies that are identical in every respect except for the monetary policy parameter μ. Assume that $\mu_H > \mu_L \geq 1$.

Not surprisingly, our model predicts that the economy with the higher money growth rate will experience a higher inflation rate; $\Pi_H > \Pi_L$. This prediction is broadly consistent with the evidence across countries. The evidence is also broadly consistent with the long-run money growth and inflation trends within an economy; see Figure 8.9.

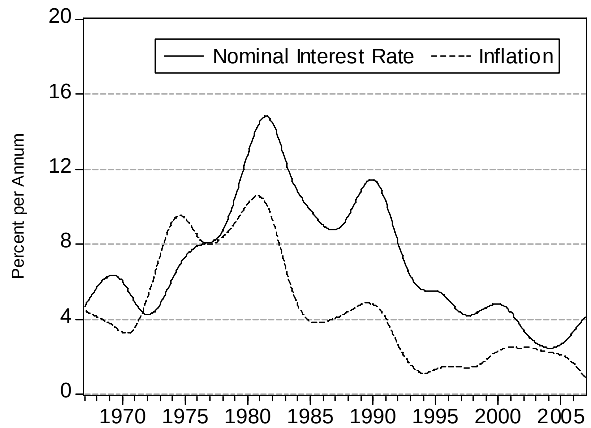
FIGURE 8.9
Base Money Growth and Inflation
Canada 1967 - 2006



Source: CANSIM II. CPI Inflation v735319; Monetary Base v37145

As the real interest rate in these two economies is the same ($R = z_2$), it follows that the economy with the higher inflation rate will also have a higher nominal interest rate; $R_H^n > R_L^n$. This too is broadly consistent with the cross-country evidence; and is also broadly consistent with interest rate and inflation trends within an economy; see Figure 8.10.

FIGURE 8.10
Nominal Interest Rate and Inflation
Canada 1967 - 2006



Source: CANSIM II. Nominal interest rate v122484 (91 day TBill)

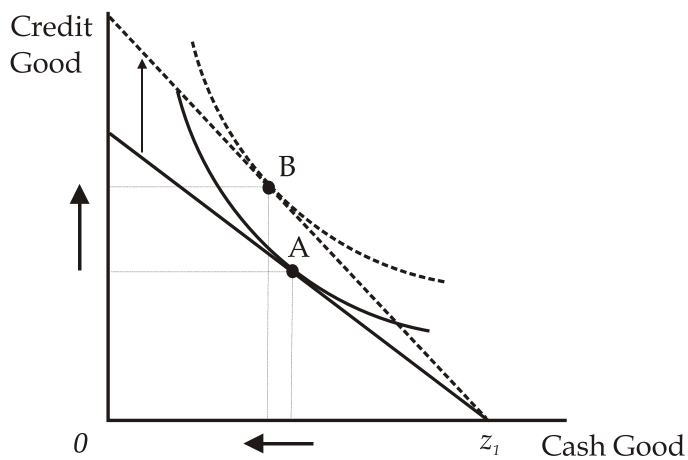
4.4 News Shock

As many economists and other pundits appear to make a big deal about "aggregate demand" shocks and how they influence economic activity and the price-level, let us see how this phenomenon might be understood in the context of our model. I propose to model an "aggregate demand shock" as an exogenous increase in z_2. The shock is imagined to occur at some date t. Note that this shock does not affect current productivity (which remains fixed at z_1). Instead, the shock leads agents to revise upward their estimate on the return to capital at date $t+1$. Because the actual shock occurs at date t, I will refer to it (as in earlier chapters) as a "news shock." Because a positive news shock increases the expected return to capital, the demand for investment (and hence, aggregate demand) will increase at date t. This is the sense in which I think a "news shock" corresponds to an "aggregate demand shock."

As $R = z_2$, the effect of this news shock will be to increase the real interest rate. You can think of this as being caused by the increase in demand for loanable funds. This shock will have no effect on the **expected** inflation rate; as this variable is pinned down by the money growth parameter $\Pi = \mu$ (as we shall see, however, the **realized** rate of inflation may differ from the expected inflation rate in the period of the shock). In what follows, I assume that $\Pi = \mu = 1$ (the price-level is expected to be constant).

Now, take a look at Figure 8.8 and keep in mind that we are assuming here that $\Pi = 1$. Then an increase in R has the effect of making the budget line steeper (the nominal interest rate increases along with the real interest rate); see Figure 8.11.

FIGURE 8.11
A Positive News Shock



Since point B lies to the left of point A in Figure 8.11, I am assuming here that the substitution effect dominates the wealth effect in terms of the demand for the cash good. That is, the higher interest rate leads agents to demand less cash and more interest-bearing money. Hence, the effect of this shock will lead to a "surprise" increase in x^* and a "surprise" decrease in q^*.

From equation (62), the effect of this news is to increase the total money supply $M1_t^*$. Note that as the monetary base M_t is assumed not to change at date t, the entire increase in the money supply here comes from the **private sector** (the money multiplier increases). We also see that this sudden increase in M1 is associated with a "surprise" increase in the price-level; see (61). But the reason for this here is not because M1 increases; instead, it is because the demand for real money balances q^* declines.

Finally, note that this news shock has no effect on the current period real GDP:

$$Y_t = Nz_1 + Nz_2 x_{t-1}^*.$$

This is because Y_t depends on current productivity z_1 and the productivity of capital put in place during the **previous** period (remember that investment becomes productive capital with a one-period lag). Nevertheless, as the news shock causes current period capital to expand, this implies that the **future** real GDP will expand. Hence, our model is consistent with the idea that a current period increase in $M1_t$ is associated with a future increase in real GDP Y_{t+1}.

Exercise 8.5: In the data, it appears that $Cor(M1_t, Y_{t+1}) > 0$. Explain why it might be wrong to assume that an increase in the money supply causes an increase in real GDP. Hint: reverse causality.

5. International Money Systems

A nominal exchange rate represents the relative price of two monies. The behavior of nominal exchange rates are frequently difficult to explain. Sharp movements in nominal exchange rates frequently lead to much commentary and debate. Should the value of a country's currency be determined in the market place (the foreign exchange market) like any other good? Or should a country endeavor to fix the value of its currency relative to another? Should countries attempt to establish a multilateral fixed exchange rate system, such as the Bretton Woods agreement (1945-1971)? Or should countries adopt even more radical measures; for example, by adopting a common currency (like the Euro) or by unilaterally deciding to adopt the currency of another country (Dollarization)?

To give you some idea of what the fuss is all about, consider Figure 8.12.[58] The data plots two pieces of information for four countries: Canada, Japan, Hong Kong, and France. The solid line in each figure represents the real per capita GDP of a country relative to that of the U.S. The dashed line in each figure represents the value of a country's currency relative to that of the U.S. dollar. The shaded regions in each figure (except for the Euro era for France) represent episodes in which the exchange rate was fixed either through a multilateral agreement (Bretton Woods) or through a unilateral arrangement (such as Hong Kong's Linked Exchange Rate System).

If there is anything that should strike you from this data, it is the following: the relative value of a country's currency appears in no way related to economic fundamentals (measured here as relative GDP). If we look at Canada, for example, we see that for the first half of the sample, the Canadian dollar averaged about 0.95 USD; in the last half of the sample, it averaged about 0.75 USD. For the entire sample period, however, Canada's real per capita GDP fluctuated randomly around 80% relative to U.S. real per capita GDP.

Take another look at Figure 8.12. Can you detect any systematic movement between the relative value of a country's currency and its relative performance?

[58] This data is from: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

I can't–so let's move on.

FIGURE 8.12
Real per capita GDP relative to U.S. and Nominal Exchange Rates
Selected Countries

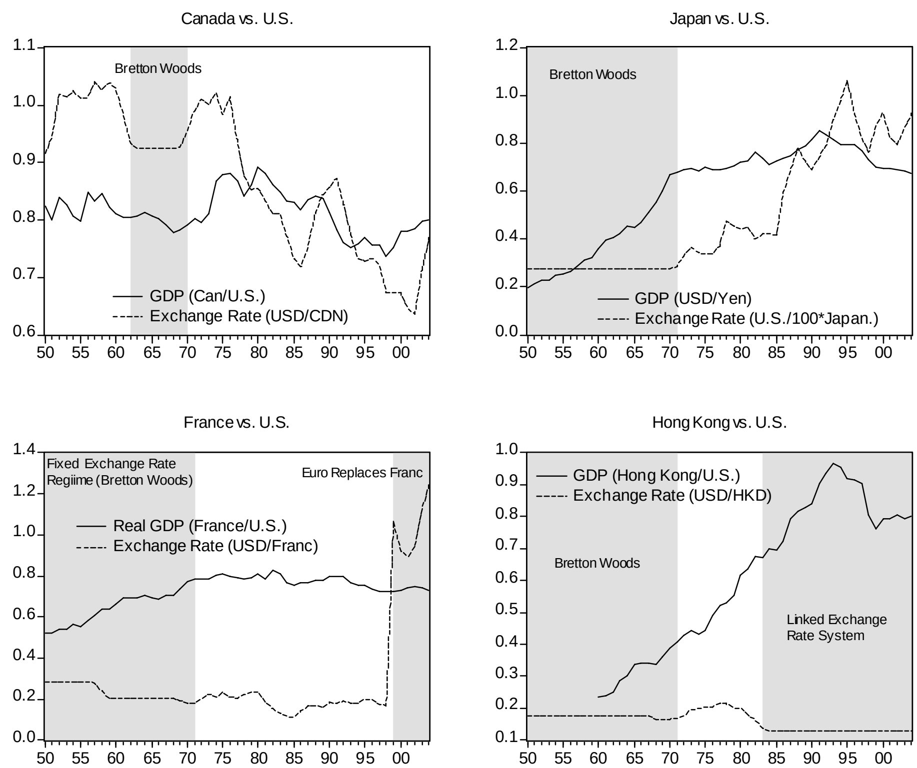


Figure 8.12 lends some credence to the view exposited by some that nominal exchange rates, if left to be determined in the market, frequently (but not always) appear to exhibit behavior that is detached from economic fundamentals. I used to be highly sceptical of this idea. But recently, I've become less sceptical. Let me explain why with the aid of a simple model.

5.1 Nominal Exchange Rate Determination

Let me start off with a very simple model. Agents are endowed with y units of output when young and nothing when old. These agents only value consumption when old. For simplicity, assume that agents have linear preferences; $u(c) = c$ (in this case, they are risk-neutral). The initial old are endowed with M dollars.

Since young agents do not value output when young, they will want to save it all. But as output is nonstorable, the only way they can do this is by selling their output (to the old) for money. Hence, each young person's demand for real money balances is simply given by $q = y$. Their budget constraint looks like this:

$$p_{t+1}c = p_t q;$$

or

$$c = \Pi^{-1} q.$$

The market-clearing condition is given by:

$$\frac{M}{p_t} = Nq.$$

Since $q = y$, it follows that:

$$p_t = \frac{M}{Ny};$$

so that the equilibrium inflation rate is $\Pi = 1$ (the young end up consuming y when old).

Now, imagine that there are two countries, a and b, and that each country issues its own currency, M^a and M^b. The two countries are otherwise identical (in particular, the produce the same output, although possibly at different levels, y^a and y^b). If the two countries operated in complete isolation from each other, then the price-level in each country would simply be given by:

$$p^a = \frac{M^a}{Ny^a} \text{ and } p^b = \frac{M^b}{Ny^b}. \tag{63}$$

So far, not very interesting. But now, imagine that the agents of different countries are allowed to trade with each other. Since the two goods are the same, the **real** exchange rate is fixed here at unity. While the output across countries is the same, they may still want to trade. For example, a young a agent may be willing to sell some of his output to an old b agent for his b money. In general, a young person may sell his output in exchange for some a money and some b money; and then carry these monies forward to the next period where they are used to purchase output (from agents of either country). Let q^a and q^b denote the real money balances carried forward to the next period by an a agent.

Now, imagine that–for some reason–young a and b agents holding foreign exchange did not want to carry it forward in time. A more realistic reason might be that their countries have imposed **foreign currency controls** that prevent them from doing so. But whatever the reason, the implication is that the young must dispose of their foreign money very soon after acquiring it. They can do so on a foreign exchange market (the young in each country would swap currencies). For a type a agent, this implies $q^a = y^a$ and $q^b = 0$ (and the opposite is true for a young b agent).

Let e_t denote the value of b money (euros) measured in units of a money (dollars). How is e_t determined here? Remember that the output of each country is the same. The value (purchasing power) of one dollar is $(1/p_t^a)$. That is, one dollar buys $(1/p_t^a)$ units of output. Similarly, one euro buys $(1/p_t^b)$ units of output. Therefore, the value of a euro relative to a dollar must satisfy:

$$e_t = \left(\frac{1/p_t^b}{1/p_t^a} \right) = \left(\frac{p_t^a}{p_t^b} \right). \tag{64}$$

Under foreign currency controls, we have $q^a = y^a$ for a agents and $q^b = y^b$ for type b agents. Therefore, the equilibrium price-levels in each country are determined by conditions (63). Substituting these into (64) we have:

$$e = \left(\frac{M^a}{M^b} \right) \left(\frac{y^b}{y^a} \right).$$

In other words, the equilibrium exchange rate is determined by fundamentals (relative money supplies and relative output). But I have already explained that this is not a very good theory.

In many countries, people are free to hold whatever currency they wish. When this is the case, the determination of the exchange rate becomes more problematic. To illustrate why, let me now reconsider what could happen in the absence of any foreign currency controls. To ease notation, let me also assume that $y^a = y^b = y$ (this is not important for the argument). Let me further assume that all agents (both a and b) hold identical money portfolios, so that $q^a + q^b = y$ for all young agents. In this case, each young agent faces the following budget constraint:

$$c = \Pi_a^{-1} q^a + \Pi_b^{-1} q^b.$$

Assuming that both currencies are willingly held by young agents, the two currencies must yield the same real rate of return. This implies that the **expected** (not realized) inflation rates across countries must be equal; i.e.,

$$\Pi_a = \Pi_b = \Pi. \tag{65}$$

In this case, the two currencies are viewed as **perfect substitutes** by young agents. What this means is that while each person has a determinate demand for real money balances equal to $q = y$, the actual **composition** $(q^a + q^b) = y$ of this money portfolio is **indeterminate** (agents do not care which money they hold–they both earn the same expected rate of return).

In the world described here, there is no separate market-clearing condition for each country. This is because the two countries are fully integrated (free trade in goods and monies). Thus, the relevant market-clearing condition is one that equates the world supply of money with the world demand for money; i.e.,

$$\frac{M^a}{p_t^a} + \frac{M^b}{p_t^a} = 2N\left(q^a + q^b\right).$$

Using condition (64) and the fact that $q^a + q^b = y$, this condition can be rewritten as:

$$M^a + e_t M^b = p_t^a 2Ny. \tag{66}$$

As this condition must hold at every date, the **realized** equilibrium inflation rate must satisfy:

$$\Pi_{t+1} = \frac{M^a + e_{t+1}M^b}{M^a + e_t M^b}.$$

There is something very curious about the equilibrium here. To see this, let me limit attention (for the moment) to an equilibrium where all variables remain constant over time. In this case, $\Pi_{t+1} = \Pi = 1$; and equation (66) becomes:

$$M^a + eM^b = p^a 2Ny.$$

Note that this market-clearing condition constitutes one equation in the two unknowns (e, p^a). The theory here places no other restrictions on these variables. Therefore, there are an infinite number of (e, p^a) combinations that are consistent with an equilibrium. To put things another way, the equilibrium nominal exchange rate is **indeterminate** (knowledge of fundamentals does not help pin down the exchange rate). Such an indeterminacy opens the door to speculative forces or "animal spirits." That is, the equilibrium nominal exchange rate is determined by whatever the foreign exchange market expects it to be (**self-fulfilling prophesies**).

It is also easy to see–although I will not demonstrate here–that there are equilibria in which (e_t, p_t^a) do not remain constant over time. In other words, this model is consistent with the idea that–if left to the foreign exchange market–nominal exchange rates might fluctuate over time for no apparent reason (or reasons unrelated to any underlying fundamentals).

5.2 Understanding Indeterminacy

The idea that a free market cannot determine the relative price of monies as a function of fundamentals sounds strange to economists. The market appears to do a good job of determining the relative price of apples and oranges, for example. When there is a frost in Florida, the supply of oranges contract, and oranges become more expensive relative to apples. Why does the same logic not apply to different monies?

The key to understanding this discrepancy is to recognize that government monies are not like other goods or assets. That is, other goods and assets have an underlying intrinsic value (people like to consume oranges; and assets produce a real flow of income). In contrast, the monies that governments issue today are **fiat** in nature. That is, they are intrinsically useless objects (and represent no legal claim against any intrinsically valuable object). Perhaps we should not be surprised that a free market is incapable of determining the "fundamental" relative price of two intrinsically useless objects.

Let me explain further by way of examples. Imagine that you sit down to play poker one night with your friends. Your host brings out the poker chips. These chips are intrinsically useless objects; they are simply distinguished by color (red, white, and blue). Your group must decide on an exchange rate system for poker chips. Are there any fundamental reasons to prefer one set of exchange rates over another? This seems unlikely; the group simply picks one out of many possibilities. Imagine further that once this exchange rate is determined, that your host discovers ten more blue chips that were hidden under the couch. Is there any reason to expect this discovery (an increase in the supply of blue chips) to depress their relative value?

Alright, let's consider another example. Imagine that the Bank of Canada prints two types of paper notes: blue (Lauriers) and green (Queens). Imagine that the Bank neglects to place any numbers on the notes, choosing instead to let the market determine their relative value. Are there any fundamental reasons for why the nominal exchange rate between these two notes should be equal to any particular number? How is the nominal exchange determined between Lauriers and Queens in reality? Simple: the Bank simply picks the exchange rate out of thin air (currently, the exchange rate is 4:1; that is, a Laurier is worth $5 and a Queen is worth $10). To put things another way, the government of Canada has put in place a system of **fixed exchange rates** between the notes that it issues.

Does anything in this logic change when we instead consider two notes issued by different countries? Imagine that Canada issues a blue Laurier and that the United States issues a grey Lincoln. Both of these notes have the number "5" printed on them; but this is like printing the number 5 on poker chips of different colors.

If fixing the nominal exchange rate between different monies within a nation makes sense, why does it also not make sense to fix the nominal exchange rate between monies across nations? Doing so would then determine the exchange rate (and possibly eliminate exchange rate fluctuations driven by speculation). In fact, why even bother with different national currencies? Why do nations simply not agree to adopt a single common currency?

5.3 Multilateral Fixed Exchange Rate Regimes

If nominal exchange rate fluctuations are driven primarily by speculation (exogenous shifts in expectations that are not based on fundamentals), then the resulting uncertainty is likely welfare-reducing. In principle, nominal exchange rate uncertainty would pose little concern if individuals held their wealth in a diversified portfolio consisting of assets denominated in all the world's currencies. But for various reasons, individuals do not appear to behave in this manner. Likewise, individuals could, in principle, try to hedge foreign exchange risk by purchasing insurance. In fact, many companies do behave in this manner. But many do not and, in any case, hedging is costly. Many economists believe that international trade would be facilitated (and welfare improved) in the absence of nominal exchange rate risk. One way to eliminate such risk is to enter into a multilateral agreement with other countries to fix the exchange rate. The Bretton-Woods arrangement (1946–1971) is a classic example of such an agreement.[59]

A multilateral fixed exchange rate regime sounds like a great idea–in principle. However, as you may have guessed from the collapse of the Bretton-Woods arrangement in 1971, fixed exchange rate regimes are not without their problems. One of the fundamental problems with maintaining a fixed exchange system is that it requires a high degree of coordination between countries in how fiscal and monetary policies are to be conducted. The basic idea is very easy; let me explain.

Consider our model above with two countries a and b. With free trade in goods and monies, the nominal exchange rate is indeterminate. Imagine then that the two countries agree to peg their currencies at par ($e = 1$). Then we are done, right? Well, not quite.

When the exchange rate is fixed, we see that the price-level is determined by (66); at par exchange, the price-level must be the same in both countries. Note that this equilibrium price-level is determined by the **total** supply of money $M^a + M^b$. What this means is that if one country decides to increase its money supply (say, to finance domestic government purchases), then the equilibrium value of **all** monies ($1/p$) must fall.[60] That is, by increasing (say) M^a, country a is in effect **exporting inflation** to country b. The resulting inflation tax is borne by all world citizens (even though the tax was used only to finance expenditures in one country).

Hence, we see that under a fixed exchange rate system, each domestic government has a strong incentive to expand their money supplies. Doing so allows them to "free-ride" off of other nations. This is arguably what led to the collapse of the Bretton Woods system. That is, as the war in Vietnam escalated in the late 1960s, the growing fiscal pressures applied to the U.S. government ultimately led to more rapid expansion in the supply of U.S. dollars. The resulting inflation was effectively exported to all countries belonging to the Bretton Woods arrangement. When other member countries finally got sick of this, they

[59] See: http://en.wikipedia.org/ wiki/ Bretton_Woods_Conference

[60] This holds true within a country as well. For example, if the Bank of Canada expands the Canadian money supply by printing more Queen's, then the value (purchasing power) of all Canadian monies (including Lauriers) must fall in proportion.

withdrew from the system.

For a fixed exchange rate system to work well across countries, the governments of all countries involved must somehow agree to discipline their monetary and fiscal policies. Needless to say, this is easier said than done.

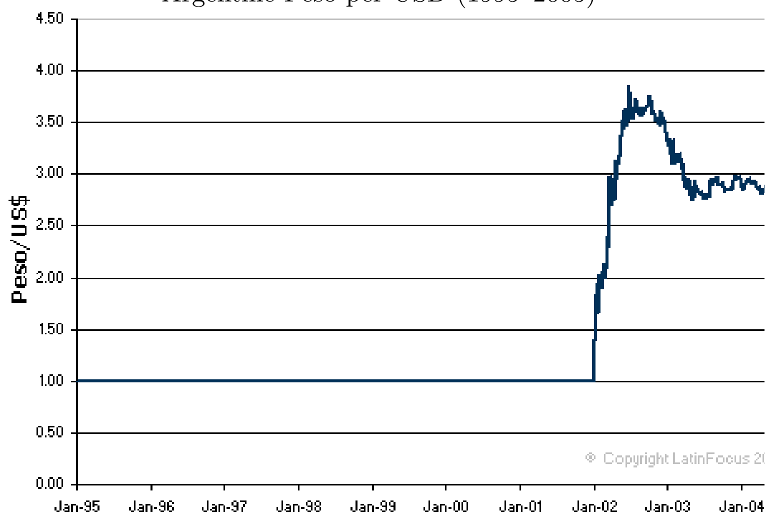
5.4 Speculative Attacks

The key to maintaining a fixed exchange rate system is obtaining a **credible commitment** on the part of all member governments to exchange different monies at the stated rates. Such a commitment is necessary to defend the exchange rate system against speculative attacks. To see why, consider the following scenario. Imagine that Canada and the U.S. have entered into a bilateral fixed exchange rate agreement. Now, imagine that participants in the foreign exchange (FX) market 'speculate' that the Canadian dollar may depreciate. Such a speculation may lead market participants to 'dump' their Canadian dollars on the FX market (in exchange for U.S. dollars). But **if** the U.S. government stands ready to print all the U.S. dollars demanded by speculators (in exchange for Canadian dollars) at the stated rate, then the speculative attack **must** fail. Understanding that this must be the case, there is no reason to engage in speculative activity in the first place.

However, many countries attempt to fix their exchange rate **unilaterally**, often by way of a currency board. For example, Argentina adopted a currency board from April 1, 1991 through January 6, 2002 with the stated intent of fixing the value of its Peso to the U.S. dollar at par. Defending one's exchange rate against speculative attacks is more difficult to do unilaterally than it is via a bilateral agreement. This is because the commitment to defend the exchange rate must rest solely on the country imposing the peg. In particular, the United States did not promise to help Argentina defend the Peso in the event of a speculative attack. To defend its currency unilaterally, Argentina had to convince FX participants that it stood ready to do whatever it took to maintain the exchange rate. One way to do this is for the currency board to hold one U.S. dollar in reserve for every Peso it prints (this reserve currency must ultimately be acquired via taxation, if the Fed has no desire to hold Pesos). Alternatively, the Argentine government must stand willing to tax its citizens to acquire the U.S. dollars it needs to meet the demands of any speculators. More importantly, FX participants must *believe* that the Argentine government would be willing to take such an action; i.e., the stated policy must be perceived to be **credible**.

Figure 8.13 plots the exchange rate between the Argentine Peso and the USD for 1995-2005. For a period of time, the Argentine currency board appeared to work well, at least, in terms of maintaining a fixed rate of exchange (par) with the USD. However, for a variety of reasons, the currency board was compelled to abandon its peg against the USD in January 2002. Figure 11.2 shows that following this abandonment, the Argentine Peso devalued sharply and is presently worth around \$0.34 USD (about one-third of its former value).

FIGURE 8.13
Argentine Peso per USD (1995–2005)



What went wrong in Argentina? According to my Argentine friends, nothing went wrong–what happened was perfectly normal (which is to say that everything is always going wrong in Argentina). Some people place the 'blame' on the U.S. dollar, which strengthened relative to most currencies over the 1990s. Since the Peso was linked to the U.S. dollar, this had the effect of strengthening the Peso as well, which allegedly had the effect of making Argentina's exports uncompetitive on world markets. While there may be an element of truth to this argument, one wonders how the U.S. economy managed to cope with the rising value of its currency over the same period (in which the U.S. economy boomed). Likewise, if the rising U.S. dollar made Argentine exports less competitive, what prevented Argentine exporters from cutting their prices?

A more plausible explanation may be the following. First, the charter governing Argentina's currency board did not require that Pesos be fully backed by USD. Initially, as much as one-third of Pesos issued could be backed by Argentine government bonds (which are simply claims to future Pesos). In the event of a major speculative attack, the currency board would not have enough USD reserves to defend the exchange rate. Furthermore, it would likely have been viewed as implausible to expect the Argentine government to tax its citizens to make up for any shortfall in reserves. Second, a combination of a weak economy and liberal government spending led to massive budget deficits in the late 1990s. The climbing deficit led to an increase in devaluation concerns. According to

Spiegel (2002), roughly \$20 billion in capital 'fled' the country in 2001.[61] Market participants were clearly worried about the government's ability to finance its growing debt position without resorting to an inflation tax (Peso interest rates climbed to between 40-60% at this time). In an attempt to stem the outflow of capital, the government froze bank deposits, which precipitated a financial crisis. Finally, the government simply gave up any pretense concerning its willingness and/or ability to defend the exchange rate. Of course, this simply served to confirm market speculation.

At the end of the day, the currency board was simply not structured in a way that would allow it to make good on its promise to redeem Pesos for USD at par. In the absence of full credibility, a unilateral exchange rate peg is an inviting target for currency speculators.

Exercise 8.6: Explain why speculating against a currency that is pegged unilaterally to a major currency like the USD is close to a 'no-lose' betting situation. Hint: explain what a speculator is likely to lose/gain in either scenario: (a) a speculative attack fails to materialize; and (b) a speculative attack that succeeds in devaluing the currency.

5.5 Currency Union

A currency union is very much like a multilateral fixed exchange rate regime. That is, different monies with fixed nominal exchange rates essentially constitute a single money. The only substantive difference is that in a currency union, the control of the money supply is taken out of the hands of individual member countries and relegated to a central authority. The central bank of the European Currency Union (ECU), for example, is located in Frankfurt, Germany, and is called the European Central Bank (ECB). The ECB is governed by a board of directors, headed by a president and consisting of the board of directors and representatives of other central banks in the ECU. These other central banks now behave more like the regional offices of the Federal Reserve system in the United States (i.e., they no longer exert independent influence on domestic monetary policy).

Having a centralized monetary authority is a good way to mitigate the lack of coordination in domestic monetary policies that may potentially afflict a multilateral fixed exchange rate system. However, as the recent European experience reveals, such a system is not free of political pressure. In particular, ECB members often feel that the central authority neglects the 'special' concerns of their respective countries. There is also the issue of how much seigniorage revenue to collect and distribute among member states. The governments of member

[61] I presume what this means is that Argentines flocked to dispose of \$20 billion in Peso-denominated assets, using the proceeds to purchase foreign (primarily U.S.) assets.

countries may have an incentive to issue large amounts of nominal government debt and then lobby the ECB for high inflation to reduce the domestic tax burden (spreading the tax burden across member countries). The success of a currency union depends largely on the ability of the central authority to deal with a variety of competing political interests. This is why a currency union within a country is likely to be more successful than a currency union consisting of different nations (the difference, however, is only a matter of degree).

5.6 Dollarization

One way to eliminate nominal exchange rate risk that may exist with a major trading partner is to simply adopt the currency of your partner. As mentioned earlier, this is a policy that has been adopted by Panama, which has adopted the U.S. dollar as its primary medium of exchange.

One of the obvious implications of adopting the currency of foreign country is that the domestic country loses all control of its monetary policy. Depending on circumstances, this may be viewed as either a good or bad thing. It is likely a good thing if the government of the domestic country cannot be trusted to maintain a 'sound' monetary policy. Any loss in seigniorage revenue may be more than offset by the gains associated with a stable currency and no exchange rate risk. On the other hand, should the foreign government find itself in a fiscal crisis, the value of the foreign currency may fall precipitously through an unexpected inflation. In such an event, the domestic country would in effect be helping the foreign government resolve its fiscal crisis (through an inflation tax).

Exercise 8.7: If the Argentine government had simply dollarized instead of erecting a currency board, would a financial crisis have been averted? Discuss.

6. Summary

Money is an object that serves as a means of payment. Most economies use both government fiat currency and private debt instruments as a means of payment. The use of money in transactions allows trade to occur where it might otherwise not.

Governments that control the supply of fiat currency can potential use money creation as a revenue device. Such a seigniorage policy results in inflation, which acts as a tax on all activities that require monetary exchange. That is, inflation reduces the real rate of return on money, and distorts the relative price of cash and non-cash goods. But because the demand for real money balances is likely to contract when inflation rises, there are limits to how much seigniorage revenue a government can extract from the economy.

Private agencies like chartered banks created debt instruments (demand deposits) that are ultimately backed by the capital they invest in. An expansion in the demand for investment will expand the supply of private money necessary to finance investment expenditure. The result may be changes in the broad money supply (M1) that are positively correlated with future changes in real GDP. The direction of causality here likely works in reverse.

The nominal exchange rate measures the relative price of two currencies. If these currencies are fiat in nature, and if people are free to trade in goods and monies, then there are no fundamental determinants of the nominal exchange rate. In this case, the nominal exchange rate is determined entirely by speculation. When this is the case, there are benefits associated with fixing the exchange rate or adopting a common currency. There are costs associated with such regimes, however, when international governments cannot agree to coordinate their monetary and fiscal policies.

Appendix 8.1
The Asian Financial Crisis

Perhaps you've heard of the so-called *Asian Tigers*. This term was originally applied to the economies of Hong Kong, South Korea, Singapore and Taiwan, all of which displayed dramatic rates of economic growth from the early 1960s to the 1990s. In the 1990s, other southeast Asian economies began to grow very rapidly as well; in particular, Thailand, Malaysia, Indonesia and the Philippines. These 'emerging markets' were subsequently added to the list of Asian Tiger economies. In 1997, this impressive growth performance came to a sudden end in what has subsequently been called the Asian Financial Crisis. What was this all about?

Throughout the early 1990s, many small southeast Asian economies attracted huge amounts of foreign capital, leading to huge net capital inflows or, equivalently, to huge current account deficits. In other words, these Asian economies were borrowing resources from the rest of the world. Most of these resources were used to finance domestic capital expenditure. As we learned in Chapters 5 and 7, a growing current account deficit may signal the strength of an economy's future prospects. Foreign investors were forecasting high future returns on the capital being constructed in this part of the world. This 'optimism' is what fuelled much of the growth domestic capital expenditure, capital inflows, and general growth in these economies. Evidently, this optimistic outlook turned out (after the fact) to be misplaced.

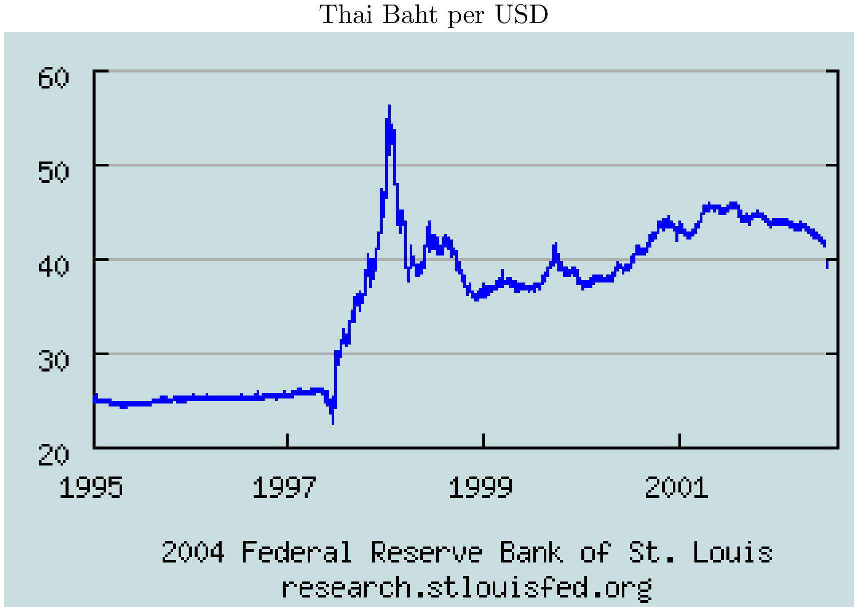
What went wrong? One possible is that nothing went 'wrong' necessarily. After all, rational forecasts can (and often do) turn out to be incorrect (after the fact). Perhaps what happened was a growing realization among foreign investors that the high returns they were expecting were not likely to be realized. Investors who realized this early on pulled out (liquidating their foreign asset holdings). As this realization spread throughout the world, the initial trickle in capital outflows exploded into a flood. Things like this happen in the process of economic development.

Of course, there are those who claim that the 'optimism' displayed by the parties involved was 'excessive' or 'speculative;' and that these types of booms and crashes are what one should expect from a free market. There is another view, however, that directs the blame toward domestic government policies. For example, if a government stands ready to bailout domestic losers (bad capital projects), then 'overinvestment' may be the result as private investors natural downplay the downside risk in any capital investment. To the extent that foreign creditors are willing to lend to domestic agents against future bail-out revenue from the government, unprofitable projects and cash shortfalls are refinanced through external borrowing. While public deficits need not be high before a crisis, the eventual refusal of foreign creditors to refinance the country's cumulative losses forces the government to step in and guarantee the outstanding stock of external liabilities. To satisfy solvency, the government must then undertake appropriate domestic fiscal reforms, possibly involving recourse to seigniorage revenues. Expectations of inflationary financing thus cause a collapse of the currency and anticipate the event of a financial crisis.

There is also evidence that government corruption may have played a significant role. For example, a domestic government may borrow money from foreigners with the stated intent of constructing domestic capital infrastructure. But a significant fraction of these resources may simply be 'consumed' by government officials (and their friends). For example, in 2001 Prime Minister Thaksin (of Thailand) was indicted for concealing huge assets when he was Deputy Prime Minister in 1997. Evidently, Mr. Thaksin did not dispute the charge. Instead, he said that the tax rules and regulations were 'confusing' and that he made an 'honest mistake' in concealing millions of dollars assets, manipulating stocks and evading taxes.[62]

The Asian crisis began in 1997 with a huge speculative attack on the Thai currency (called the *Baht*). Prior to 1997, the Thai government had unilaterally pegged their currency at around 25 Baht per USD. Many commentators have blamed this speculative attack for precipitating the Asian crisis. A more plausible explanation, however, is that the speculative attack was more of a **symptom** than a cause of the crisis (which was more deeply rooted in the nature of government policy).

Thai Baht per USD



In any case, financial crisis or not, our theory suggests that the Thai government could have maintained its fixed exchange rate policy and prevented

[62] See: www.pacom.mil/ publications/ apeu02/ 24Thailand11f.doc.

a speculative attack on its currency if it had either: (1) maintained sufficient USD reserves; or (2) been willing to tax its citizens to raise the necessary USD reserves. Evidently, as the Thai economy showed signs of weakening in 1997, currency speculators believed that neither of these conditions held (and in fact, they did not).

Would the crisis in Thailand have been averted if the government had maintained a stable exchange rate? It is highly doubtful (in my view) that this would have been the case. If the crisis was indeed rooted in the fact that many bad investments were made (the result of either bad decisions or corruption), then the contraction in capital spending (and the corresponding capital outflows) would have occurred whether the exchange rate was fixed or not.

Problems

1. If money creation is a tax, then why do governments simply not tax people directly, rather than by depreciating the value of their money? That is, list the circumstances under which an inflation tax might be preferable to an income tax.

2. Political elements in the Canadian province of Quebec regularly express their desire to separate Quebec from the rest of Canada. They came very close to doing so in a 1995 referendum (49.4% voted yes; and 50.6% voted no). In the run up to the vote, Quebec separatists claimed that they would continue to use the Canadian dollar following separation. This claim generated quite an uproar in the rest of Canada. Some went so far as to suggest that Canada should refuse to allow a separate Quebec to use the Canadian dollar. Can you explain to me two things: (a) Precisely how might it be possible to prevent foreigners from using our currency; and (b) Why on earth would we ever want to do this, even if we could?

3. In the mid 1970s, the Canadian dollar was trading at close to par with the U.S. dollar. For the next twenty years, the Canadian dollar depreciated steadily in value, reaching a low of around 62 cents in the early 1990s (it has recently moved back to par again). During this period of depreciation, many economists were arguing that Canada "Dollarize;" by which they meant that Canada should abandon its dollar in favor of the U.S. dollar. Explain the likely costs and benefits of such a policy.

CHAPTER 9

Economic Growth and Development

1. World Income and Population Dynamics

The earliest known civilization–the Sumerians–appeared in Iraq about 6,000 years ago. We've come a long way since then. According to one estimate, the average real per capita income in the world today is approximately 65 times higher than it was 6,000 years ago. According to the same source, the world's population today is almost 900 times higher than it was in 4000 B.C.[63]

One might be inclined to think that this pattern of development occurred on a more or less steady basis for most of human history. In fact, almost all of the gains in income and population have occurred in just the last 200 years. The average level of material living standards in 1700 A.D. was not that much different than it was thousands of years ago. On the other hand, the world's population tended to grow at a relatively stable pace; roughly doubling in size every 1000 years up to about the year 1600 A.D. From 1600 to 1800–a mere two hundred years–the population doubled again (to almost 1 billion). It only took another 100 years for the population to double again. Over the last 100 years, the world's population quadrupled (to over 6 billion).

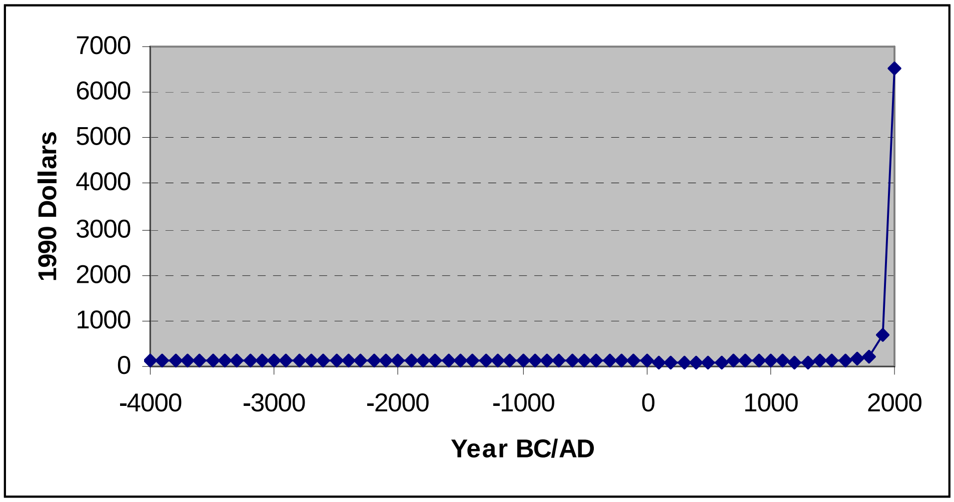
Thus, for almost all of recorded history, the general pattern of development was characterized by either stagnant to modest gains in material living standards and a steadily increasing world population. About 200 years ago, material living standards began to grow; and about 400 years ago, the world's population began to grow at an accelerating rate; see Figure 9.1.

Why did average world living standards change so little over so long a period of time? Most people are likely to guess that the primary reason for this is because of an absence of technological progress. But even a casual glance at the evidence suggests that this cannot be the case. According to the historian Joel Mokyr, there is much evidence of technological progress throughout the ages.[64] For example, the Greeks of classical antiquity made several advances in mathematics, medicine, science, architecture, construction, and political organization. Rome in 100 A.D. had better paved streets, sewage, water supply, and fire protection than the capitals of Europe in 1800 A.D. The middle ages witnessed the introduction of the heavy plow, the wheel barrow, the three-field system, windmills, the weight-driven mechanical clock and the printing press (to name just a few).

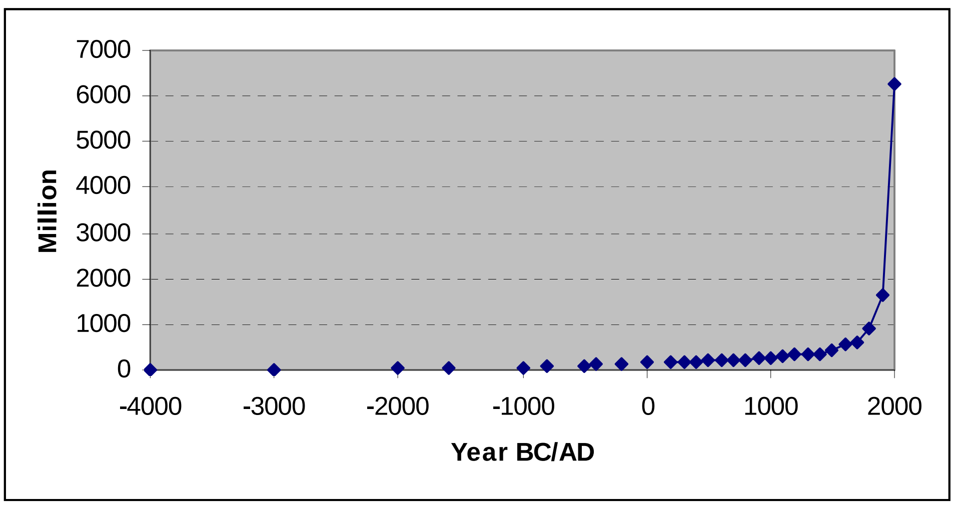
[63] From J. Bradford DeLong's webpage:
http://econ161.berkeley.edu/TCEH/1998_Draft/World_GDP/Estimating_World_GDP.html

[64] Mokyr, Joel (1990). *The Levers of Riches: Technological Creativity and Economic Progress*, Oxford University Press, New York.

FIGURE 9.1
Average World Real GDP per Capita



World Population



In our current age of iphones and PCs, perhaps you are not so impressed with an invention like the wheel barrow. But imagine that prior to this invention, the best technology for clearing a field of stones involved two men with a stretcher. With a wheel barrow, however, only one man would be required to move the same quantity of stones. Two men doing the same job with two wheelbarrows would then double the daily production. We are talking here about a 100% improvement in the productivity of labor!

Hence, whatever the reason for little or no change in material living standards over thousands of years, the answer cannot simply be a lack of technological progress. Something else must be at work to keep living standards low despite improvements in technology. One possible explanation might be related to population growth–an hypothesis first touched upon by Thomas Malthus (1766-1834).

2. A Malthusian Model

Given that Malthus died in 1834, we should keep in mind that his view of the history of economic development would not have included the last little bit in Figure 9.1. From Malthus' perspective, the material living standards of the average person (typically a peasant farmer) would have appeared to have remained unchanged forever. At the same time, he would have noted the phenomenon of an increasing population.

The way Malthus appears to have explained this phenomenon is as follows. Consider an economy that produces food using land and labor (agriculture being the most important industry in his day). Land is relatively fixed in supply. But the supply of labor expands with the size of the population. More people working the same amount of land implies less output/income per person. Most of the output is consumed (rather than saved). This latter assumption reflects the idea that most people were living at close to subsistence levels. Moreover, even if the average peasant did want to save, the means of saving were largely absent (owing to a lack of well-developed financial markets).

But why does the population change over time? The net population growth rate is equal to the difference in the birth rate and the mortality rate. According to Malthus, the birth rate was largely insensitive to economic conditions; being determined primarily by a hunger for sex. On the other hand, the mortality rate did appear to be sensitive to economic conditions. Mortality rates appear to be inversely related to the average level of income. More income, for example, means that children are better fed and more resistant to disease.

This simple theory implies that absent any technological improvement, the per capita income and population will converge to a steady-state. That is, income will converge to some subsistence level and the population will remain constant over time (so that the birth rate equals the mortality rate).

Imagine now the appearance of the wheel barrow. The immediate impact is to increase productivity and hence increase per capita income. But as per capita income rises, the mortality rate falls–so that the population expands. As the population expands, per capita income falls, and the mortality rate begins to rise. Incomes continue to fall and mortality rates continue to rise until, in the long-run, the new steady-state is characterized by a larger population–all of whom are living at the original subsistence level.

3. Children as an Investment Good

Malthus' interpretation of the pattern of development no doubt contains an element of truth. His theory, however, treats people as behaving rather robotically. In particular, the people in his model simply consume all their income and mindlessly create children. Perhaps some people did (and do) behave in this manner. But there is evidence to suggest that fertility is a rational choice, even among those that live in lesser developed economies. A 1984 World Bank report puts it this way:[65]

> All parents everywhere get pleasure from children. But children involve economic costs; parents have to spend time and money bringing them up. Children are also a form of investment–providing short-term benefits if they work during childhood, long-term benefits if they support parents in old age. There are several good reasons why, for poor parents, the economic costs of children are low, the economic (and other) benefits of children are high, and having many children makes economic sense.

I want to focus on this idea of children as constituting a form of investment. While the concept may seem odd to those of us who live in well-developed economies, it is not nearly so odd when we think of the savings opportunities available to people in primitive economies. Let us try to imagine then a very primitive economy with a complete absence of savings opportunities. The only asset in the economy is a fixed amount of land L. Assume that there are no private property rights in this asset; that is, assume that land is **common property**.

Time is discrete and denoted by $t = 1, 2, ..., \infty$. Let N_t denote population size at date t (and assume that the labor input is proportional to population). Land and labor together produce an aggregate level of real GDP equal to $Y_t = zF(L, N_t)$, where $z > 0$ is an exogenous productivity parameter. Assume that $F(L, N)$ is increasing and strictly concave in both L and N; and that F exhibits constant returns to scale (CRS).[66] A function F displays CRS if for any number $\lambda > 0$,

$$F(\lambda L, \lambda N) = \lambda Y.$$

Imagine, for example, that $\lambda = 2$. Then CRS implies that if we double the quantity of land and labor, then the level of output will double as well.

Imagine instead that $\lambda = 1/N$. Then CRS implies:

$$\frac{Y}{N} = F\left(\frac{L}{N}, 1\right).$$

[65] Quoted from: Razin, Assaf and Efraim Sadka (1995). *Population Economics*, The MIT Press, Cambridge, Massachussetts.

[66] Mathematically, one would say that F is linearly homogeneous in L and N.

The left-hand-side of this expression is just per capita income; which I denote here by $y \equiv Y/N$. Under CRS, per capita income is an increasing function of the land-labor ratio. Since land is assumed to be fixed in supply, it follows that an increase in population will lead to a decline in per capita income. Hence, the function $f(N) \equiv F(L/N, 1)$ is a decreasing function of N; and we can write the relationship conveniently as:

$$y = zf(N).$$

Now, I want to consider an economy that literally (rather than metaphorically– as in Chapter 8) consists of overlapping generations of people. For simplicity, assume that people live only for two periods, so that they are either young or old. The population of young parents at date t is given by N_t. The young work the land, earning income $y_t = zf(N_t)$. The old do not work; their consumption must be provided to them (if at all) by their children. As there are no other means of providing for old age, the number of children here will constitute a form of investment. The family planning decision (the number of children to have) is a choice that must be made when young.

Assume that the young value consumption when young and old. Let c_t^y denote consumption when young at date t; and let c_{t+1}^o denote consumption when old at date $t+1$. Then preferences can be written as $u\left(c_t^y, c_{t+1}^o\right)$ with an associated $MRS(c_t^y, c_{t+1}^o)$. There is also an "initial old" generation that lives for one period only; their utility function is simply given by c_1^o.

I depart from Malthus in assuming that young families can choose their desired family size. Let n_t denote the number of children chosen by a young family. At the individual level, the number of children must be a non-negative integer $\{0, 1, 2, ...\}$. But I am going to assume here (for simplicity only) that is possible to have fractions of children. That is, it is possible to have, for example, 0.8 or 1.2 children. (Do not get too hung up on this; the basic idea will remain intact if I restrict individual family size to integer numbers).

Given these assumptions, the average number of children per young family is given by n_t. For simplicity, assume that all children survive into adulthood. In this case, the future population of young families is given by:

$$N_{t+1} = n_t N_t.$$

Raising children is costly. Assume that each child requires one unit of output.

There is one important remaining detail to consider. In particular, what is it that motivates the young to transfer some of their income to their parents? And if they are motivated to transfer some of their income in this manner, exactly how much should they hand over?

One might be inclined to think that these considerations are dictated by feelings of love, duty, and fairness. But there may be an economic rationale for transferring income to parents as well. For example, imagine that as a child, you see that your parents do nothing to support your grandparents (their parents), even though they clearly have the capacity to do so. Then I assume that you will treat your parents in a **tit-for-tat** manner. That is, children will treat their parents in exactly the same way that they treated their parents. If people behave in this way, then the young will be motivated to transfer income to their parents. Doing so is the only way to ensure that their "gift" is reciprocated by their own children in the future (the alternative is starvation).

The question remains as to how much the young should transfer to their parents. Here, I make the simplifying assumption that the young transfer a constant fraction $0 < \theta < 1$ of their income to their parents. In this case, the future consumption of parents is dictated primarily by the number of children they choose to have (rather than the size of the gift they made to their parents). Thus, the budget constraints for a young family are given by:

$$\begin{aligned} c_t^y + n_t &= (1-\theta)y_t; \\ c_{t+1}^0 &= n_t \theta y_{t+1}. \end{aligned}$$

The initial old, of course, simply consume $c_1^o = n_0 \theta y_1$.

The two constraints above can, as usual, be combined to form a single lifetime budget constraint:

$$c_t^y + \frac{c_{t+1}^o}{\theta y_{t+1}} = (1-\theta)y_t.$$

The right-hand-side denotes the young family's disposable income (the income left over after they transfer resources to their parents). Note that the term θy_{t+1} acts just like an interest rate. That is, it is the rate of return (in terms of additional future consumption) that is related to the investment in children. For a given (y_t, y_{t+1})–which is beyond the control of any individual family–the solution to the choice problem can be depicted by a point like A in Figure 9.2.

Exercise 9.1: Consider Figure 9.2. Suppose that a young family's income y_t increases exogenously. How does this "income shock" affect the family's fertility choice? Explain. (Hint: consumption smoothing).

Exercise 9.2: Consider Figure 9.2 again. Now imagine that a young family receives information that leads it to revise upward its forecast of the income that will be earned by its children y_{t+1}. How is this "news shock" likely to influence fertility choice? (Hint: there will be substitution and wealth effects).

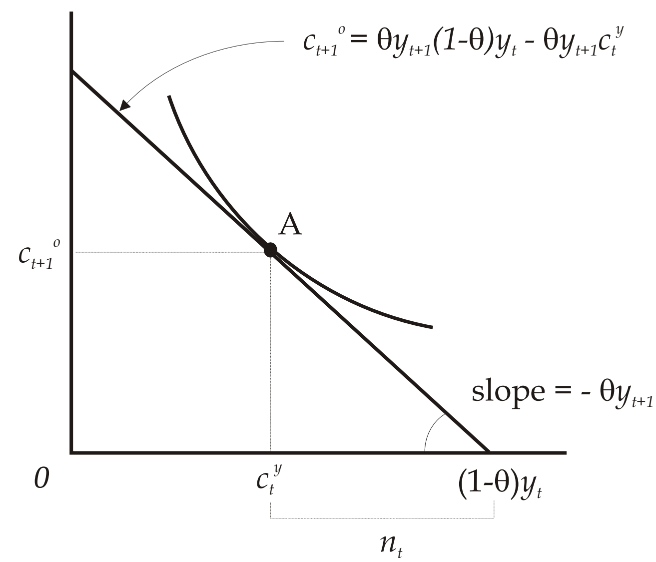
To explore the implications of this model further, let me assume that preferences take a particular form; i.e.,

$$u(c_t^y, c_{t+1}^o) = \ln(c_t^y) + \beta \ln(c_{t+1}^o). \tag{67}$$

In this case, we have:

$$MRS(c_t^y, c_{t+1}^o) = \frac{1}{\beta}\left(\frac{c_{t+1}^o}{c_t^y}\right). \tag{68}$$

FIGURE 9.2
Optimal Family Size



We know that point A in Figure 9.2 is described by the following two conditions:

$$\begin{aligned} \frac{1}{\beta}\left(\frac{c_{t+1}^o}{c_t^y}\right) &= \theta y_{t+1}; \\ c_t^y + \frac{c_{t+1}^o}{\theta y_{t+1}} &= (1-\theta)y_t. \end{aligned}$$

Exercise 9.3: Using the two conditions above, solve for the desired consumption pattern (c_t^y, c_{t+1}^o). Next, using the fact that $n_t = (1-\theta)y_t - c_t^y$, solve for the desired family size variable n_t and show that:

$$n_t = \sigma y_t; \tag{69}$$

where $\sigma = \beta/(1+\beta)$ is the saving rate. Explain why desired family size here does not depend on θ or y_{t+1} (refer to Exercise 9.2).

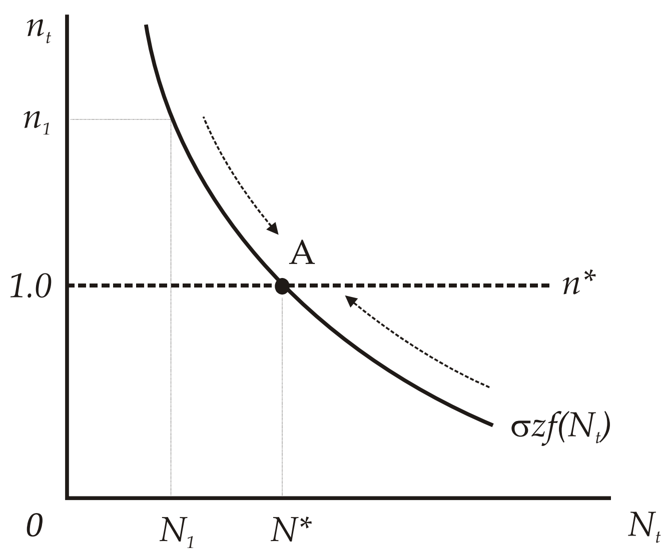
Describing the equilibrium dynamics of this economy will now be easy. In particular, note that $y_t = zf(N_t)$. Combining this with the fertility choice variable (69), we can determine the equilibrium fertility choice as a function of N_t and z; i.e.,

$$n_t = \sigma z f(N_t).$$

This equation implies that as population size expands, the population growth rate will decline. The reason for this is because as population expands, real per capita income declines. The decline in wealth then reduces the demand for all normal goods (both current and future consumption). The decline in the demand for future consumption implies a smaller desired family size (as children are the way in which future consumption is attained).

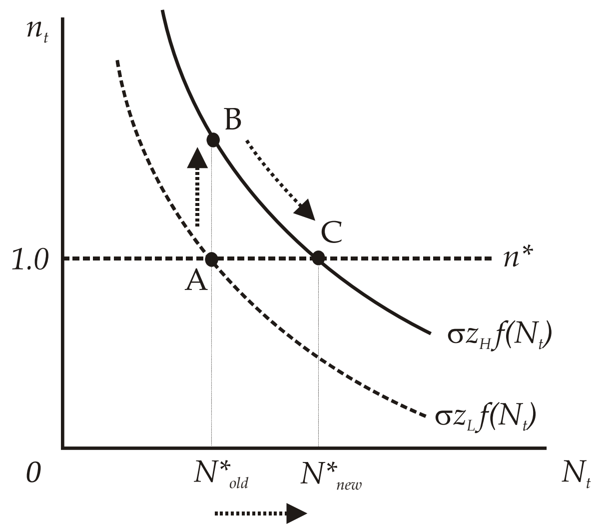
Now to describe the dynamics, assume that the population at date t is such that $n_t > 1$. Then the population will expand and incomes will fall. Falling incomes will reduce the population growth rate. This process will continue until a steady-state is reached at which the population remains constant; $n^* = 1$. Imagine instead that the population at date t is such that $n_t < 1$. Then the population will shrink and incomes will rise. Rising incomes will increase the population growth rate. This process will again continue until a steady-state $n^* = 1$ is reached; see Figure 9.3.

FIGURE 9.3
Equilibrium Population Dynamics



Now, imagine that the economy begins in a steady-state, as described by point A in Figure 9.3. Imagine further that some smart person invents the wheel barrow, and that this invention spreads very quickly (within the date of its invention). The effect of the invention is to increase the productivity of labor from z_L to z_H. What are the resulting population and income dynamics? Figure 9.4 illustrates the effects.

FIGURE 9.4
Population Dynamics Following
a Technological Advancement



The initial steady-state is depicted by point A in Figure 9.4. Associated with this steady-state is a population of size N^*_{old}. Associated with this population is per capita income $y^*_{old} = z_L f(N^*_{old})$. The effect of a technological improvement is to shift the $\sigma z f(N)$ function "up." That is, the immediate impact is to increase the level of income for young households from $z_L f(N^*_{old})$ to $z_H f(N^*_{old})$. This higher income implies an increase in the fertility rate (point B); as families rationally plan to smooth this income shock over their lifetimes. But as all young families make the same calculation, the effect is to make the population grow. As the population expands over time, productivity and income declines. The steady decline in income eventually reduces the population growth rate back to zero (point B to point C). At point C–the new steady-state, the population is now higher at N^*_{new}. Associated with this higher long-run population, however, is exactly the **same** level of per capita income; i.e.,

$$y^*_{new} = z_H f(N^*_{new}) = z_L f(N^*_{old}) = y^*_{old}.$$

Hence, the effect of a **permanent** technological improvement is to increase living standards only **temporarily**. In the long-run, the higher level of productivity is entirely dissipated by a higher population. The basic conclusion here is entirely consistent with the Malthusian growth model. The only difference is that here, I have shown the same result occurring through individually rational fertility choices; while Malthus stressed the role of an economically sensitive mortality rate.

Exercise 9.4: In an attempt to curb the growth rate of its huge population, Chinese authorities implemented a one-child policy in 1979 (a legal restriction on the number of children allowed per household). Use the theory developed here to rationalize the adoption of such a policy. Explain why such a policy is likely to improve the welfare of future generations, at the expense of reducing the welfare of current generations (relative to the welfare they could achieve in the absence of any legal restriction).

4. Asset Markets and Fertility

The Malthusian models studied here provide plausible interpretations of why, in the face of continued technological advancements over time, the material living standards of the average person did not change very much over long periods of time. Of course, the Malthus model does not explain why living standards (and population) began to grow rapidly in recent history. Average living standards began to grow rapidly around two hundred years ago; and the population began to grow more rapidly even before this. Evidently, something in the world–or at least, in some parts of it–changed. The question is what changed (and why)?

Given that these changes occurred somewhere around the so-called "Industrial Revolution," some may be inclined to explain recent history simply by appealing to "technological progress." Certainly, there must be an element of truth to this, as technological progress does appear to be advancing at a more rapid pace since 1800. But as I have argued above, we cannot appeal solely to the effects of productivity gains to explain the facts. Prior to 1800, productivity improvements ultimately led to higher populations and relatively stable long-run living standards. Why has not the same thing happened in recent history?

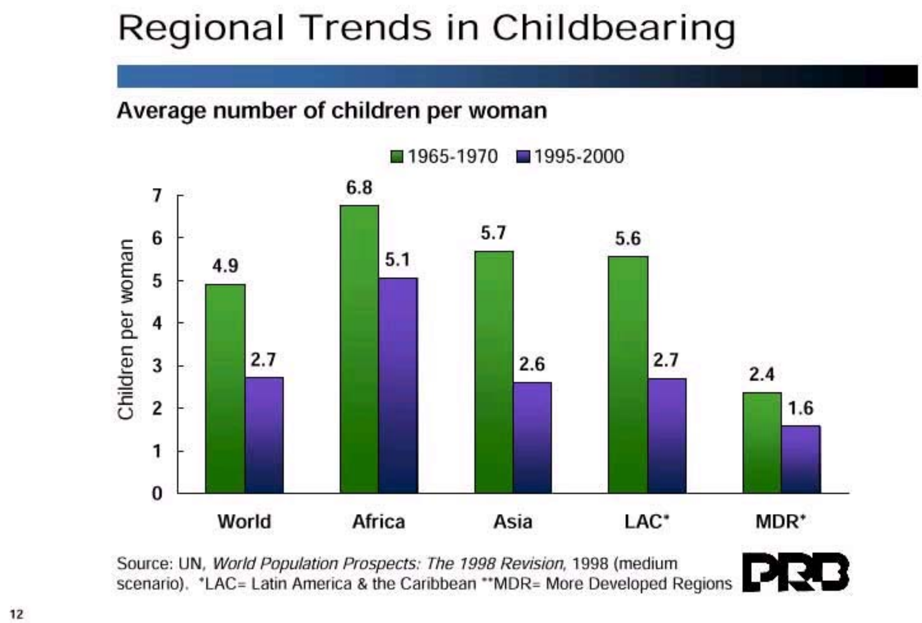
In my version of the Malthus model, an increase in per capita income leads to an increase in the birth rate. And yet, when we take a look at recent fertility trends, it appears that the opposite is true. We know, for example, that world fertility rates have been falling over time (as per capita incomes have been rising; and that, at any given point in time, poor countries tend to have higher fertility rates than rich countries; see Figure 9.5.[67]

One possible explanation for the decline in the fertility rate is a largely corresponding decline in the mortality rate, brought about by advances in health technology. But even when one corrects for this, it appears that the net popu-

[67] www.uwsp.edu/business/economicswisconsin/e_lecture/pop_sum.htm

lation growth rate is negatively correlated with per capita income.

FIGURE 9.5



Thus, while rich countries are obviously more productive than poor countries, the key to translating higher productivity into higher material living standards (and not hugely larger populations) appears to be related to the fact that people in rich countries choose to have smaller families. The question is why?

One hypothesis is that rich countries have adopted different institutions; in particular, institutions that define and enforce **private property rights** in personal asset holdings. Let me explain.

In the primitive Malthusian world described earlier, I assumed that land was common property (not private property). This assumption precluded individuals from saving; in particular, by purchasing titles to land. More generally, we can think of primitive economies as being characterized by the absence of any well-developed financial market. An important modern day financial institution is the retail banking sector, which allows people with even modest means to accumulate savings that are protected by the law. The emergence of retail banking was likely only made possible (or at least greatly facilitated) by the emergence of a legal institution that protected the deposits of individuals from theft (from other individuals or even the banker).

The idea then is as follows. In a world where saving through asset accumulation is difficult (or impossible), households will rationally choose to save by investing in children. If the institutional environment changes in a way that makes saving through asset accumulation easier or more attractive, then households may be induced to divert their saving into assets and away from children. In a world with highly developed asset markets, parents may still choose to have children–but they will do so for reasons that are entirely divorced from relying on children as a method of saving.

5. Modeling a Modern Economy

By a "modern" economy, I mean an economy with a well-developed system of private property rights that allows for the operation of an asset market. While this can be modeled in many different ways, let me consider the "primitive" economy I described earlier and transform it into a "modern" economy by assuming that land is private property and that titles to land can be traded.

In the economy I describe here, the young will be able to save by purchasing titles to land. To simplify matters, I am going to assume that because of this, they need not save by investing in children. In the model described above, this would mean that $n = 0$, so that the population would quickly go to zero (this is because the only incentive for having children there was for saving purposes). To avoid having the population go to zero, let us simply assume that every person has one child and that this child costs nothing. In this case, the population will remain constant over time.

To begin, assume that the fixed stock of land L is owned by the initial old and that the land is divided equal among them (so that they each have L/N units of land).

The services of land and labor are sold in a competitive market. Let r_t denote the rental price of land and let w_t denote the rental price of labor (the real wage). It is an easy matter to demonstrate that these rental rates will be determined, in equilibrium, by their respective marginal products (see Appendix 9.1); i.e.,

$$\begin{aligned} r &= zF_L(L, N); \\ w &= zF_N(L, N). \end{aligned} \tag{70}$$

In fact, if F displays CRS, then these equilibrium factor prices will only depend on z and the land-labor ratio L/N. For example, if $F(L, N) = L^{\alpha}N^{1-\alpha}$ with $0 < \alpha < 1$, then $F_L(L, N) = \alpha(N/L)^{1-\alpha}$ and $F_N(L, N) = (1-\alpha)(L/N)^{\alpha}$.

Exercise 9.5: For $Y = zF(L, N) = zL^{\alpha}N^{1-\alpha}$, demonstrate that if factors are paid their marginal product (condition 70), then $Y = rL + wN$. That is, the total GDP is equal to total factor income.

Now, at any given date t, the old own s_t shares of the land. They rent this land (say, to a firm that organizes production) at rental rate r. Hence, they earn rental income rs_t. But as the old expect to die, they will want to sell their shares. Assume that they can do so on an asset market; and let q_t denote the price per share. The old receive no transfer of income from their children here; so that their budget constraint is given by:

$$c_t^o = [r + q_t]\, s_t.$$

Now let us consider a young household. The young rent their labor (again, to a firm that organizes production) at rental rate w. This is all the income they earn (and they make no transfers to the old). They may either consume this income or save some of it by purchasing shares in land. The shares that they purchase when young only entitle them to the land's rental income in the future. So, in effect, they purchase shares to future land, s_{t+1}. Hence, a young agent faces the following set of budget constraints:

$$\begin{aligned} c_t^y + q_t s_{t+1} &= w; \\ c_{t+1}^o &= [r + q_{t+1}] s_{t+1}. \end{aligned}$$

Combining these two budget constraints (substituting out for s_{t+1}), we can derive the lifetime budget constraint:

$$c_{t+1}^o = \left[\frac{r + q_{t+1}}{q_t}\right] w - \left[\frac{r + q_{t+1}}{q_t}\right] c_t^y.$$

The term in square brackets represents the (gross) rate of return associated with purchasing a share of land. That is, the share costs q_t units of output and it delivers $r + q_{t+1}$ units of output in the future. Hence, the rate of return is given by:

$$R_{t+1} \equiv \left[\frac{r + q_{t+1}}{q_t}\right];$$

so that the lifetime budget constraint may alternatively be written in a more familiar form:

$$c_{t+1}^o = R_{t+1} w - R_{t+1} c_t^y.$$

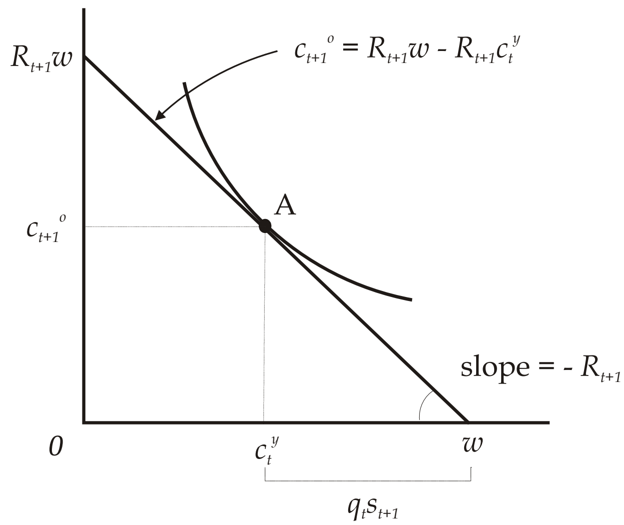
The desired consumption-saving choice can be depicted by a point like A in Figure 9.6.

As usual, any change in the interest rate (R_{t+1}) will have offsetting substitution and wealth effects on the level of desired saving. Let us assume, for simplicity, that these two effects exactly cancel. This can be guaranteed here by assuming a utility function of the form used in (67); with the associated MRS given by (68).

Excercise 9.6: *Use the MRS in (68) to show that desired saving here is given by:*

$$q_t s_{t+1}^D = \left(\frac{\beta}{1+\beta}\right) w.$$

FIGURE 9.6
Desired Consumption and Saving



From Exercise 9.6, we see that the demand for land shares is given by:

$$s_{t+1}^D = \left(\frac{\beta}{1+\beta}\right)\left(\frac{w}{q_t}\right). \tag{71}$$

That is, the young demand more shares as their income rises, and demand fewer shares as the price of shares increases. Hopefully, this makes sense to you.

Now, the initial old each start off with one share of land, so that the aggregate supply of shares is just given by N. The aggregate demand for shares is given by Ns_{t+1}^D. In equilibrium then, it must be the case that $s_{t+1}^D = 1$ (each young agent purchases one share of land). That is, using (71), we see that the following must hold:

$$1 = \left(\frac{\beta}{1+\beta}\right)\left(\frac{w}{q_t}\right).$$

We can solve this expression for the equilibrium share price of land; i.e.,

$$q = \left(\frac{\beta}{1+\beta}\right)w.$$

That is, the share price remains constant over time, so that the equilibrium return on land remains constant over time as well;

$$R = \left(\frac{r+q}{q}\right). \tag{72}$$

Exercise 9.7: If we interpret R as the (gross) real rate of interest, use (72) to show that that the equilibrium share price of land can also be expressed as the present value of the rental income generated by land.

Let me summarize. Equilibrium in the factor markets implies that rental prices (w, r) are determined by productivity z and the land-labor ratio (L/N). Equilibrium in the asset market then determines q as a function of preferences (β) and the equilibrium wage rate (which itself, depends on z and L/N). The old survive by renting their land to firms and then selling their land to the young. The young generate income by working and use their earnings to consume and purchase shares in land. When the future rolls around, the young become old; and the whole process repeats itself.

Note that while there are similarities in the way this "modern" economy functions in relation to the "primitive" economy studied earlier, there are important differences as well. They are similar in the sense that agents would naturally like to smooth consumption over the life cycle. The only way this can occur in the primitive economy, however, is by way of **intergenerational transfers**. That is, the old generation requires income support from the young generation (their children). For this reason, the young are motivated to increase family size when their incomes improve (say, because productivity increases). We have already studied the long-run effects of having children serve as a form of retirement saving.

In the modern economy, there is no need for intergenerational transfers from young to old.[68] If the old are in possession of capital (land), they can live off of their capital. The young who desire to smooth their consumption can now do so by purchasing assets (land). If their incomes rise (because of an increase in productivity), they will want to increase their savings. But they need not do this by increasing their family size. In this way, an increase in productivity will not manifest itself in the form of a larger population. Instead, it manifests itself as an increase in real per capita income.[69]

6. An Interpretation of the Great Transition

Consider a world consisting of many regions. Each region is initially in a "primitive" stage; so that each economy is described by Malthusian dynamics. Technological advancements appear throughout time and slowly diffuse

[68] Indeed, in modern economies, it is typically the case that intergenerational transfers move in the opposite direction; i.e., from the old to the young.

[69] This result also follows from the Solow growth model. People frequently suggest that the Solow growth model differs from the Malthusian model in that the former allows for an endogenous accumulation of physical capital (whereas the Malthus model assumes a fixed stock of land). As the analysis here demonstrates, this perception is incorrect. The key difference between Malthus and Solow is that the latter assumes that net population growth is insenstive to economic conditions. This has nothing to do with an endogenous capital stock and everything to do with the use of assets as an alternative saving vehicle.

throughout the regions of the world. For simplicity, assume no migration across regions (which is not quite accurate). As new advancements are implemented across regions, they experience transitory increases in living standards, followed by population growth, and living standards that–in the long-run–do not change very much. The world's population grows.

Now, imagine that at some point in time, say around 1750, the pace of technological advancement increases in one small region of the world (England). The population there begins to rise rapidly. At the same time, imagine the appearance of some "institutional shock" that slowly transforms this region into a "modern" economy. Population growth will slow, and per capita incomes remain at a higher level permanently.

Next, imagine that these advancements (both technological and institutional) begin to spread to surrounding areas (continental Europe and North America). Populations there begin to rise along with material living standards.

Imagine further that the technological advancements begin to spread around the globe at a relatively rapid pace–but that the institutional innovation spreads much more slowly (if at all). Then the world's population will begin to grow rapidly. But for those regions that have not adopted the new institution, per capital incomes remain low. At the world level, average per capita incomes rise–but the rise is concentrated in those economies that have adopted the new institution. Population growth, on the other hand, is centered among those regions that have adopted the more productive technologies, but have not adopted the new institution.

This interpretation is, I think, broadly consistent with the pattern of development we have observed. I would not go so far as to say that it is the only thing that accounted for the great transition (documented in Figure 9.1); but it certainly seems plausible that it played some role. One might note too that this idea is not inconsistent with cross-country correlations between material living standards and historical rates of population growth and saving (the correlation is negative for the former and positive for the latter).

If this interpretation is correct, then it suggests that current foreign aid policies are largely misguided. Introducing synthetic nitrogen fertilizers (a part of the so-called "Green Revolution") in lesser-developed regions of the world may improve agricultural productivity, but absent institutional reform, any increase in material living standards is likely to be thwarted by a rapidly expanding population (a concern that has been widely expressed by critics of the Green Revolution). Foreign aid might be better directed at developing institutions that allow poor people to save through means other than investing in children.

7. Recent Developments

FIGURE 9.7

Real per Capita GDP Relative to the United States

Selected Countries 1950-2000

Consider the data plotted in Figure 9.7.[70] This figures plots a country's real per capita GDP relative to the United States. You should keep in mind that the U.S. was generally growing throughout this sample period at around 2% per annum. Therefore, a "flat" trajectory on this graph means that the country in question was growing as quickly as the U.S. A rising trajectory means that the country was growing more quickly than the U.S.; and of course, a declining trajectory means that growth was slower than the U.S. (it does not necessarily mean that per capita GDP was declining).

Our transition story above might be consistent with what has happened with a number of these countries. A number of Latin American, Middle Eastern, and African countries may, for example, be held back from increasing their material living standards by the Malthusian forces described above. Some historically lesser-developed regions of the world–for example, Japan, South Korea, and Hong Kong–have evidently "modernized" along some dimension. One cannot help but note that these latter countries now have very well-developed financial markets (whether this development caused economic growth or was the product of economic growth is the matter of some debate).

But there must be something else going on. For example, even among the set of "modern" economies of Western Europe, evidence of income disparity across countries remains. Likewise, the idea that Malthusian dynamics are holding back countries like Israel or Argentina do not seem compelling. And before one gets too excited about Malthusian interpretations, one should keep in mind that there are several regions of the world with very high population densities and relatively high living standards (Japan, Hong Kong, Taiwan, etc.).

One interesting line of enquiry rests on the observation that there are persistent productivity differences across countries when one controls for the level of inputs. That is, if one tries to estimate aggregate production possibilities across countries based on a neoclassical production function of the form:

$$Y = zF(K, N),$$

where K here is now interpreted to be a broad measure of capital inputs (including land), then most of the cross-country differences in per capita income Y/N appear to be accounted for by differences in the parameter z (**total factor productivity**, or **TFP** for short).

Of course, even if this is true, it does not explain **why** TFP differs across countries. Researchers, like Stephen Parente and Ed Prescott argue that a country's TFP is a function of "barriers" that constrain the technology choice of firms located there.[71] The "barriers" that these authors have in mind consist of the successful political lobbying of various **special interest groups** that

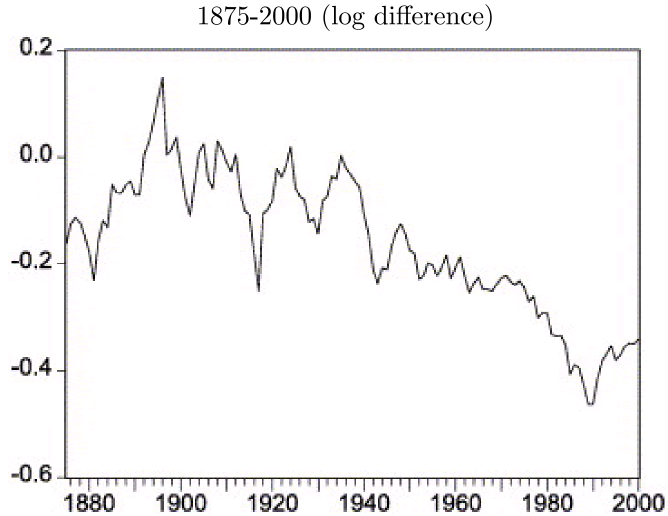
[70] Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

[71] Parente, Stephen and Ed Prescott (1999). *Barriers to Riches*, Third Walras Lecture, University of Lausanne.

prevent the adoption of technologies or work methods that threaten their own well-being (even if the well-being of the broader economy may be improved by adoption).[72] This idea rests on the plausible notion that any given technological advancement is not likely to affect all people in the same way–there will be winners and losers. If the potential losers form an effective lobby group, technology adoption will be constrained. Societies with institutions in place that are better able to resist the efforts of lobby groups to block the adoption of new technology will therefore be more productive on average. Technology change will always hurt and benefit some; but in the long-run, all will be better off if a society simply remains open to technological advancement.

The Parente-Prescott hypothesis may or may not be true; but it does have an air of plausibility to it. It seems clear enough that at some level, a country's institutions play some role in either enhancing or prohibiting economic growth and development. One interesting case study involves comparing Argentina to Australia and Canada.[73] About 100 years ago, these three countries were very similar in many respects. All three were "settler countries," drawing large numbers of migrants from Europe. All three countries had vast expanses of land and were rich in resources. All three countries had very similar real per capita incomes. And all three countries held the promise of future riches for their inhabitants.

FIGURE 9.8
Real per capita GDP in Argentina Relative to Canada
1875-2000 (log difference)



But at the turn of the last century, their growth paths began to diverge; with Canada and Australia heading in one direction, and Argentina heading in the

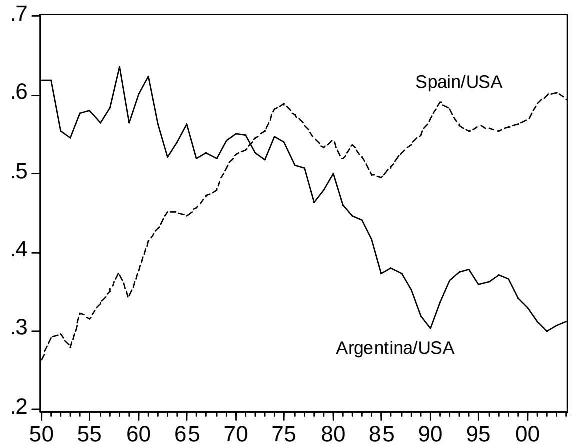
[72]See Appendix 9.2 for some examples throughout history.

[73]Sanz-Villarroya, Isabel (2005). "The Convergence Process of Argentina with Australia and Canada: 1875-2000," *Explorations in Economic History*, 42(3): 439-458.

other; see Figure 9.8. According to Figure 9.8, average material living standards in Argentina today are almost 40% lower than in Canada. Malthusian dynamics do not explain this divergence.

One might be inclined to explain this divergence in terms of the different English/Spanish heritages of these economies. But Figure 9.9 demonstrates that any such hypothesis is likely to fail.

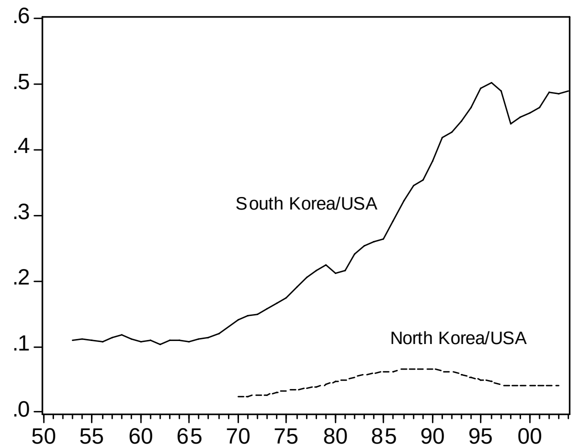
FIGURE 9.9
Real per Capita GDP Relative to U.S.
1950 - 2004



To demonstrate further that "cultural" differences alone cannot account for development patterns, consider the Korean peninsula, which consists of a relatively homogeneous population with a common tradition. This peninsula was divided into two during the Korean war (1950-1953), with the northern part adopting "communist" institutions and the southern part adopting "capitalist" institutions. This same culture, with different institutions, has produced very different results; see Figure 9.10.

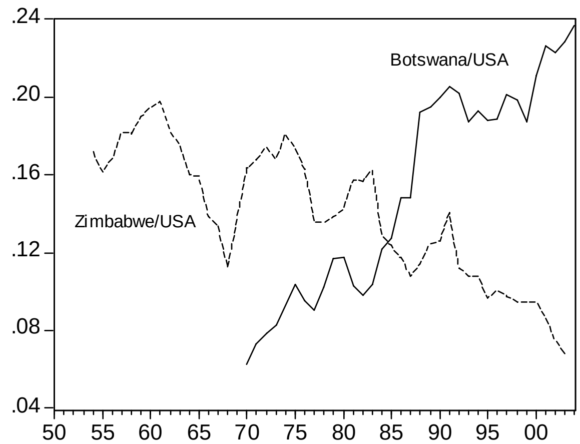
Figure 9.10 certainly does seem to suggest that institutions matter and that cultural traditions do not. It does not, however, explain why two apparently similar cultures ended up adopting very different institutions.

FIGURE 9.10
Real per Capita GDP Relative to U.S.
1950 - 2004



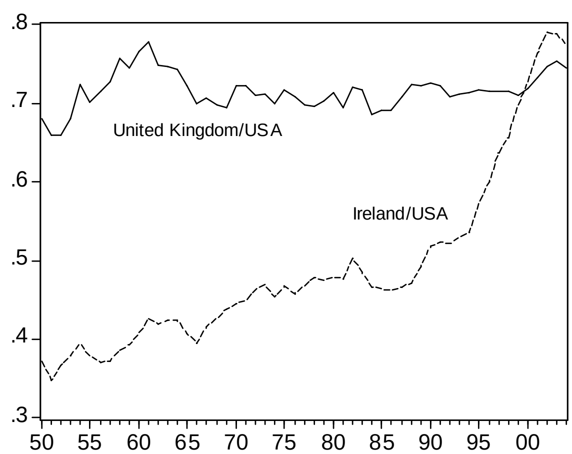
Perhaps you have in mind that the recent history of "colonialism" plays some role in explaining why some countries today are so poor and why they fail to catch up with the rest of the world. This is an argument that is frequently applied to the sub-Saharan African continent. But consider two neighboring countries in southern Africa formerly under British colonial rule: Botswana and Zimbabwe (formerly, Rhodesia); see Figure 9.11.

FIGURE 9.11
Real per Capita GDP Relative to U.S
1950 - 2004



An interesting case study pertains to Ireland. While the Celtic and Anglo-Saxon cultures between Ireland and the United Kingdom may differ along various dimensions, take a look at Figure 9.12.

FIGURE 9.12
Real per Capita GDP Relative to U.S.
1950 - 2004



The development of Ireland throughout the 1990s is intriguing. Long considered to be one of Europe's "basket-cases," economic growth began to take off in the late 1980s. The growth path has been so dramatic that average living standards in Ireland now exceed those of the United Kingdom. A natural question to ask is what happened? What happened is succinctly summarized by the following quote:[74]

> While 1987 marked the bottom of a long recession, it was also the year Charles Haughey took over as prime minister and decided that the economic system should be rebuilt from scratch. He even managed to sell his idea to the opposition and to the most important interest groups, including the unions. What would later be called a miracle started with a social contract between the government, the employers, and the unions. The contract included tax cuts and some financial support for those worst off.

Another important element of the Irish reform was opening the doors for foreign investment. In short, a whole series of reforms were introduced that altered incentives in a manner that promoted growth. The reforms were made politically feasible by two factors. First, given the poor economic record, there was a general consensus that something had to be done differently; and second, the reforms included provisions that to some extent mitigated the pain of transition.

8. Summary

The "problem" of economic development remains a largely unresolved issue. The issues involved are complicated and far beyond the scope of this chapter. Nevertheless, there are some lessons to be gleaned from the data.

We know, for example, that for most of recorded history, technological progress has manifested itself primarily in the form of population growth, rather than growth in material living standards. For "primitive" economies, the Malthusian model provides a plausible interpretation of this pattern of economic development. It also suggests that for primitive economies in the modern world, erecting institutions that induce people to alter their fertility choices is likely important. Draconian policies, such as China's "one-child" policy may be one way to go. But perhaps a better alternative would be to develop financial institutions that allow people to save by means other than by having children.

Malthusian dynamics alone cannot fully explain the persistent discrepancy in living standards that one observes even across "modern" economies. Richer countries tend to be more productive economies. But this does not explain why poor countries simply do not copy the technology and institutions employed by

[74] Sigfrid, Karl (2004). *The Irish Miracle*, www.fee.org/pdf/the-freeman/sigfrid0404.pdf

rich countries. Differences in culture or historical experiences may be important; but the differences we observe in living standards do not correlate perfectly–or even very well–with such differences. There is some evidence, however, which suggests that "pro-market" reforms tend to promote growth. The key constraint here appears to be political in nature. Especially in democratic countries, reform packages must be structured in a manner that renders them politically implementable. Evidently, this is easier said than done.

Appendix 9.1
Factors are Paid Their Marginal Products

Consider an economy where firms operate a CRS production technology $Y = zF(L,N)$. These firms hire factors of production (L,N) on a competitive factor market, with rental rates (r,w). A representative firm's real profit is therefore given by:

$$zF(L,N) - rL - wN.$$

Maximizing this profit function entails choosing factor demands (L^D, N^D) that satisfy:

$$\begin{aligned} zF_L(L^D, N^D) &= r; \\ zF_N(L^D, N^D) &= w. \end{aligned}$$

If the supply of these factors is fixed (as in our model in the body of the chapter), then the factor-market clearing conditions require:

$$\begin{aligned} L^D(w^*, r^*, z) &= L; \\ N^D(w^*, r^*, z) &= N. \end{aligned}$$

These two conditions determine the equilibrium factor prices (w^*, r^*). But if we combine the latter set of equations with the former, we see that:

$$\begin{aligned} r^* &= zF_L(L,N); \\ w^* &= zF_N(L,N). \end{aligned}$$

Since F is linearly homogenous in (L,N), it follows that Euler's theorem):

$$zF(L,N) = zF_L(L,N)L + zF_N(L,N)N.$$

Using the fact that factors are paid their marginal products, this may alternatively be expressed as:

$$Y = r^*L + w^*N.$$

That is, the total GDP is paid out to land and labor (so that profits are zero). The relation above also motivates the use of the income approach to computing the GDP (the sum of all incomes paid to domestic factors of production).

Appendix 9.2
Special Interests Throughout History

It is important to realize that barriers emanating from special interests have always been present in all economies (from ancient to modern and rich to poor), so that any differences are really only a matter of degree. For example, as early as 1397, tailors in Cologne were forbidden to use machines that pressed pinheads. In 1561, the city council of Nuremberg, apparently influenced by the guild of red-metal turners, launched an attack on Hans Spaichl who had invented an improved slide rest lathe. The council first rewarded Spaichl for his invention, then began to harass him and made him promise not to sell his lathe outside his own craft, then offered to buy it from if he suppressed it, and finally threatened to imprison anyone who sold the lathe. The ribbon loom was invented in Danzig in 1579, but its inventor was reportedly secretly drowned by the orders of the city council. Twenty five years later, the ribbon loom was reinvented in the Netherlands (and so became known as the Dutch loom), although resistance there too was stiff. A century and a half later, John Kay, the inventor of the flying shuttle, was harassed by weavers. He eventually settled in France, where he refused to show his shuttle to weavers out of fear. In 1299, an edict was issued in Florence forbidding bankers to use Arabic numerals. In the fifteenth century, the scribes guild of Paris succeeded in delaying the introduction of the printing press in Paris by 20 years. In the sixteenth century, the great printers revolt in France was triggered by labor-saving innovations in the presses.

Another take on the special interest story pertains to case in which the government itself constitutes the special interest, as in the case of autocratic rulers. It seems as a general rule, weaker governments are able to exert less resistance to technological adoption. With some notable exceptions, autocratic rulers have tended to be hostile or indifferent to technological change. Since innovators are typically nonconformists and since technological change typically leads to disruption, the autocrat's instinctive desire for stability and suspicion of nonconformism could plausibly have outweighed the perceived gains to technological innovation. Thus, in both the Ming dynasty in China (1368–1644) and the Tokugawa regime in Japan (1600–1867) set the tone for inward-looking, conservative societies. Only when strong governments realized that technological backwardness itself constituted a threat to the regime (e.g., post 1867 Japan and modern day China) did they intervene directly to encourage technological change.

During the start of the Industrial Revolution in Britain, the political system strongly favored the winners over the losers. Perhaps this was because the British ruling class had most of its assets in real estate and agriculture which, if anything, benefited from technological progress in other areas (e.g., by increasing land rents). However, even in Britain, technological advances were met by stiff opposition. For example, in 1768, 500 sawyers assaulted a mechanical saw mill in London. Severe riots occurred in Lancashire in 1779, and there many instances of factories being burned. Between 1811 and 1816, the Midlands and the industrial counties were the site of the 'Luddite' riots, in which much damage was inflicted on machines. In 1826, hand-loom weavers in a few Lancashire towns rioted for three days. Many more episodes like these have been recorded.

But by and large, these attempts to prevent technological change in Britain were unsuccessful and only served to delay the inevitable. An important reason for this is to be found in how the government responded to attempts to halt technological progress. In 1769, Parliament passed a harsh law in which the wilful destruction of machinery was made a felony punishable by death. In 1779, the Lancashire riots were suppressed by the army. At this time, a resolution passed by the Preston justices of peace read: "The sole cause of the great riots was the new machines employed in cotton manufacture; the country notwithstanding has greatly benefited by their erection and destroying them in this country would only be the means of transferring them to another...to the great detriment of the trade of Britain."

The political barriers to efficiency manifest themselves in many ways, from trade restrictions and labor laws to regulatory red tape. For example, a recent World Bank report (Doing Business in 2004: Understanding Regulation) documents the following. It takes two days to register a business in Australia, but 203 days in Haiti. You pay nothing to start a business in Denmark, while in Cambodia you pay five times the country's average income and in Sierra Leone, you pay more than 13 times. In more than three dozen countries, including Hong Kong, Singapore and Thailand, there is no minimum on the capital required by someone wanting to start a business. In Syria, the minimum is 56 times the average income; in Ethiopia and Yemen, it's 17 times and in Mali, six. You can enforce a simple commercial contract in seven days in Tunisia and 39 days in the Netherlands, but in Guatemala it takes more than four years. The report makes it clear, however, that good regulation is not necessarily zero regulation. The report concludes that Hong Kong's economic success, Botswana's stellar growth performance and Hungary's smooth transition (from communism) have all been stimulated by a good regulatory environment. Presumably, a 'good' regulatory environment is one which allows individuals the freedom to contract while at the same time providing a legal environment that protects private property rights

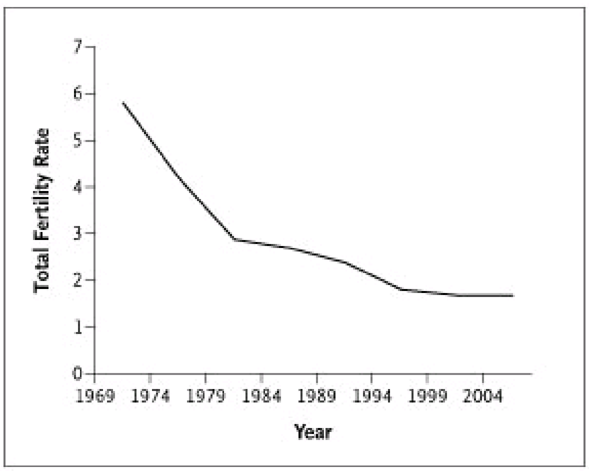
While there are certainly many examples of special interests working against the implementation of better technology, our political economy story is not without shortcomings. In particular, special interest groups are busy at work in all societies. The key question then is why different societies confer more or less power to various special interests. Perhaps some societies, such as the United States, have erected institutions that are largely successful at mitigating the political influence of special interests. These institutions may have been erected at a time when a large part of the population shared similar interests (e.g., during the American revolution).

But even if new technologies have sectoral consequences for the economy, it is still not immediately clear why special interests should pose a problem for the way an economy functions. For example, in the context of the model developed above, why do individuals not hold a diversified portfolio of assets that would to some extent protect them from the risks associated with sector-specific shocks? In this way, individuals who are diversified can share in the gains of technological progress. Alternatively (and perhaps equivalently), why do the winners not compensate (bribe) the losers associated with a technological improvement? These and many other questions remain topics of current research

Problems

1. In 2005, the United Nations sponsored a so-called "International Year of Microcredit." In 2006, the Bangladeshi economist Muhammad Yunnus won a Nobel Peace prize for his Grameen Bank; an institution that specialized in lending to the poor. If you perform a search on Google, you will discover that these microcredit facilities are spreading around the world. Explain the likely consequences for population growth.

2. As explained in the text, in 1979 China implemented its infamous "one-child" policy in an attempt to curtail population growth. Since 1979, the Chinese fertility rate has indeed declined. However, based on the evidence (see figure below), would one necessarily attribute this decline in fertility to the one-child policy?[75] What other factors may have been responsible for the decline in the fertility rate?



3. In one of the greatest pandemics recorded in history, the "Black Death" (bubonic plague) is estimated to have wiped out between one-third and two-thirds of Europe's population over the period 1347-48. In the context of the Malthusian model, this shock can be modeled as an exogenous decline in the population. Explain the model's predictions concerning the time path of incomes and population following this event.

4. Contrary to popular belief, real wages in Europe did not rise immediately following the Black Death.[76] In fact, it appears that wages began to fall just prior to the plague and continued to fall subsequent to it–not

[75] This figure is from: http://content.nejm.org/cgi/content/full/353/11/1171

[76] www.flutrackers.com/forum/showthread.php?t=21798

recovering until 1370. Is there any way to reconcile this observation with the Malthusian model? Hint: could there have been independent shocks to TFP or may have the plague in some way have affected TFP temporarily?

5. Why do eggs cost twice as much in Canada as they do in the United States?