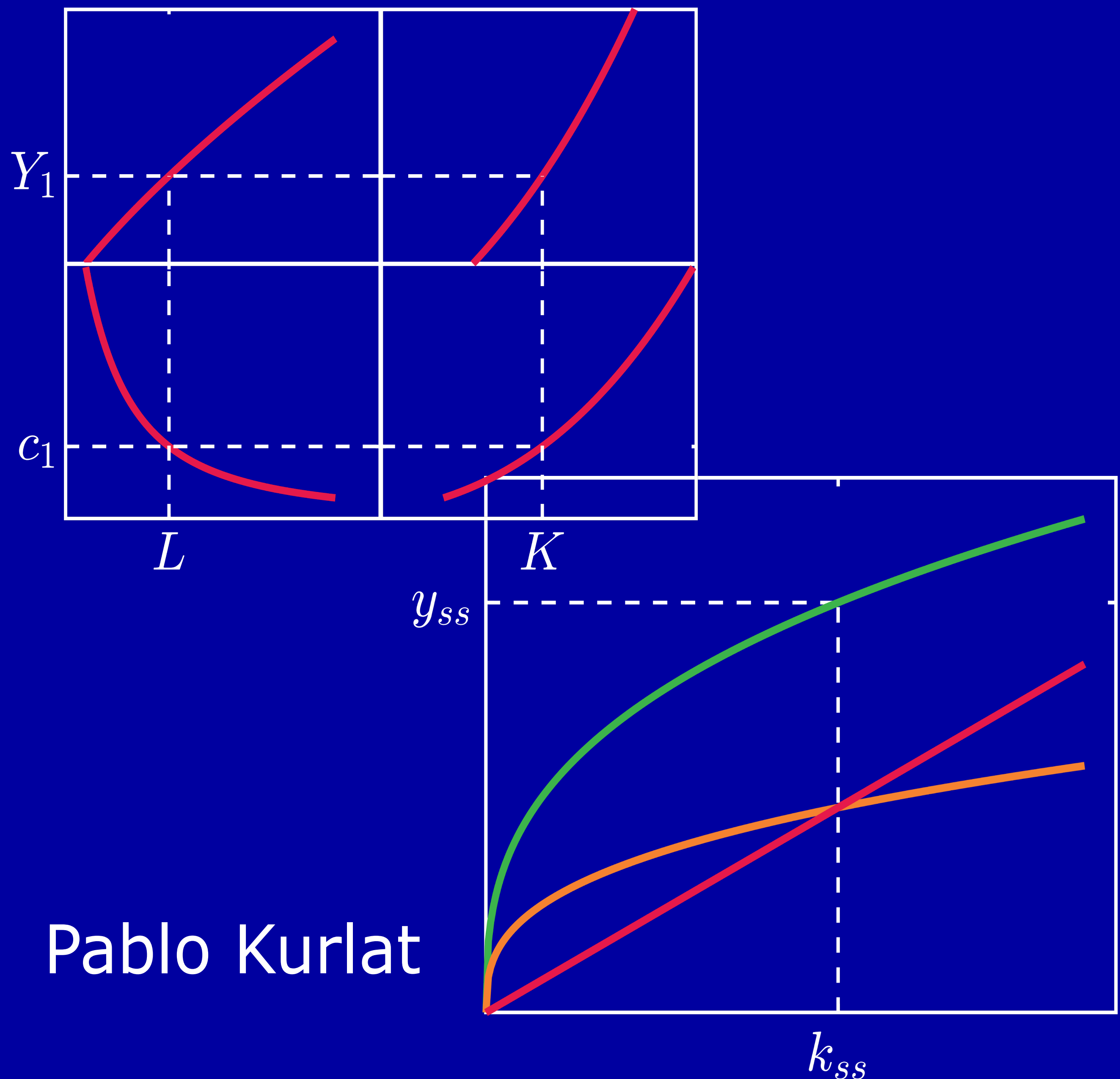
A Course in Modern Macroeconomics



Pablo Kurlat

A Course in

Modern Macroeconomics

Pablo Kurlat

https://sites.google.com/view/pkurlat

ISBN: 9781073566716

For Pam, Felix and Emilia

Introduction

This book started as a collection of my teaching notes for the ECON 52 course that I taught at Stanford University. The objective of that course, and of this book, is to introduce students to the ideas and way of thinking of modern macroeconomics in a unified way that is accessible with a moderate amount of maths. Modern macroeconomics emphasizes explicit microeconomic foundations and general equilibrium analysis, combined with various kinds of constraints and market imperfections. When preparing the class I thought none of the existing textbooks conveyed this in a way that I liked, so I prepared my own notes, which then grew into this book. While mostly self-contained, the book is probably most useful to students who are familiar with the basics of multivariable calculus and have taken a calculus-based microeconomics class.

The book is meant to be followed approximately in order. Later chapters contain many references to material in earlier chapters. However, not everything from the early chapters is indispensable for what comes next. Chapters 1 and 6-9 are the main core, but even within them everything that has to do with risk, search, adjustment costs, or infinite-horizon problems can be skipped without compromising what comes later.

At the end of each chapter there is a series of exercises. Some are relatively direct applications of the material in the chapter and others are more open-ended or explore topics related to but not directly covered in the chapter. Several of the exercises can serve as the basis for a lecture, a class discussion, or the analysis of a historical episode. The exercises vary in difficulty but are intended to be relatively hard overall.

The list of interesting areas of macroeconomics is vast and growing, and the book does not aim to be comprehensive. Probably the biggest omission is that it mostly deals with closed-economy issues and models, so there is little discussion of exchange rates, capital flows or international trade. Somewhat relatedly, the book is more US-centric than I would like. In many ways the US economy is not like that of a typical country, but it is very well studied, so many of the ideas are discussed in terms of US evidence. The book is also biased towards my own idiosyncratic tastes. For instance, there is more than one might expect on money supply and demand, which is a somewhat old-fashioned topic, and on how to define living standards.

I would like to thank several generations of teaching assistants and students for their input. Alina Arefeva, Juliane Begenau, Daniel Bennett, Ricardo de la O, Adem Dugalic, Guzman Gonzalez-Torres, Pavel Krivenko, Krishna Rao, Yevgeniy Teryoshin, Daniel Layton Wright and Victoria Zuo all contributed to making this book possible. I would also like to thank researchers who allowed me to include their findings in the book, in particular Daron Acemoglu, Daniel Andrei, Regis Barnichon, Robert Barro, Chad Jones, David Lagakos, William Mann, Nathalie Moyen and Valerie Ramey. I also benefited greatly from discussions with people who have taught similar courses, especially Sebastian Di Tella, Pete Klenow, Monika Piazzesi and Alp Simsek.

On my website (https://sites.google.com/view/pkurlat) you can find some complementary materials: data and codes for some of the exercises, clarifications, corrections, downloadable figures, etc. If you have comments, questions, suggestions, criticisms, corrections, or praise (especially praise!), you can reach me at pkurlat@gmail.com.

Contents

PART I

GDP and Living Standards

This part of the book explores the meaning and measurement of living standards.

In Chapter 1 we look at one of the main variables that macroeconomists care about: the Gross Domestic Product, or GDP. We go over its definition, the accounting conventions used to measure it and some of the conceptual issues behind the accounting conventions.

In Chapter 2 we study some of the shortcomings of GDP as a measure of living standards, some alternatives that have been proposed, and how one can use economic theory as a guide to improved measurements.

CHAPTER 1

GDP

1.1 GDP Accounting

One of the basic questions economists are interested in, when analyzing a country, is how much is produced in that country in a given year. The basic measure of this is a country's gross domestic product (GDP). The idea is simple: to record the value of everything that is produced in the country in a year and add it up. GDP accounts can be constructed in three different (but equivalent) ways, based on measuring production, income, or expenditure.

Production	Income	Expenditure
Agriculture & Mining	Employee compensation	Consumption
Construction	Proprietor's income	Investment
Manufacturing	Rental income	Government
Services	Corporate profits	Net exports
Government	Interest income	
	Depreciation[1]	

In each of the three measures we can choose how much detail to go into. For instance, in the production approach we don't need to lump all services together. We can instead separate healthcare, education, entertainment, retail trade, etc., into separate accounts.

The accounting identity from the expenditure approach is sometimes written algebraically as:

$$Y = C + I + G + X - M \tag{1.1.1}$$

where Y stands for GDP, C stands for consumption, I stands for investment, G stands for public goods and services, X stands for exports and M stands for imports. We'll return to this equation many times.

The three measures of GDP are equal to one another. The logic is that whenever goods and services are produced, whatever is spent on them will also constitute someone's income. A good description of how the accounts are constructed can be found at https://www.bea.gov/resources/methodologies/measuring-the-

[1] Including depreciation as a form of income doesn't seem to make much sense, but see below.

economy. Table 1.1 shows measures of GDP for the US for 2017 computed according to each of the three approaches.

Table 1.1: *US GDP in 2017 according to the three methods. Figures in billions of dollars. Source: BEA.*

Production		Income		Expenditure	
Agriculture	169	Employee Comp.	10,421	Consumption	13,321
Mining	269	Corporate Profits	1,807	Investment	3,368
Utilities	308	Proprietor's income	1,501	Govt. spending	3,374
Construction	781	Rental income	730	Exports	2,350
Manufacturing	2,180	Depreciation	3,116	Imports	-2,929
Wholesale + Retail	2,261	Interest Income	768		
Transport	609	Taxes	1,286		
Media	1,051	Statistical discrep.	-143		
Finance + Insurance	1,466				
Real Estate	2,591				
Professional services	2,426				
Educ. + Health	1,700				
Arts + Entertainment	805				
Other services	416				
Government	2,454				
Total	**19,485**	**Total**	**19,485**	**Total**	**19,485**

In this section we will try to understand the logic of GDP accounts through a series of examples.

Example 1.1.
Amy, who is self-employed, produces lettuce in her garden and sells it to Bob for $1. Bob eats it.

Production		Income		Expenditure	
Agriculture	1	Proprietor's income	1	Consumption	1

The production approach measures the value of the lettuce that was produced, which is $1.

The income approach looks at how much income is derived from productive activities. In our example, Amy obtains $1 of income from selling the lettuce. Since Amy is self-employed, we classify her income as proprietor's income (self-employed people are sometimes called "sole proprietors").

The expenditure approach looks at what the production was used for. Here the lettuce was consumed.

Value Added

Production typically takes place in several stages. Someone's output becomes somebody else's input. We want to measure the value at the end of the production process, avoiding double counting.

Example 1.2.

Amy is the shareholder of a corporation that operates a fertilizer plant. The corporation hires Bob to work in the plant and pays him a wage of $0.50. The corporation sells the fertilizer to Carol, a self-employed farmer, for $0.80. Carol uses it to produce lettuce, which she sells to Daniel for $1. Daniel eats the lettuce.

Production		Income		Expenditure	
Manufacturing	0.8	Wages	0.5	Consumption	1
Agriculture	0.2	Profits	0.3		
		Prop. income	0.2		
Total	**1.0**	**Total**	**1.0**	**Total**	**1.0**

Here it would be a mistake to add the value of the fertilizer to the value of the lettuce because the fertilizer was used up in producing the lettuce. The value added in the production of lettuce is just the difference between the value of the lettuce and the value of the fertilizer. Notice that doing things this way makes total GDP consistent across the three methods.

Forms of Investment

Investment can take different forms, with one thing in common: it involves producing something that will be used for production in future periods.

Example 1.3.

1. General Electric builds an X-ray machine, which it sells to Stanford Hospital for $1,000. The cost of producing it is made up of workers' wages of $600.

Production		Income		Expenditure	
Manufacturing	1,000	Wages	600	Investment	1,000
		Corporate profits	400		
Total	**1,000**	**Total**	**1,000**	**Total**	**1,000**

2. Zoe builds a house with her bare hands and sells it to Adam for $1,000.

Production		Income		Expenditure	
Construction	1,000	Proprietor's income	1,000	Investment	1,000

3. Dunder Mifflin produces 500 tons of white paper worth $40,000 and stores them in its warehouse while it waits for customers to buy them. The cost of producing them is made up of workers' wages of $50,000.

Production		Income		Expenditure	
Manufacturing	40,000	Wages	50,000	Investment	40,000
		Corporate profits	-10,000		
Total	**40,000**	**Total**	**40,000**	**Total**	**40,000**

In part 1, the X-ray machine will be used to "produce" X-ray scans in the future. In part 2, the house will be used to produce shelter ("housing services") in future periods. "Equipment" (as in part 1) and "structures" (as in part 2) are the largest components of investment.

Part 3 is a little bit more subtle. The paper was produced to be sold and used, not in order to be left lying around in the warehouse. However, sometimes production and use are not synchronized. The goods that are held in order to be used later are called "inventories" and include finished goods but also inputs and half-finished products that will be part of a further productive process. Since inventories are something that will be useful in the future, an increase in inventories is also a form of investment. In the example, we make the interpretation that Dunder Mifflin has invested in having paper available for when it manages to make sales. When the paper is finally sold and inventories go back to zero we will record that as negative investment.

Example 1.4.
Warren invests $100,000 in shares of General Motors.

Production	Income	Expenditure
0	0	0

This example is a bit tricky because the word "investment" is used somewhat differently in macroeconomics than in other contexts. In the example above there is no investment in the macroeconomic sense. There is a change in ownership but no new productive assets are created.

Durables

The distinction between consumption and investment is not always so clear. Above we saw that residential construction is an investment because it will produce "housing services" in the future. By that logic, many things could be considered investments. A refrigerator produces "refrigeration services" for a long time after it's produced. Similarly for cars, electronics, clothes, etc. How does GDP accounting treat these?

Example 1.5.

1. Panasonic builds a TV (at zero cost) and sells it to Bob for $500.

Production		Income		Expenditure	
Manufacturing	500	Corporate profits	500	Consumption of Durables	500

2. Bob watches the TV he bought last year.

Production		Income		Expenditure	
Production	0	Income	0	Expenditure	0

3. A property developer builds a house (at zero cost) and sells is to Claire for $100,000.

Production		Income		Expenditure	
Construction	100,000	Corporate profits	100,000	(Residential) Investment	100,000

4. Claire lives in the house she bought last year. In the rental market, a similar house would cost $7,000 a year.

Production		Income		Expenditure	
Housing services	7,000	Imputed owner-occupier income	7,000	Consumption	7,000

Conceptually, what's going on with the TV and with the house is very similar. They are produced one year but are enjoyed for a long time thereafter. However, GDP accounting conventions treat them differently. For most durable goods, we just treat them as being consumed at the moment of purchase, though sometimes we classify consumption of durables separately from consumption of nondurables (e.g., food and entertainment) just to emphasize that they are not quite the same. For housing, since it's such a large category and it's very long-lived, we treat the initial construction as an investment and try to measure the flow of housing services even when an homeowner is buying those housing services from herself.

Foreign Countries

GDP includes everything produced within the country, whether it's eventually used by residents or non-residents. Conversely, goods produced abroad are not included in GDP even if they are consumed in the country.

Example 1.6.
A car manufacturer buys components from Japan for $10 and uses half of those components in the production of a car, which it sells to Andy for $20. There are no other production costs. It stores the rest of the components. Amy, who is self-employed, produces lettuce in her garden and sells it to Franz (a foreigner) for $2.

Production		Income		Expenditure	
Car Manufacturing	15	Corporate profits	15	Consumption	20
Agriculture	2	Proprietor's income	2	Investment	5
				Exports	2
				Imports	-10
Total	**17**	**Total**	**17**	**Total**	**17**

The Government

The government is a major producer of goods and services. Many of those services are provided directly, so there is no real price for them. In order to add them to GDP accounts, they are valued at whatever it cost to produce them.

Example 1.7.

1. Ms. Jody teaches Kindergarten in Lucille Nixon Elementary School in Palo Alto for the entire year and earns $85,000.

Production		Income		Expenditure	
Public Education	85,000	Wages	85,000	Government	85,000

2. The City of Palo Alto hires the Los Angeles Philharmonic to play a free concert in Stanford Stadium. The musicians are paid $65,000 and renting the stadium costs $20,000. Four people show up.

Production		Income		Expenditure	
Public Concert	85,000	Wages	65,000	Government	85,000
		Rental income	20,000		
Total	**85,000**	**Total**	**85,000**	**Total**	**85,000**

Notice that GDP is the same in both examples, even though in one case the publicly provided service is something people actually value a lot and in the other case it's not.

Example 1.8.
Jack collects his $20,000 pension from Social Security.

Production		Income		Expenditure	
	0		0		0

Here the government is "spending" $20,000 but it's not in order to produce public goods and services. In terms of GDP accounting, this is just a transfer, which has no impact on any of the accounts.

Example 1.9.
The state of California builds a high-speed train from Merced to Bakersfield. It pays workers a billion dollars to build it with their bare hands.

Production		Income		Expenditure	
Railway construction	1 billion	Wages	1 billion	Government Investment	1 billion

This is an example of public investment: something the public sector does that will be useful in the future. In the expenditure approach, do we classify it as "Government Spending" or as "Investment"? In equation (1.1.1), it's included within G, but more detailed GDP accounts include a further breakdown of G into government investment and government consumption. The previous examples were all government consumption. This example is government investment.

Depreciation

Machines and buildings usually deteriorate over time, a phenomenon we call "depreciation." GDP is *gross* domestic product because it is measured before taking into account of depreciation.

Example 1.10.
Zak's Transport Co. owns a fleet of taxis. They are all new at the beginning of the year, worth a total of \$1,000. A taxi depreciates completely in 5 years. During the course of the year the company pays its workers \$200 in wages, has no other costs, and collects \$500 in fares.

Production		Income		Expenditure	
Transp. Services	500	Corporate Profits		Consumption	500
		Revenue	500		
		Wages	-200		
		Depreciation	-200		
		Total Profit	100		
		Wages	200		
		Depreciation	200		
Total	**500**	**Total**	**500**	**Total**	**500**

Since the taxis depreciate over 5 years, an estimate of the amount of depreciation is $\frac{1000}{5} = 200$. When the company computes its profits, it understands that its fleet of vehicles has lost value over the course of the year, so it subtracts the amount of depreciation. In order to compute GDP we want to get back to a before-depreciation measure, so we add back depreciation. This makes the income-based measure of GDP consistent with the production-based measure and the expenditure-based measure.

Depreciation plays an important role in the theory of economic growth that we'll study in Chapter 4.

Non-Market Activities

A lot of economic activity does not involve market transactions and is usually not included in GDP calculations. We already saw an exception to this: we impute the production of housing services even for people who live in their own home without conducting a market transaction. This particular exception is made so that GDP does not vary when housing shifts between tenant occupancy and owner occupancy. (Note that the imputed rent of owner-occupied housing accounted is almost 8% of US GDP.) Most of the time, however, we compute the value of an activity only if it is sold in the market.

Example 1.11.

1. Mary mows Andy's lawn for $25. Andy takes care of Mary's kids for $25.

Production		Income		Expenditure	
Gardening services	25	Self-employment	50	Consumption	50
Babysitting	25				
Total	**50**	**Total**	**50**	**Total**	**50**

2. Andy mows his own lawn. Mary takes care of her own kids.

Production		Income		Expenditure	
Production	0	Income	0	Expenditure	0

These two examples show that, even though the economic activity is basically the same in both cases, national accounts treat them very differently.

1.2 Making Comparisons

One of the things we often want to do is compare GDP, either across countries or within a country across time. To do this we have to be a bit careful with the units of measurement. When we compute GDP, we just add the value of everything produced in the country. If it's for the US, it will be in dollars. The problem with this measure is that the amount of goods and services you can get for one dollar is not the same in every country or in every time period, because the prices of goods and services are different.

For this reason we make a distinction between "nominal" and "real" GDP:

- Nominal GDP: the total value of goods and services produced, valued at whatever price they had at the time they were produced.
- Real GDP: the total value of goods and services produced, valued in units such that the values are comparable across time.

Real GDP

Example 1.12.
The country of Kemalchistan uses the dinar as its currency. GDP in the years 2017 and 2018, measured by the production method, was as follows:

2017		2018	
Manufact. (50 balls, 10 dinar each)	500	Manufact. (50 balls, 20 dinar each)	1,000
Educ. (10 teachers, 100 dinar each)	1,000	Educ. (10 teachers, 200 dinar each)	2,000
Total	**1,500**	**Total**	**3,000**

GDP, measured in dinar, doubled between 2017 and 2018 but the amount of goods and services was the same in both years. The reason GDP increased is because all prices increased. Often we are interested in a measure that tracks the changes in the total amount of stuff that is produced and doesn't rise just because prices have changed.

In the example above, it's clear that real GDP is the same in both years and we can express it as either "1,500 dinars in 2017 prices" or "3,000 dinars in 2018 prices." Either way, GDP did not grow between the two years. But the example is special in two ways:

1. The relative quantities of the different goods produced don't change between the two years.

2. All the prices change by the same amount.

When these conditions fail, the way to measure real GDP in a way that's comparable across years is less obvious.

Example 1.13.
The country of Expandia uses the dollar. GDP in the years 2017 and 2018, measured by the production method, was as follows:

2017		2018	
Agric.(10 tons of wheat, $50 each)	500	Agric. (11 tons of wheat, $60 each)	660
Manufact. (1 computer, $1,000)	1,000	Manufact. (2 computers, $600 each)	1,200
Total	**1,500**	**Total**	**1,860**

How much has the real output of the economy of Expandia grown? We know that agricultural output has expanded 10% (from 10 to 11) and manufacturing output has grown 100% (from 1 to 2). How should we compute the total growth? There is more than one way to do it.

Alternative 1: base year prices

One approach is to choose a "base year" and measure the value of all goods at the prices they used to have in the base year. In the example above, if we chose 2017 as the base year, we'd have the following figures for real GDP for the year 2018:

2018, at 2017 prices	
Agriculture (11 tons of wheat, $50 each)	550
Manufacturing (2 computers, $1,000 each)	2,000
Total	**2,550**

We'd say that "real GDP in 2018 was $2,550 at 2017 prices." If we want to compute the rate of growth of GDP, we would have

$$\text{growth} = \frac{2,550}{1,500} - 1 = 70\%$$

The general formula for computing real GDP this way is:

$$Y_t = \sum_i p_{i0} q_{it} \tag{1.2.1}$$

where:

- Y_t is real GDP in the year t.
- p_{i0} is the price of a certain good i in the base year (which we call year 0)
- q_{it} is the quantity of good i produced in year t

Alternative 2: final year prices

This is exactly the same, except that the base year is the last one we look at rather than the first one. In the example above, this means recomputing GDP in the year 2017 at the prices of 2018:

2017, at 2018 prices	
Agriculture (10 tons of wheat at $60 each)	600
Manufacturing (1 computers at $600 each)	600
Total	**1,200**

The general formula (1.2.1) still applies, it's just that we have changed what year we call year 0. With 2018 as the base year, we'd say that "real GDP in 2017 was $1,200 at 2018 prices," and the rate of growth of GDP is

$$\text{growth} = \frac{1,860}{1,200} - 1 = 55\%$$

Notice that the two formulas give us a different answer to the question "how much did the economy grow overall between 2017 and 2018?" This is often the case. Using an earlier year as the base year gives a higher rate of growth if the sectors that are expanding most (in the example, manufacturing) are those whose relative price is falling, and vice versa.

Alternative 3: chained prices

Neither of the above alternatives is obviously preferred, so another option is to do something in between. The idea is to:

1. Start from some base year 0
2. Compute real growth between year 0 and year 1 in two ways: at year 0 and year 1 prices
3. Average the two growth rates in some way

4. Compute real GDP in year 1 by adding the "average" growth rate to year-0 GDP

5. Repeat for years 2, 3, 4, etc.

The term "chained" comes from the fact that the estimate of real GDP in any given year will be the result of a chain of calculations linking that year to the base year. In general formulas:

$$g_t^I = \frac{\sum_i p_{it-1} q_{it}}{\sum_i p_{it-1} q_{it-1}} - 1 \qquad \text{growth based on initial year prices}$$

$$g_t^F = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{it} q_{it-1}} - 1 \qquad \text{growth based on final year prices}$$

$$g_t = \left(1 + g_t^I\right)^{0.5} \left(1 + g_t^F\right)^{0.5} - 1 \qquad \text{average growth; this is a geometric average}$$

$$Y_t = Y_{t-1}(1 + g_t) \qquad \text{real GDP one year ahead}$$

This will result in a measure of GDP in "chained" prices of the base year.

Comparisons Across Countries and PPP

Suppose we want to compare GDP across countries.

Example 1.14.
In 2018, GDP in the US and Mexico were as follows:

	United States	Mexico
GDP	20.5 trillion dollars	23.5 trillion pesos
Population	327 million	127 million
GDP per capita	62,700 dollars per person	185,000 pesos per person

Suppose we wanted to ask: did US residents produce more output per person than Mexican residents in 2018? The figures above don't quite give us the answer because they are in different units: GDP in the US is measured in dollars while GDP in Mexico is measured in pesos. How do we convert everything to the same units?

One approach is to look up the exchange rate between the Mexican peso and the US dollar. On average during 2018, you could trade one dollar for about 19 Mexican pesos in foreign exchange markets; equivalently, you could trade one Mexican peso for 0.053 US dollars. Using this exchange rate, we can restate Mexican GDP in US dollars as:

$$\begin{array}{ccccc} \text{GDP in Foreign Country} & = & \text{market exchange rate} & \times & \text{GDP in Foreign Country} \\ \text{(in dollars, at market} & & \text{(dollars per unit of foreign} & & \text{(in foreign currency)} \\ \text{exchange rates)} & & \text{currency)} & & \end{array}$$

Using this approach, we'd conclude that Mexico's GDP in 2014 was 1.24 trillion dollars, or 9,700 dollars per person.

One drawback of this approach is that it doesn't take into account that, even after converting currencies, prices are different in different countries. In other words, if you take one dollar, use it to buy Mexican pesos, go to Mexico and go shopping, the amount of stuff you'd be able to afford need not be equal to the amount of stuff you'd be able to afford if you had just stayed in the US. When we see that a country has low GDP when converted at market exchange rates, it could mean that their output is low or that prices, converted to dollars, are low. How do we distinguish between these possibilities?

One way to do it is to change the way we assign dollar values to goods produced in foreign countries. Instead of measuring their value in local currency and converting to dollars at the market exchange rate, we look up an equivalent good in the US, see its US price and value the foreign goods at their US price. In formulas:

$$\underset{\text{(in dollars, at PPP)}}{\text{GDP in Foreign Country}} = \sum_{i=1}^{N} p_i^{US} \times q_i$$

where N is the number of different goods that we are adding up, p_i^{US} is the market price of good i in the US and q_i is the quantity of good i produced in the foreign country. This is known as the "Purchasing Power Parity" or PPP approach because it aims to adjust for the fact that the purchasing power of a dollar is different in different countries. In practice, PPP calculations are harder to do than converting GDP at market exchange rates: one needs to figure out what US good is the correct equivalent to each foreign good, which is not so easy because the goods available in each country are different. For Mexico, most estimates of PPP put its per capita GDP at around 18,000 dollars, almost twice as high as using market exchange rates, reflecting the fact that goods tend to be cheaper than in the US.

A byproduct of computing GDP at PPP is to define a "PPP exchange rate." This is an answer to the following question: "what would market exchange rates need to be for GDP at market exchange rates and GDP at PPP to coincide?" In formulas:

$$\underset{\text{(in dollars, at PPP)}}{\text{GDP in Foreign Country}} \equiv \underset{\text{(dollars per unit of foreign currency)}}{\text{PPP exchange rate}} \times \underset{\text{(in foreign currency)}}{\text{GDP in Foreign Country}}$$

or

$$\text{PPP exchange rate} \equiv \frac{\text{GDP in dollars at PPP}}{\text{GDP in foreign currency}}$$

If PPP exchange rates and market exchange coincide it means that on average goods cost as much in the foreign country as in the US. For many years, *The Economist* magazine has computed a simple indicator of PPP exchange rates: instead of looking for the exchange rate that would make goods overall cost the same in the US and in foreign countries, they focus on a single good: the Big Mac. This has the advantage of being highly standardized across countries.[2] The Big Mac index is simply

$$\text{Big Mac exchange rate} = \frac{\text{Big Mac price in US (dollars)}}{\text{Big Mac price in Foreign Country (foreign currency)}}$$

[2]Though not fully: the price of Big Macs includes the location, cleanliness, etc. of the McDonald's restaurant and these vary across countries.

Exercises

1.1 **Accounting**

How does GDP accounting record the following events? For each of them, describe how they would be computed in GDP accounts using the income method, the production method and the expenditure method.

(a) A car manufacturer buys components from Japan for \$1 to be used in production later on and stores them at its warehouse.

(b) A car manufacturer buys components from Japan for \$1 and uses half of those components in the production of a car, that it sells to Andy for \$2. It stores the rest of the components.

(c) An army battalion is deployed to the border to repel a threatened Canadian invasion. The soldiers earn wages of \$10,000 and use ammunition that the government buys for \$5,000. The ammunition is produced using \$2,000 of imported steel and 100 hours of work, for which the workers were paid \$1,000.

(d) Walmart sells 1000 bottles of Coca-Cola for \$1,500. It had previously paid \$1,200 for them.

(e) A shipyard builds a cruise ship. It pays wages of \$200,000, interest on loans (from US residents) of \$100,000 and \$300,000 for imported raw materials. The ship is sold for \$1,000,000 to a cruise company. In the same year, the cruise company has revenue for \$50,000 from operating cruises, pays wages of \$20,000 to its workers and has no other expenses. Half the cruise revenue comes from tourists who reside in the United States and half comes from tourists who reside abroad.

(f) The government collects \$1000 in income taxes from Roger.

(g) Roger earns \$4000 for working as a babysitter and pays \$1000 in income taxes.

1.2 **Comparisons Across Time and Across Countries**

Suppose these are the prices (in US dollars) and quantities of goods A and B produced in the US in 2017 and 2018:

	p_A	q_A	p_B	q_B
2017	4	5	3	3
2018	1	10	4	2

(a) What was nominal GDP in 2017?

(b) What was nominal GDP in 2018?

(c) What was real GDP in 2018 at 2017 prices (computed using fixed 2017 prices)? Using this measure, how much did GDP grow between 2017 and 2018?

(d) What was real GDP in 2017 at 2018 prices (computed using fixed 2018 prices)? Using this measure, how much did GDP grow between 2017 and 2018? What explains the difference between the two measures?

(e) How much did GDP grow between 2017 and 2018 using the chain-weighted method?

In Thailand, prices (in Thai baht) and quantities in 2017 were:

	p_A	q_A	p_B	q_B
2017	30	1	10	2

(f) What was nominal GDP in Thailand in 2017, expressed in baht.

(g) Suppose the exchange rate in 2017 was 25 baht per dollar. What was GDP in Thailand in 2017, expressed in US dollars at market exchange rates?

(h) What was GDP in Thailand in 2017 at PPP? What accounts for the difference between the market exchange rate measure and the PPP measure?

(i) What was the PPP exchange rate between the baht and the dollar?

1.3 Chained GDP

The country of Fructus produces Apples, Bananas and Cherries. Its production statistics are given below:

	Apples		Bananas		Cherries	
	quantity	price	quantity	price	quantity	price
2000	100	40	50	30	200	20
2001	103	39	47	35	206	20
2002	107	39	47	38	213	19
2003	109	39	45	38	215	18
2004	112	40	46	41	224	18
2005	112	39	46	39	227	17
2006	114	40	47	36	237	16
2007	115	40	46	42	249	16
2008	119	40	43	44	255	16
2009	118	42	40	46	262	15
2010	116	43	41	44	268	15
2011	118	42	40	50	280	15
2012	117	42	41	50	289	15
2013	119	43	40	53	290	14
2014	120	44	40	56	297	13
2015	125	45	41	59	308	13
2016	128	45	38	66	320	13
2017	129	46	38	61	330	13
2018	131	47	37	64	344	13
2019	136	47	37	61	353	13
2020	143	47	34	61	361	12

(you can download them as an Excel spreadsheet from the book website)

(a) Compute a real GDP series at year-2000 prices using base year prices and using the chained method.

(b) Plot both series and comment on any differences.

(c) What was the average growth rate according to each method?

1.4 **Drugs and Prostitution**
Read the following article: http://www.economist.com/news/finance-and-economics/21603073-italys-inclusion-illicit-activities-its-figures-excites-much-interest-sex. What do you think? Should drug production and prostitution be included in the calculation of GDP?

1.5 **Changes in Relative Prices**
We saw that using an earlier year as the base year to compute real GDP results in a high rate of growth if the sectors that are expanding most are those whose relative price is falling. Can you think of reasons why that should be the case (i.e., economic forces that make the same types of goods become relatively cheaper and be produced in higher quantities)? Can you think of reasons why the opposite should be the case (i.e., economic forces that make the same types of goods become relatively *expensive* and be produced in higher quantities)?

CHAPTER 2

Beyond GDP

GDP is an incomplete indicator of standards of living, and people have realized this for a long time. For instance, here is a speech by Robert Kennedy in 1968:

> Our gross national product, now, is over $800 billion dollars a year, but that gross national product–if we judge the United States of America by that–that gross national product counts air pollution and cigarette advertising, and ambulances to clear our highways of carnage. It counts special locks for our doors and the jails for the people who break them. It counts the destruction of the redwood and the loss of our natural wonder in chaotic sprawl. It counts napalm and counts nuclear warheads and armored cars for the police to fight the riots in our cities. It counts Whitman's rifle and Speck's knife, and the television programs which glorify violence in order to sell toys to our children. Yet the gross national product does not allow for the health of our children, the quality of their education or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages, the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage, neither our wisdom nor our learning, neither our compassion nor our devotion to our country, it measures everything in short, except that which makes life worthwhile. And it can tell us everything about America except why we are proud that we are Americans.

There have been efforts to construct broader measures of living standards that address some of the limitations of GDP.

2.1 The Human Development Index

The United Nations has for some time constructed a measure called the Human Development Index (HDI). The HDI takes as a starting point that, in addition to high output, two other things contribute to "Human Development": a long life and good education. One might argue for the inclusion of all sorts of other things in the index but the ones that are included are not unreasonable. Furthermore, some of the other things that one might consider including ("the beauty of our poetry"?) are much harder to measure.

The HDI is constructed as follows. First, construct indices of each of the three components:

$$\text{Life Expectancy Index} = \frac{\text{Life Expectancy} - 20}{85 - 20}$$

$$\text{Education Index} = \frac{1}{2}\left(\frac{\text{Avg. school yrs. 25-year-olds}}{15} + \frac{\text{Expected schools yrs. 5-year-olds}}{18}\right)$$

$$\text{Income Index} = \frac{\log(\text{GNP per capita}) - \log(100)}{\log(75{,}000) - \log(100)}$$

The logic of the index is to convert each of the three categories into a number between 0 and 1. Life expectancy for the healthiest countries is around 80 years and for the least healthy countries is around 40 years, so the Life Expectancy Index of different countries will range somewhere between 0 and 1. Similarly, children spend somewhere between 0 and 18 years of life in school and GNP per capita ranges between about 400 to $75,000$ US dollars (at PPP).[1]

The HDI is a geometric average of the three indices:

$$\text{HDI} = \sqrt[3]{\text{Life Expectancy Index} \times \text{Education Index} \times \text{Income Index}}$$

Figure 2.1.1 shows a scatterplot of the HDI against GDP per capita. Each observation represents a different country. The correlation between the two measures is 0.94. With few exceptions, countries that have high GDP per capital also have a high HDI.

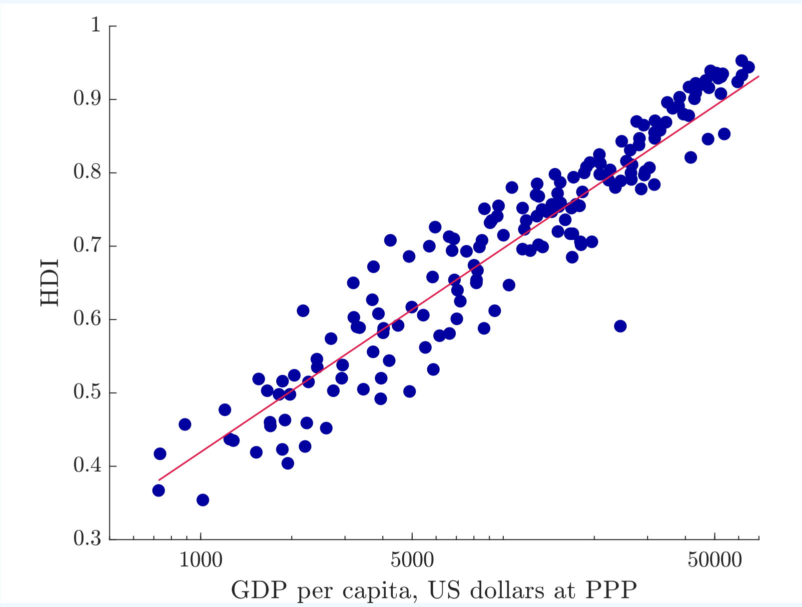


Fig. 2.1.1: *GDP per capita and the HDI for 2017. Source: World Bank for GDP per capita, UN for HDI.*

One unsatisfying aspect of the HDI is that the scaling and weighting of the various factors is somewhat

[1]The HDI uses GNP rather than GDP. The main difference between them is the income that is produced in the country but belongs to nonresidents, like interest paid on debt owed to foreigners. The difference is usually small. Some countries' GNP actually exceeds 75,000 dollars, so their Income Index is greater than 1.

arbitrary. What exactly is the HDI measuring? Why convert variables into indices? Why equal weights on the three factors? We turn next to a study that builds a measure of well-being that is more firmly grounded in economic theory.

2.2 Beyond GDP

Jones and Klenow (2016) propose a measure of living standards that is similar in spirit to the HDI but quite different in methodology. Again, the idea is to construct a measure that includes important aspects of quality of life that GDP does not capture. Jones and Klenow focus on the following variables:

- consumption (as opposed to production),
- the value of leisure and nonmarket production,
- life expectancy,
- inequality.

Defining Living Standards

What do we mean when we talk about "living standards"? We'll define them as the answer to the following question. Suppose you take a random person (Jones and Klenow, who have a sense of humor, call him "Rawls") and invite him to spend a year living as a resident in one of two countries.[2] The rules of the experiment are that, beforehand, Rawls will not know what specific individual within that country he will be: young or old, rich or poor, etc.. The two options are:

- some country whose living standards we are trying to measure, let's call it Utilia,
- a country that is exactly like the US except that everyone's consumption is multiplied by some number λ.

What number λ would make Rawls indifferent between Utilia and the rescaled US? λ will be our measure of living standards in Utilia. You might recall from microeconomics that λ is known as an "equivalent variation." Being a random resident of Utilia is equivalent, in utility terms, to being a random resident of the US and having your consumption multiplied by λ.

In order to measure λ we have to come up with a way to decide *how much* people care about all the variables, other than consumption, that are different in different countries. We are going to assume that Rawls evaluates the choice according to the following utility function:

$$u(c, l, a) = \mathbb{E}\left[\left(\bar{u} + \frac{c^{1-\sigma}}{1-\sigma} - \theta (1-l)^2\right) a\right] \tag{2.2.1}$$

Utility is a function of three variables:

[2] They also analyze a hypothetical scenario where Rawls will spend his whole life in one of these countries, but we'll focus on the single-year experiment.

- c is the level of consumption. As we'll see below, this is a random variable. Different people within both the United States and Utilia consume different amounts. Rawls does not know who he will be so he does not know how much he will end up consuming. The operator $\mathbb{E}$ is there to indicate that we need to compute the expected value of this function of a random variable.
- l stands for leisure: it's the fraction of time that people devote to activities that are not counted in GDP.
- a stands for "alive". This is also a random variable. It takes the value 1 if Rawls turns out to be alive and 0 if he turns out to be dead. Since a multiplies the rest of the formula, it means Rawls will only get utility if he is alive. The chances of this happening depend on the country's life expectancy.

Parameters $\bar{u}$, σ and θ govern how much people care about each of the variables. In order to evaluate utility we are going to have to:

1. Find a way to put concrete numbers for $\bar{u}$, σ and θ. We try to find examples of situations where people actually make choices in which they trade off different goals, showing how much they care about each of them.
2. Find data for many countries on the distribution of consumption, on life expectancy and on leisure.
3. Obtain our measure of welfare for each country by solving for λ in the following equation:

$$u\left(\lambda c^{US}, l^{US}, a^{US}\right) = u\left(c^{Utilia}, l^{Utilia}, a^{Utilia}\right) \tag{2.2.2}$$

Let's look at each issue in turn.

Consumption

GDP is a measure of production, not consumption. But what enters the utility function (2.2.1) is consumption. The logic is that what determines people's standards of living is how much they get to consume, not how much they are able to produce. The output that Utilia dedicates to investment will not give Rawls any utility during the year he stays in Utilia. This doesn't mean it's wasted: it will benefit future residents of Utilia after Rawls has left. Similarly, the output that Utilia exports will not give Rawls any utility, and, conversely, the consumption goods that Utilia imports *will* give Rawls utility.

A slightly subtler question is how to treat government consumption. As we saw in Chapter 1, government purchases of goods and services are counted the same regardless of how much people actually value them. It could be that on average a unit of public goods gives Rawls a lot more utility than a unit of private goods (he really likes to feel protected by police officers), or that it gives him less. We are going to assume that on average public and private consumption give the same utility. In terms of the GDP accounting identity (1.1.1) we'll assume that what goes into the utility function is $C + G$.[3]

[3] To be more precise, G includes both government consumption and government investment. We'll assume that the utility function values C plus the part of G that is government consumption.

The Veil of Ignorance, Inequality and Risk Aversion

Philosophers refer to a hypothetical choice that people must make before knowing what place in society they are going to occupy as a decision made "behind the veil of ignorance." The philosopher John Rawls famously argued that the way to determine what is just is to ask what sort of society people would organize from behind the veil of ignorance.[4]

From behind the veil of ignorance, Rawls faces risk. He knows that, once he enters the experiment to spend a year in Utilia, he might turn out to be rich and enjoy high consumption or poor and have low consumption. How do we model his attitude towards these possibilities? Suppose there are N people in Utilia. Rawls knows how much each inhabitant of Utilia consumes and he knows that his experience in Utilia is going to look like one of them, but he does not know which one. Therefore, behind the veil of ignorance, he measures his expected utility by:

$$u(c,l,a) = \sum_{n=1}^{N} \frac{1}{N} \left(\bar{u} + \frac{c_n^{1-\sigma}}{1-\sigma} - \theta\left(1-l\right)^2 \right) a \tag{2.2.3}$$

where c_n denotes the consumption of the nth resident of Utilia. Rawls believes there is a probability $\frac{1}{N}$ of him ending up in each point of the distribution, so he just computes the average utility he will get across all of these possibilities.[5] This way of modeling attitudes towards risk is known as "expected utility theory" because it says that people evaluate uncertain prospects according to the expected utility that they will obtain.

Under this theory, Rawls's attitude towards risk is related to the concavity of the function $u\left(c,l,a\right)$. Before we go into the maths of why that is, an important warning: the important economic idea here is that (most) people don't like risk. Concave functions are just a mathematical representation of this idea. It's wrong to say "people dislike risk because their utility function is concave". Instead, one should say: "in economic models, we describe people's preferences with concave utility functions to capture the fact that people dislike risk". Do not put the mathematical cart before the conceptual horse!

Now let's look at the maths of utility functions and risk aversion. Hold l and a constant for now (for instance, because there is no uncertainty about them) and imagine that u is just a function of c. Suppose that there are two people in Utilia, one is rich and the other is poor, with consumption c_{rich} and c_{poor} respectively. Therefore, from Rawls' point of view, there are two possible "states of the world": one in which he is rich and one in which he is poor. We say that Rawls is "risk averse" if

$$u\left(\frac{c_{rich}+c_{poor}}{2}\right) > \frac{u\left(c_{rich}\right)+u\left(c_{poor}\right)}{2} \tag{2.2.4}$$

What does this mean? Suppose someone offered Rawls insurance. Instead of consuming c_{rich} in one state of the world and c_{poor} in the other, Rawls gets to consume the average $\frac{c_{rich}+c_{poor}}{2}$ for sure. That would give him

[4] Rawls also proposed an answer to the question of what people would choose behind the veil of ignorance, which he called the "difference principle": society would be organized in whatever way benefits the least-well-off person. You'll see in Exercise 2.10 that, according to standard economic theory, the difference principle would follow from behind-the-veil-of-ignorance choices only if people were extremely risk averse. Interestingly, Rawls himself was adamant that the difference principle had nothing to do with risk aversion.

[5] Implicit in formula is the assumption that all inhabitants of Utilia enjoy the same leisure. Otherwise we would have to have l_n in the formula instead of the same value l for everyone. This is obviously not correct but it's hard to obtain data on inequality in leisure, so assuming it's the same for everyone is a simplification.

the utility in the left-hand-side of (2.2.4). Instead, in the actual world where he faces risk, his expected utility is the right-hand-side of (2.2.4). Therefore (2.2.4) just says that Rawls, in expected-utility terms, would prefer a world without risk. Generalizing from the example, Rawls is risk averse if, whenever c is uncertain,

$$u\left(\mathbb{E}\left(c\right)\right) > \mathbb{E}\left(u\left(c\right)\right) \tag{2.2.5}$$

Inequality (2.2.5) holds whenever u is a concave function. Figure 2.2.1 illustrates this general principle. $\mathbb{E}\left(c\right)$ is the midpoint between c_{Rich} and c_{Poor} and $u\left(\mathbb{E}\left(c\right)\right)$ is just evaluating the function u at this point. $\mathbb{E}\left(u\left(c\right)\right)$ is the midpoint between $u\left(c_{Rich}\right)$ and $u\left(c_{Poor}\right)$. Since the function u is concave, $\mathbb{E}\left(u\left(c\right)\right)$ lies below $u\left(\mathbb{E}\left(c\right)\right)$.

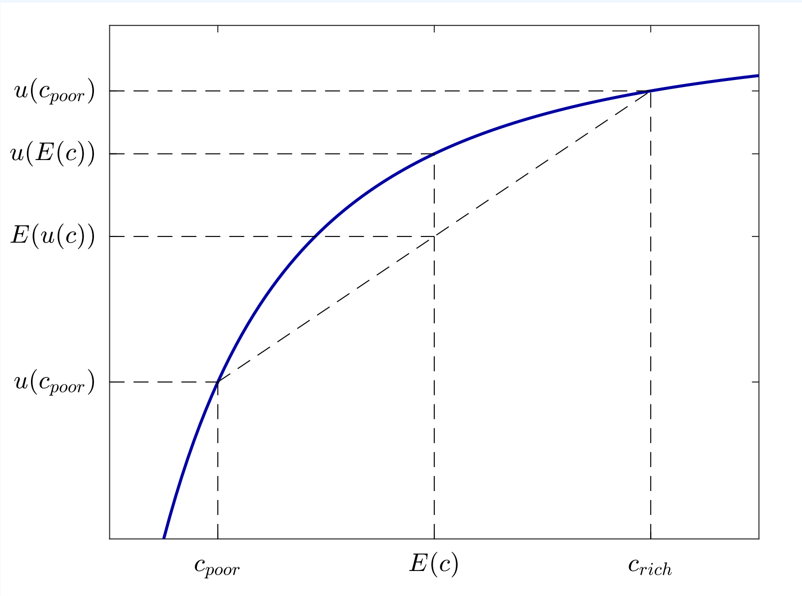


Fig. 2.2.1: *Concavity of the utility function and risk aversion.*

In function (2.2.1), the concavity of u is governed by parameter σ: when σ is high the function is very concave, when σ approaches 0 the function is close to linear. If we take the derivative of utility with respect to consumption, we get a formula for marginal utility:[6]

$$u'\left(c\right) = c^{-\sigma} \tag{2.2.6}$$

Figure 2.2.2 plots marginal utility for different values of σ. For all values of $\sigma > 0$, we have that marginal utility is positive but decreasing, but it's decreasing faster for higher values of σ. This gives us another way

[6]This assumes that $a = 1$, i.e. it's the marginal utility of consumption for those who are alive. Notice that if $\sigma < 1$, then $\frac{c^{1-\sigma}}{1-\sigma}$ is a positive number and if $\sigma > 1$, then $\frac{c^{1-\sigma}}{1-\sigma}$ is a negative number. Whether utility is positive or negative doesn't mean much because we don't really have an interpretation of what the *level* of utility means: we just care about what a utility function says about how people compare different alternatives. Regardless of whether $\sigma > 1$ or $\sigma < 1$, *marginal* utility is positive. For $\sigma = 1$, the function $\frac{c^{1-\sigma}}{1-\sigma}$ is not well defined. However, its properties approach those of the function $\log\left(c\right)$. For instance, formula (2.2.6) works fine if we just adopt the convention that when $\sigma = 1$ the function becomes $\log\left(c\right)$.

of thinking about the relationship between risk aversion and the shape of the utility function. An individual who faces risk will have a lot of consumption in some states of the world and less consumption in others. If the utility function is very concave, this means that the difference in marginal utility between high and low consumption states of the world will be large and the individual would have a strong preference to make consumption more even between the different states of the world, i.e. the individual will be very risk averse.

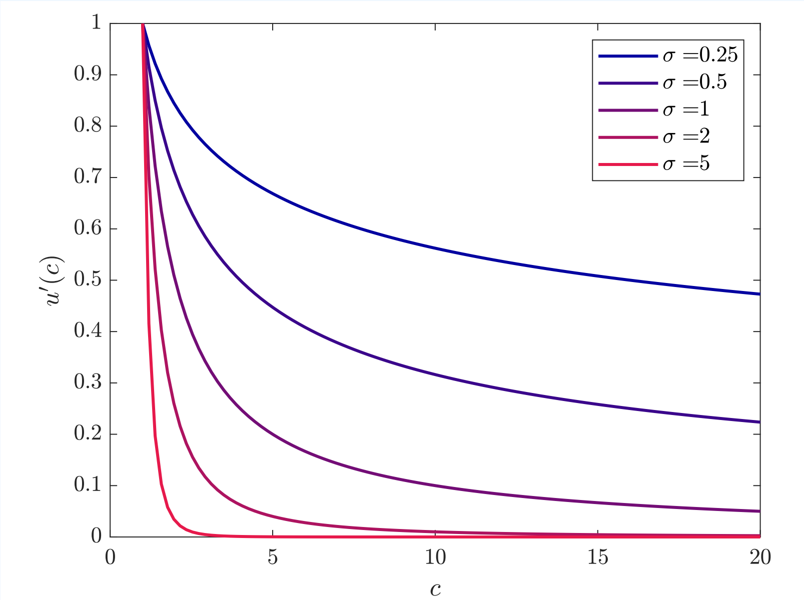


Fig. 2.2.2: *Marginal utility schedule for different values of σ.*

So far we've established that different values of σ can be used to represent different attitudes towards risk. The next step is to decide what value of σ captures people's actual attitudes towards risk. One obvious caveat to this analysis is that different people have different attitudes towards risk, so at best we will find a value of σ that roughly represents the behavior of some average person.

How can we measure people's attitudes towards risk? We need to find environments where people are actually trading off higher average consumption for more risk. Two situations in which people make this sort of decision are in making financial investments and buying insurance.

On average, risky investments like the stock market give higher returns than safer investments like US government bonds, so someone who invested their wealth in risky investments would obtain higher *average* consumption than someone who chose safer investments. However, there are states of the world where risky investments turn out poorly and lead to very low consumption. One way to measure people's attitude towards risk is to look at their investments: to what extent are they willing to bear risk in exchange for a higher average consumption? Several studies have measured σ by doing exactly that. Friend and Blume (1975) were among the first to do so. We'll think more about risky investments in Chapter 8.

When someone buys insurance, they are moving consumption across different states of the world. Take the example of fire insurance. There is a state of the world in which my house burns down. If I had no insurance, in that state of the world I would have to lower consumption in order to pay for the repairs on my house.

An insurance contract lets me pay the insurance company some money in the state of the world where my house is fine in order to get the insurance company to pay me when my house burns down. Typically, buying insurance makes my average consumption go down because on average the premium I pay is higher than the benefits I collect: that's how the insurance company covers administrative costs and makes profits. However, if I'm risk averse I'll still be willing to buy insurance: it lets me consume more in those states of the world where I need it the most. One way to measure people's attitudes towards risk is to measure the extent to which they decide to buy insurance, an approach used by Szpiro (1986) among others.

There is a fair amount of disagreement about what the right value of σ is; estimates range from about 1 to about 10 (Jones and Klenow use $\sigma = 1$ as a baseline, near the less-risk-averse end of the range of empirical estimates). Part of the reason why estimates of σ differ is that people's attitudes towards risk depends to some extent on the context where they are making this choice. Indeed, some researchers argue that this means the expected utility model of risk attitudes is unsatisfactory.

In terms of comparing standards of living across countries, σ is what determines how much Rawls cares about inequality. From Rawls' point of view, a very unequal country is risky. If Rawls is very risk-averse (as represented by high σ), a more egalitarian society may look more attractive to him, behind the veil of ignorance, than a more unequal society that is richer on average.

The Value of Life Expectancy

Behind the veil of ignorance, Rawls does not know how old he'll be. Let's assume his age is a number randomly drawn from 0 to 100. If his age turns out to be higher than life expectancy in the country he's going, he'll be dead; otherwise he'll be alive. According to (2.2.1), if he turns out to be dead he'll get zero utility and if he turns out to be alive he'll get $\bar{u} + \frac{c^{1-\sigma}}{1-\sigma} - \theta\,(1-l)^2$. In this formula, everyone who is alive gets $\bar{u}$ in addition to however much utility they get from c and l. Therefore the value of $\bar{u}$ governs how attractive it is to live in a society with high life expectancy.

How can we measure people's preferences for high life expectancy? We need to find environments in which people trade off years of life against higher consumption. We can find evidence on this in how much people are willing to pay for safety features in cars or in the extra money that people demand in exchange for doing dangerous jobs. In these situations people have to choose between lower consumption but probably a longer life (get a car with airbags, work as a librarian) or higher consumption but probably a shorter life (don't pay for airbags, work as a drug dealer). By observing what choices people make at various prices we get a sense of how they are willing to trade these off.

The Value of Leisure and Non-Market Production

There are many alternative uses of people's time. Some result in output that is counted in GDP and some do not. But even activities that are not counted in GDP can contribute to utility. We saw some examples of this in Chapter 1: cleaning one's own house, cooking for friends or taking care of one's own children are all non-market activities that nevertheless produce something valuable. Moreover, people also enjoy time spent in pure leisure activities like reading books or watching TV. In equation (2.2.1), the variable l stands for the fraction of time that people, on average, spend on all these non-market activities, so $1 - l$ is the fraction of time spent at a job counted in GDP. The parameter θ governs how much people dislike working in the market sector (notice that there's a negative sign) or, equivalently, how much utility they derive from the time they spend in non-market activities.

How can we measure θ? We need to find instances of people trading off higher consumption against more free time. A direct source of evidence on this is in people's choices of how much to work: at what age they enter the labor force, when they retire, how many holidays they take, how many hours per week they work, etc.[7] Under the preferences given by (2.2.1), higher values of θ imply that people will choose to work less and have more non-market time. Since we can measure how much time people on average spend on market and non-market activities, we can estimate θ by asking what the value has to be to match our empirical observations.

Data

Once we've settled on values for $\bar{u}$, σ and θ, we need to get actual data from Utilia to plug into formula (2.2.1). Ideally, we would need to have data on:

1. the consumption of every individual in Utilia,
2. the average fraction of time that residents of Utilia spend working,[8]
3. life expectancy in Utilia.

Life expectancy is the easiest, because most countries measure it relatively reliably. The consumption of every individual, of course, is impossible to know. However, many (but not all) countries conduct surveys where they ask a lot of households about their income or their consumption. These surveys are not always as accurate as we would like, but at least they give a rough estimate of what the distribution of consumption looks like. As to the fraction of time worked, the quality of data varies by country. Some countries have detailed time-use surveys while others just report employment rates but not hours of work per employed person.

Results

Figure 2.2.3 shows a scatterplot of GDP per capita relative to the US on the horizontal axis and λ, obtained using formula (2.2.2), on the vertical axis. Clearly the two measures are very highly correlated, although not identical. This means that just looking at GDP per capita as a measure of welfare is not so bad (or it could be that some other variable that is not included in the derivation of λ is important).

There are some interesting patterns. Some Western European countries like France look better in the welfare measure than in GDP per capita. This is because they have higher life expectancy, more leisure, and less inequality than the US. Rich East Asian countries like South Korea, Hong Kong, and especially Singapore look worse in terms of welfare than in GDP per capita. This is mostly because they have low consumption relative to GDP: they produce a lot but dedicate a large fraction to investment and net exports. This is also true of oil-rich countries like Kuwait and Saudi Arabia (which are also very unequal). Many Sub-Saharan

[7] This is valid as long as we believe that people are actually choosing how much to work. We'll return to this when we study labor markets in Chapter 7.

[8] Recall that we are assuming that everyone gets the average amount of leisure.

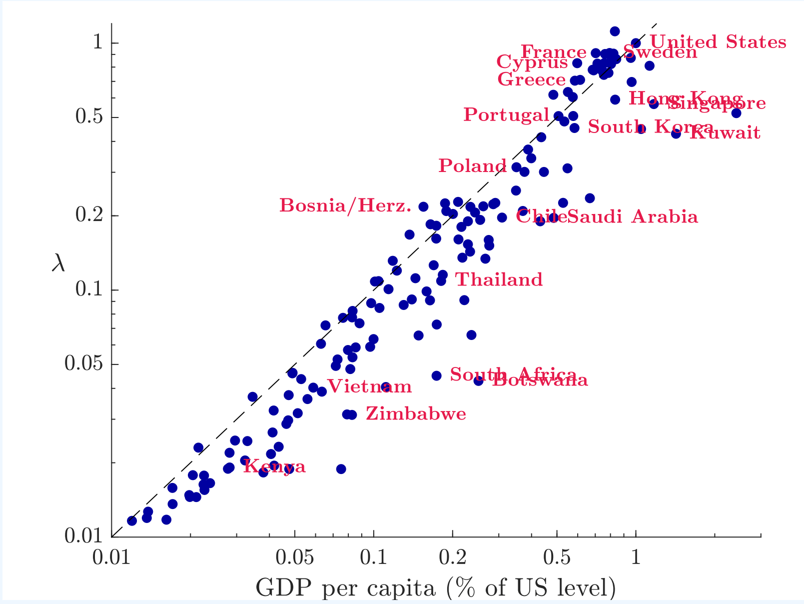


Fig. 2.2.3: *Relative GDP and welfare. Source: Jones and Klenow (2016).*

African countries look worse in welfare than in GDP per capita, especially some not-so-poor ones like South Africa and Botswana. In large part this is due to low life expectancy (which itself is the result of the AIDS epidemic), though inequality plays a role as well.

Exercises

2.1 **Comparing the HDI and the Jones & Klenow Welfare Measure**

Download the UN data that goes into building the HDI from http://hdr.undp.org/en/data and the Jones & Klenow Data from http://web.stanford.edu/~chadj/papers.html#rawls. For each country, construct:

(a) Relative GDP:

$$\frac{\text{GDP per capita}}{\text{GDP per capita in the US}}$$

(b) Welfare-to-GDP:

$$\frac{\lambda}{\text{Relative GDP}}$$

(c) HDI-to-GDP:

$$\frac{\text{HDI}}{\text{Relative GDP}}$$

and plot a scatterplot of Welfare-to-GDP against HDI-to-GDP. What do each of these ratios measure? Is the impression we get from the Jones & Klenow measure very different from the one we get from the HDI? What countries look better under each measure and why?

2.2 Other Things that Matter

The welfare measure λ takes into account data on consumption levels, inequality, leisure and mortality.

(a) Name one other variable that might be an important determinant of welfare that is not included in standard GDP calculations.

(b) What data would you need in order to figure out how much weight to give to this variable? Describe how you would use that data to come up with the correct way to include the variable in question in the welfare calculation. What choices that people make might reveal how much they care about this variable? Don't worry too much about the feasibility of the data-collection procedure, but think carefully about why the observed choices would be informative.

2.3 Healthcare

When we calculate consumption, one of the (many) categories of consumption is healthcare services.

(a) Look up how much healthcare is consumed in the United States per year for a recent year. State the total dollar amount and also what fraction of GDP and what fraction of consumption is accounted for by healthcare.

A good place to look for this data is the National Income and Product Accounts (NIPA) at https://apps.bea.gov/iTable/index_nipa.cfm. Browse around a little to get a sense of how the NIPA data is presented, and find the correct place to look up this particular fact.

(b) Suppose that we are calculating welfare in the style of Jones and Klenow, taking into account the impact of life expectancy on utility. Should we therefore subtract consumption of healthcare services from our measure of consumption? What do you think?

2.4 Construction and the Armed Forces

Suppose we compared welfare in two otherwise identical countries. Country A dedicates 50% of GDP to maintaining a large army. Country B dedicates 50% of GDP to building houses. In which of the two countries would welfare be higher, according to the measure used by Jones and Klenow. What do you think of this?

2.5 Measuring the Value of Life Expectancy

Would each of the following observations be useful in determining how much to weigh life expectancy in measured welfare? Explain.

(a) Observing how life expectancy varies across people with different income levels within a country.

(b) Observing how life expectancy varies across countries.

(c) Observing how much people are willing to pay for funerals.

(d) Observing the difference in house prices in neighborhoods with different murder rates.

2.6 The Value of Life Expectancy

In the US, average consumption per capita is about \$35,000. Life expectancy is 79 years, so in the Jones and Klenow experiment, Rawls would have a 79% chance of being alive. In Bolivia, average consumption per capita is $3,700 and life expectancy is 68 years. Assume that there is no inequality in either country so that everyone who is alive gets the average level of consumption. Assume the following utility function:

$$u(c, a) = \left[\bar{u} + \frac{c^{1-\sigma}}{1-\sigma}\right] a$$

with $\sigma = 0.5$ and $\bar{u} = 48.1$.

(a) Suppose you made a US resident the following offer: you can buy a health plan that costs x dollars per year and will extend your lifetime by one year. What is the price x that would make them indifferent between getting the health plan or not?

(b) How much would a Bolivian resident be willing to pay for such a plan?

(c) Write down the the special case of equation (2.2.1) that applies to this example and solve for λ for Bolivia. Don't replace any numbers yet, leave it in terms of a_{US}, a_{Bol}, c_{US}, c_{Bol}, σ and $\bar{u}$.

(d) Replace the values of a_{US}, a_{Bol}, c_{US}, c_{Bol}, σ and $\bar{u}$ to find a value for λ. How does it compare to $\frac{c_{Bol}}{c_{US}}$? Why?

2.7 Measuring Risk Aversion

The students of Concave University are all identical in their ability and preferences, which are well described by the function:

$$u(c) = \mathbb{E}\left[\frac{c^{1-\sigma}}{1-\sigma}\right]$$

Two employers recruit CU graduates. Stabilis offers each graduate a fixed salary of $50,000. Lotteris instead offers them a base salary of $20,000 plus a bonus scheme that depends on their performance, which is entirely driven by luck. The bonus is either $10,000 or $100,000, with equal probability. The job itself is the same in both firms. About half the CU graduates choose to work for Stabilis and half prefer Lotteris, and they all say they found it hard to choose because both offers were similarly attractive. What value of σ is consistent with their decisions?

2.8 Inequality and Risk Aversion

Compare the following two countries. Both have a population of 100. Within each country, we label individuals in order of increasing consumption. In country A, the consumption of individual j is:

$$c^A(j) = 100 + 8j$$

while in country B the consumption of individual j is:

$$c^B(j) = 200 + 4j$$

(a) Plot the consumption patterns of each country, with an individual's label j (which ranges from 1 to 100) on the horizontal axis and their consumption on the vertical axis.

(b) Compute per capita consumption in each country. [Note: if you want, you can approximate sums with integrals]

(c) Suppose the utility function in both countries is:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

For what values of σ is expected utility higher in country A? Interpret your results.

2.9 **Risk Aversion and Welfare in Western Europe**

If instead $\sigma = 1$, Jones and Klenow had set a higher value of σ, what would this do to measured welfare in Western European countries? Explain.

2.10 **Risk Aversion and the Difference Principle**

Suppose a individual is trying to evaluate a society from behind the veil of ignorance. He knows that he can either be rich or poor, with equal probability. Expected utility is

$$\mathbb{E}[u(c)] = \frac{u(c_{Rich}) + u(c_{Poor})}{2}$$

where the utility function u is

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

Let's now define an indifference curve. For any utility level $\bar{U}$, the indifference curve plots all the combinations of c_{Rich} and c_{Poor} such that expected utility is $\bar{U}$.

(a) Take the following equation that defines an indifference curve:

$$\bar{U} = \frac{1}{2}\frac{c_{Rich}^{1-\sigma}}{1-\sigma} + \frac{1}{2}\frac{c_{Poor}^{1-\sigma}}{1-\sigma}$$

and solve for c_{Rich}. This will tell you what value of c_{Rich} corresponds to each value of c_{Poor} in the indifference curve with level $\bar{U}$.

(b) Plot (using Excel, Matlab or some other software that produces nice graphs) a map of indifference curves for the following values of σ:

i. $\sigma = 0.5$. Plot three curves, for $\bar{U} = 2$, $\sqrt{10}$ and 4 respectively

ii. $\sigma = 2$. Plot three curves, for $\bar{U} = -1$, -0.4 and -0.25 respectively.
Note that when $\sigma > 1$ utility values will be negative, but that's OK.
Also note that

$$\bar{U} > \frac{1}{2}\frac{c_{Poor}{}^{1-\sigma}}{1-\sigma}$$

is impossible, i.e. even if $c_{Rich} = \infty$ expected utility will be lower than $\bar{U}$, so when you are plotting this indifference curve, whenever you set c_{Poor} to the left of $\left[2\bar{U}(1-\sigma)\right]^{\frac{1}{1-\sigma}}$ you should set $c_{Rich} = \infty$.

iii. $\sigma = 5$. Plot three curves, for $\bar{U} = -0.25$, $-\frac{4}{5^4}$ and -4^{-5} respectively.

iv. $\sigma = 11$. Plot three curves, for $\bar{U} = -0.1$, $-\frac{4^{10}}{10^{11}}$ and $-\frac{4^{-10}}{10}$ respectively.

In all cases, have c_{Poor} on the horizontal axis and c_{Rich} on the vertical axis and let the range of the horizontal axis be $[0, 10]$ and the range of the vertical axis be $[0, 10]$.

(c) Notice that as σ becomes large, the indifference curves start to look like right angles. Explain how this relates to Rawls's "difference principle."

PART II

Economic Growth

This part of the book looks at economic growth.

In Chapter 3, we start by reviewing some facts about long-run economic growth.

In Chapter 4, we study the Solow growth model, which provides a theory of how and why economic growth takes place.

In Chapter 5, we look at how the model can be used to make sense of the evidence and how the evidence can be used to test, quantify and apply the lessons from the model.

CHAPTER 3

Basic Facts about Economic Growth

3.1 The Very Long Run

Let's start by looking at the distant past. Following Maddison (2001), a number of economists have attempted to measure GDP per capita for many countries going back hundreds of years. Of course, this is very hard and involves quite a bit of guesswork, but we have some clues. First of all, there is some minimum level of consumption (the "subsistence level") below which people starve, so we know that in all societies that didn't starve GDP per capita must have been at least that. Estimates of how much that is vary, but they are in the order of about 400 dollars a year at current prices, close to what is nowadays considered extreme poverty. Beyond this, before we had proper national accounts, we had pieces of data on things like people's average height (from bones), the total number of livestock, crop yields or total output of specific industries like iron that can be used to piece together rough estimates of GDP.

The left panel of Figure 3.1.1 shows the evolution of GDP per capita in the UK in the very long run.[1] We focus on the UK because it has the best data but also because it was the first country to show fast and sustained economic growth, starting in the early 19th century. The first fact that emerges from Maddison's data is that, even before 1800, GDP per capita in the UK was well above subsistence levels, and growing slowly. These are somewhat controversial points among economic historians, some of whom believe GDP per capita was stagnant and closer to subsistence. The second fact that emerges is that something happened in the 19th century that led to an acceleration in the rate of economic growth (on this there is less disagreement). This change is known as the "Industrial Revolution" since one of the things that took place at the time was a shift in production from agriculture to industry. We don't have a definitive answer as to what caused the the industrial revolution and why it first took place in the UK in the 19th century, but it's a major question among economic historians.

Other countries went through a similar process of an acceleration in the rate of economic growth. They started from levels of GDP per capita not too far from subsistence and, at different initial dates, started

[1] You'll notice that most of the graphs in this section are in log scale. What this does is convert proportional differences into absolute differences: the vertical distance between $1,000$ and $2,000$ is the same as the vertical distance between $10,000$ and $20,000$. Hence a constant proportional rate of growth (e.g. 2% per year) looks like a straight line, and the slope of this line indicates the rate of growth. If we plotted GDP in regular scale, then a constant rate of growth would look like an exponential function.

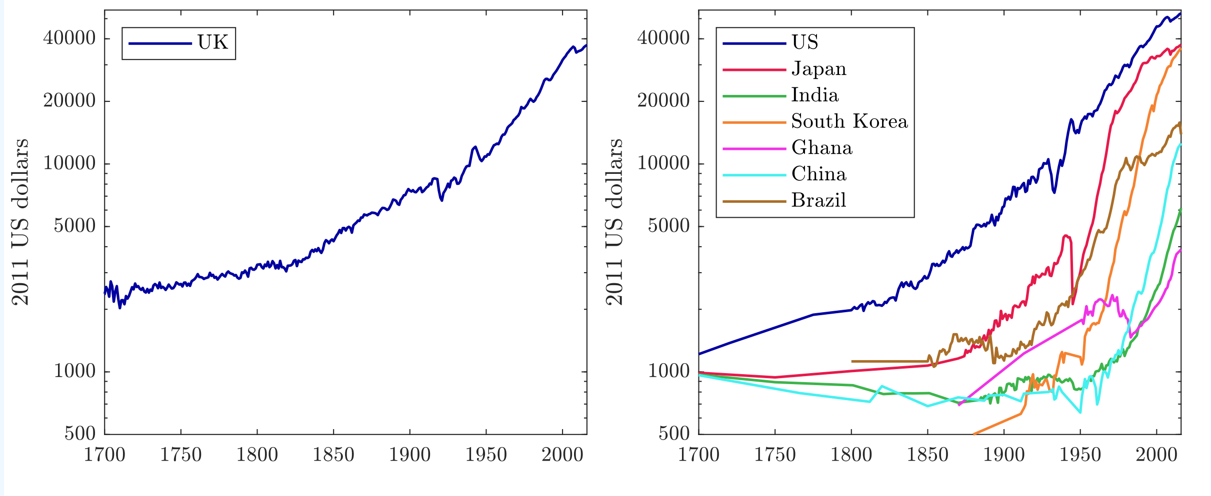


Fig. 3.1.1: *GDP per capita in the UK and selected countries. Source: Bolt et al. (2018).*

growing. The right panel of Figure 3.1.1 shows some examples.

3.2 The Kaldor Facts

We turn now to some patterns that we observe in economies that are growing. Kaldor (1957) summarized some of the main facts about economic growth in advanced economies. He called them "remarkable historical constancies" and they became known as the "Kaldor Facts". Let's have a look at some of those facts and ask whether they are still approximately true, focusing on the US.[2]

1. The rate of growth of GDP per capita is constant.

   Figure 3.2.1 shows the evolution of GDP per capita in the US from 1800 to 2016. A straight line (in log scale) seems to do quite well in describing how the US economy has grown for many decades. GDP per capita has been growing at a rate of approximately 1.5% per year for a long time. Notice that while 1.5% doesn't seem like a lot, compounded over time it amounts to a huge increase in GDP per capita. From 1800 to 2016, GDP per capita has become about 27 times higher.

2. The ratio of the total capital stock to GDP is constant

   The capital stock is the sum of the value of all the machines, buildings, etc. that are currently available for use in production. It's not an easy thing to measure. A standard way to do it is by keeping track of investment and depreciation over time. Exercise 5.4 asks you to think more about this. Figure 3.2.2 shows the evolution of $\frac{K}{Y} = \frac{\text{Capital Stock}}{\text{GDP}}$ over time. Consistent with the Kaldor Facts, this ratio has

[2]Kaldor described them slightly differently and in different order

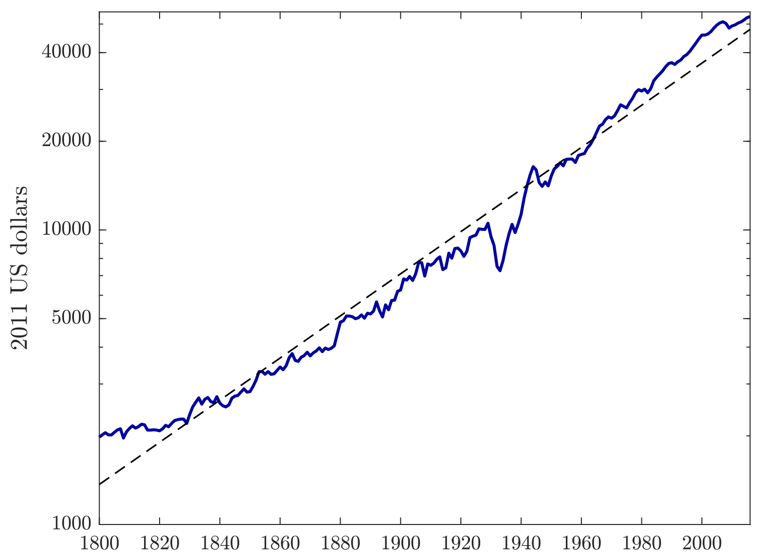


Fig. 3.2.1: *GDP per capita in the US. Source: Bolt et al. (2018).*

remained more or less constant over time at about 3.2. This means that the total value of all the capital that the US economy has accumulated is about the same as the economy produces in 3.2 years.

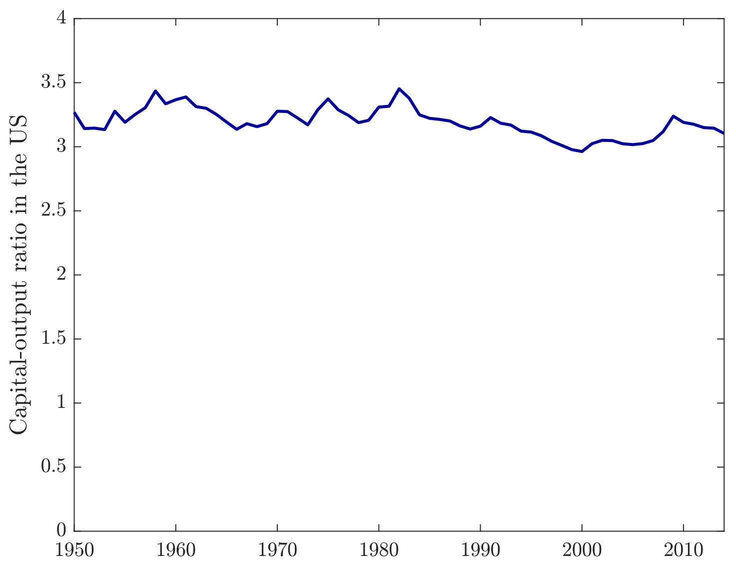


Fig. 3.2.2: *Capital-to-output ratio in the US. Source: Feenstra et al. (2015).*

3. The shares of labor and capital income in GDP are constant

Recall from Chapter 1 the income method of measuring GDP. Let's take a simplified view of the types of income and classify them into just two categories: labor income (that is earned for work done in the current period) and capital income (that is earned by those who own some form of capital). Some forms of income are easy to classify: workers' wages are labor income, corporate profits and real estate rents are capital income. Others are a little bit trickier: is the income earned by small business owners a reward for the work they do or for the investment they put into the business? One way of addressing the issue is to leave "proprietors' income" our of the calculation entirely, which is equivalent to assuming that the split between labor and capital income is the same in the sole proprietor sector as in the rest of the economy. This method is not entirely satisfactory but it is often adopted. Figure 3.2.3 shows how the labor share of GDP in the US has evolved over time. Until about 2000 or so, it seemed that this fact continued to hold: the share of labor income in GDP was very stable at around 65%. More recently, there has been a noticeable fall in this percentage: the share of GDP going to workers has fallen by about 3 percentage points.

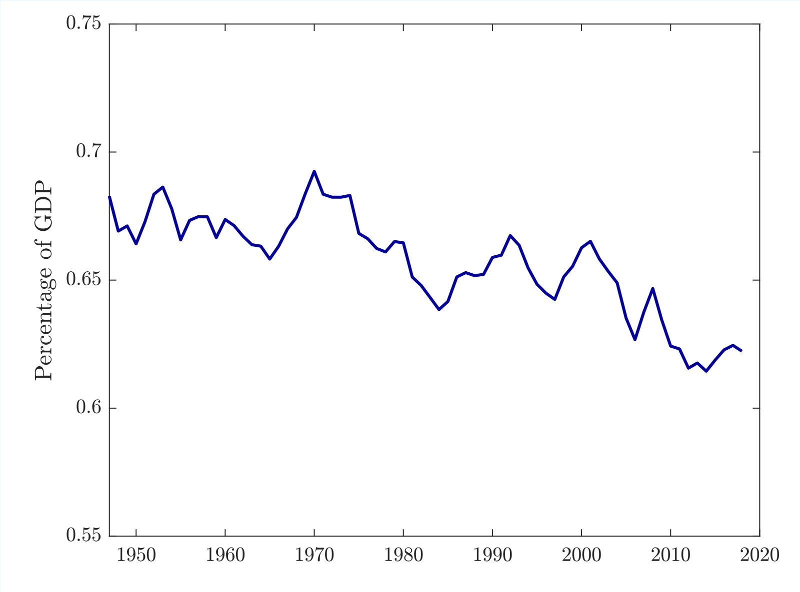


Fig. 3.2.3: *The labor share of GDP in the US. Labor income is Compensation of Employees. Capital income is Corporate Profits + Rental Income + Interest Income + Depreciation. Source: NIPA.*

4. The average rate of return on capital is constant

   By the rate of return on capital we mean how much income is earned by the owner of capital per unit of capital that they own. This fact says that this has stayed constant over time. Strictly speaking, it's not a separate fact since it's implied by facts 2 and 3. Let's see why:[3]

$$\text{Return on capital} \equiv \frac{\text{Capital income}}{\text{Capital stock}}$$
$$= \frac{\frac{\text{Capital income}}{\text{GDP}}}{\frac{\text{Capital stock}}{\text{GDP}}}$$

[3] This refers to the *gross* rate of return on capital. To get the *net* rate of return we should subtract depreciation from capital income. If the depreciation rate is constant, the net return on capital will be constant too.

Fact 2 says that the denominator is constant and fact 3 says the numerator is constant, so if these facts are true then the return on capital must be constant too. Nevertheless, we state it as a separate fact because the behavior of the rate of return on capital over time is an important aspect of many theories of economic growth and it's useful to keep this fact in mind.

3.3 Growth Across Countries

Figure 3.3.1 shows the growth of countries of different initial income levels since 1960. Initially-rich countries have growth rates that are quite similar to each other and near the middle of the range of other countries. Among initially-poor countries there is a lot more variation. Some countries like South Korea (KOR) and Botswana (BWA) have very rapid rates of growth, so that over time they are catching up to the living standards in rich countries while some others like Congo (COD) and Madagascar (MDG) have very low or even negative growth rates, meaning that they are falling further behind rich countries.

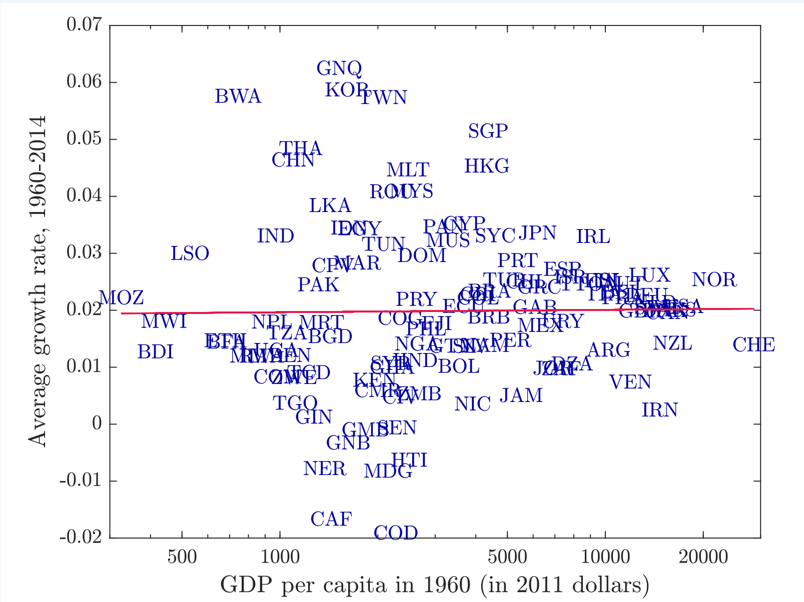


Fig. 3.3.1: Growth across countries since 1960. Source: Feenstra et al. (2015).

Exercises

3.1 **The Past is a Foreign Country**

Find the data compiled by Bolt et al. (2018) at https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2018. Starting from 1800, look up the US GDP per capita at intervals of one decade. For each of these points in time, find the country that currently has the closest GDP per capita to the past US level. When did the US become richer than present-day India, present-day China, and present-day Portugal?

3.2 The Kaldor Facts in Other Countries

Look up GDP accounts for some countries other than the US (a good source is https://www.rug.nl/ggdc/productivity/pwt/). Reproduce Figures 3.2.1-3.2.3 using data from those countries (try a rich country, a middle-income country and a poor country). Do the Kaldor facts seem to hold for those countries as well?

3.3 Within-Region Convergence

Look up cross-country GDP data (a good source is https://www.rug.nl/ggdc/productivity/pwt/). For each country, compute average GDP-per-capita growth 1960-2014. Produce a scatterplot of growth against initial GDP per capita like Figure 3.3.1 but separately for countries in each of three regions of the world: Europe, Latin America and Africa. Is it the case in any of these regions that initially-poor countries have grown faster?

CHAPTER 4

The Solow Growth Model

Solow (1956) proposed a simple model that can help us to start to think about the process of economic growth.

4.1 Ingredients of the Model

Production Function

The first ingredient of the model is a production function:

$$Y = F\left(K, L\right) \tag{4.1.1}$$

Formula (4.1.1) says that the output of any productive process (denoted Y) depends on:

- K: the amount of capital (machines, buildings, etc.) dedicated to the production process.
- L: the amount of labor that is dedicated to the production process.

One way to interpret a production function is as a book of recipes: for any given combination of ingredients, it says how much stuff will be produced.

Example 4.1. One page of the recipe book says:

$$F\left(\underbrace{\text{100 acres of Iowa land} + \text{1 tractor}}_{K=\$800{,}000}, \underbrace{1,000 \text{ hours of work}}_{L=1{,}000}\right) = \underbrace{18,000 \text{ bushels of corn}}_{Y=\$66{,}000}$$

Another page of the book says:

$$F\left(\underbrace{\text{1 garage in Palo Alto}}_{K=\$150{,}000}, \underbrace{3,000 \text{ hours of work by Stanford dropouts}}_{L=3000}\right) = \underbrace{\text{1 app}}_{Y=\$100{,}000}$$

We are going to assume that everyone in the country knows the production function; anyone can set up a firm, hire L workers and K units of capital and obtain $F(K, L)$ units of output. Furthermore, we are going to make the following assumptions about the production function:

Assumption 4.1 (Constant Returns to Scale)**.**

$$F(\lambda K, \lambda L) = \lambda F(K, L) \text{ for any } \lambda > 0$$

The standard justification for assuming constant returns to scale is that production processes can, at least approximately, be replicated. If I have a factory that produces paint and I want to produce twice as much paint, I build a replica of the original factory next to it, hire replicas of all the workers and I'm done. Obviously, there are many objections to this argument. Maybe some factors of production (natural resources, workers with specific skills) are not easily replicable: this would push towards having decreasing returns to scale. Alternatively, one could imagine that one large factory can be run more efficiently than two small ones because not everything needs to be exactly duplicated: this would push towards having increasing returns to scale. We are going to stick with the assumption of constant returns to scale: any productive process can be exactly scaled up or down by increasing or decreasing the use of capital and labor in the same proportion. We'll see that this assumption has profound implications.[1]

Assumption 4.2 (Positive Marginal Product)**.**

$$F_K(K, L) > 0$$
$$F_L(K, L) > 0$$

Assumption 4.2 has a straightforward interpretation: adding additional workers or additional capital to a productive process adds at least a little bit to total output.

Assumption 4.3 (Diminishing Marginal Product)**.**

$$F_{KK}(K, L) < 0$$
$$F_{LL}(K, L) < 0$$

Assumption 4.3 says that adding just one of the factors (workers without extra machines or machines without extra workers) becomes less and less useful the more you do it.

Assumption 4.4 (Inada Conditions)**.**

1. $\lim_{K \to 0} F_K(K, L) = \infty$
2. $\lim_{K \to \infty} F_K(K, L) = 0$

[1] We adopt the following notation for partial derivatives: $F_K(K, L) \equiv \frac{\partial F(K,L)}{\partial K}$, $F_{KK}(K, L) \equiv \frac{\partial^2 F(K,L)}{\partial K^2}$

Assumption 4.4 is slightly more technical. It says that if there is very little capital then a little bit of capital is extremely useful. Conversely, if there is a lot of capital then additional capital becomes almost useless (because there are no workers to operate the additional machines). It's similar in spirit to Assumption 4.3 (diminishing marginal product) although mathematically one does not imply the other.

We'll see the role that each assumption plays later on.

One example of a production function that we'll often resort to is the so-called Cobb-Douglas production function, shown in Figure 4.1.1:

$$Y = K^{\alpha} L^{1-\alpha} \tag{4.1.2}$$

It's easy to verify that this satisfies Assumptions 4.1-4.4. We'll see that parameter α in this formula has a natural interpretation.

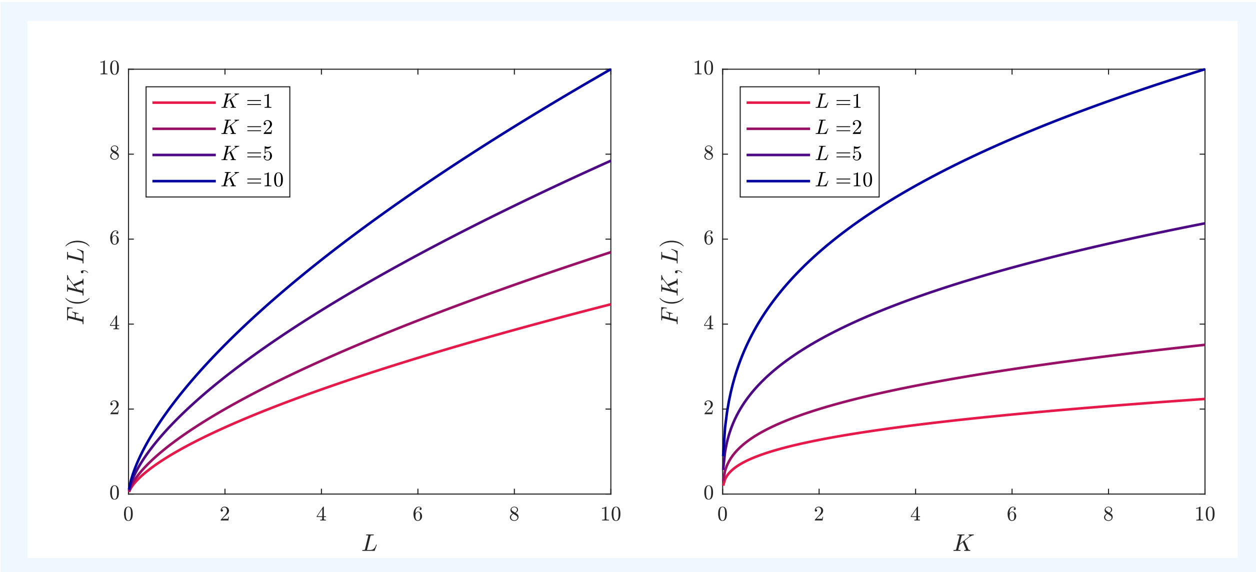


Fig. 4.1.1: *The Cobb Douglas production function for $\alpha = 0.35$.*

Population and the Labor Supply

Assumption 4.5. The population grows at a constant, exogenous rate n: $L_{t+1} = (1+n)\,L_t$

Nowadays Assumption 4.5 is routinely made in a lot of work on economic growth but it's actually a very big deal. Historically, the possibility that population growth might be endogenous and depend on living standards (as is the case for wild animal populations) was a central preoccupation among economists. Exercise 4.5 asks you to examine some of the ideas of the 19th century economist Thomas Malthus who wrote about this issue.

We are not going to model people's decisions over how much to work. We are going to assume that everyone who is alive works, so L represents both the population and the labor force. When looking at data, it is sometimes important to distinguish between GDP per capita and GDP per worker, but we are not going to make this distinction for now. In Chapter 7 we'll go back to thinking about what incentives govern the decision over how much to work.

Consumption and Investment

Assumption 4.6. The economy is closed and there is no government

Assumption 4.6 implies that in the accounting identity (1.1.1), $X = M = G = 0$, so we are left with

$$Y = C + I$$

This means that all output is either dedicated to consumption or to investment.

Assumption 4.7. The savings rate $\frac{Y-C}{Y}$ is equal to an exogenous constant s.

Recall that by definition Y is both total output and total income. Therefore $S \equiv Y - C$ represents savings: all the income that people choose not to consume. $\frac{S}{Y} = \frac{Y-C}{Y}$ is the savings rate: savings as a fraction of income. Assumption 4.7 says that $\frac{S}{Y} = s$: people save an exogenous fraction s of their total income. In Chapter 6 we are going to think more about the incentives that shape people's decision of whether to consume or save but for now we are going to take this decision as exogenous.

An immediate consequence of Assumptions 4.6 and 4.7 is

$$I = sY \tag{4.1.3}$$

so a fraction s of output is dedicated to investment. There are actually two steps in getting to formula (4.1.3). Assumption 4.6 implies that $S = I$: savings equal investment. This is always true in a closed economy.[2] The second step uses Assumption 4.7: if savings are are a constant fraction of income then investment is a constant fraction of output.

Depreciation and Capital Accumulation

Capital depreciates. Machines wear down, computers become outdated, buildings need repairs, etc. If we want to keep track of the total capital stock it's important to keep this in mind.

[2] In fact, this does not depend on having $G = 0$. Suppose that we have $G > 0$ and the government collects τ in taxes (which may or may not be equal to G). Then private savings are $S^{Private} = Y - \tau - C$ (after-tax income minus consumption) and public savings are $S^{Public} = \tau - G$ (tax revenues minus spending). Then total savings are

$$\begin{aligned} S &= S^{Private} + S^{public} \\ &= Y - \tau - C + \tau - G \\ &= Y - C - G \\ &= I \end{aligned}$$

Assumption 4.8. The capital stock depreciates at a constant rate δ.

We are going to model depreciation in the simplest possible way: every piece of capital equipment loses a fraction δ of its value every period. Therefore the total capital stock is going to evolve according to:

$$K_{t+1} = (1-\delta) K_t + I_t \tag{4.1.4}$$

Equation (4.1.4) says that if the capital stock in this period is K_t, then the capital stock in the next period will consist of the sum of:

1. The portion of the capital stock that has *not* depreciated: $(1-\delta) K_t$ and
2. The new capital that has been created through investment I_t

4.2 Mechanics of the Model

Per Capita Production Function

First we are going to rewrite the production function in per-capita terms. Define

$$y \equiv \frac{Y}{L}$$
$$k \equiv \frac{K}{L}$$

y denotes GDP per capita and k denotes capital per worker.[3] Then using (4.1.1) we can write

$$\begin{aligned} y &= \frac{F(K,L)}{L} \\ &= F\left(\frac{K}{L}, 1\right) \\ &\equiv f(k) \end{aligned} \tag{4.2.1}$$

The first step in (4.2.1) is just using the production function (4.1.1) to replace Y. Implicitly, what we are doing is saying that all the capital and all the labor in the economy is used in one aggregate production process. Thanks to Assumption 4.1 (constant returns to scale), it wouldn't make a difference if we instead assume that it's split up into many different production processes that are just a scaled-down version of the aggregate economy. The second step uses Assumption 4.1 (constant returns to scale) directly: we are just multiplying by $\lambda = \frac{1}{L}$. The last step is a definition: we are defining the "per-capita production function" $f(k)$ as the output that would be produced by one worker with k units of capital. Equation (4.2.1) says that GDP per capita is going to the same as it would be if there was only one person and k units of capital in the economy.

[3] Recall, since everybody works, "per capita" and "per worker" is the same in this model.

Dynamics of Capital per Worker

Using (4.1.4) we can derive a formula for how the amount of capital per worker k is going to evolve over time:

$$
\begin{aligned}
\Delta k_{t+1} &\equiv k_{t+1} - k_t && \text{(definition)} \\
&= \frac{K_{t+1}}{L_{t+1}} - k_t && \text{(replacing } k_{t+1} \text{ with } \frac{K_{t+1}}{L_{t+1}}\text{)} \\
&= \frac{(1-\delta)\,K_t + I_t}{L_{t+1}} - k_t && \text{(using 4.1.4)} \\
&= \frac{(1-\delta)\,K_t + sY_t}{L_{t+1}} - k_t && \text{(using 4.1.3)} \\
&= \left[\frac{(1-\delta)\,K_t + sY_t}{L_t}\right] \frac{L_t}{L_{t+1}} - k_t && \text{(rearranging)} \\
&= \left[(1-\delta)\,k_t + sy_t\right] \frac{1}{1+n} - k_t && \text{(using Assumption 4.5: constant } n\text{)} \\
&= \frac{sy_t - (\delta+n)\,k_t}{1+n} && \text{(rearranging)} \\
&= \frac{sf\,(k_t) - (\delta+n)\,k_t}{1+n} && \text{(using 4.2.1)} \qquad (4.2.2)
\end{aligned}
$$

Formula (4.2.2) has the following interpretation. The change in the stock of capital per worker k depends on the balance of opposing forces. Investment adds to the capital stock, pushing it up. The term $sf\,(k_t) = \frac{I}{L}$ is investment per capita. Two forces push k down. The first is depreciation, which directly subtracts from the capital stock. The second is population growth. This doesn't literally subtract from the capital stock but spreads the capital stock over a larger number of workers, so it also lowers the stock of capital per worker. That's why δ and n appear together in formula (4.2.2). Figure 4.2.1 plots $f(k)$, $sf\,(k)$ and $(\delta+n)\,k$ on the same graph.

Notice that both output per capita $f(k)$ and investment per capita $sf\,(k)$ are concave functions of k. Why is that? Mathematically, it follows from Assumption 4.3 (diminishing marginal product): the derivative of the production function is positive but decreasing and $sf\,(k)$ is just multiplying by a constant, so it inherits the same properties. Economically, what's going on is that the marginal product of capital is decreasing: adding more and more machines per worker to the economy results in higher output (and therefore investment) per worker but at a diminishing rate.

Notice also that $sf\,(k)$ starts above $(\delta+n)\,k$ but ends below, i.e. the lines cross. That is a consequence of Assumption 4.4 (Inada conditions). The first part of this assumption implies that at first $sf\,(k)$ is very steep, so it must be above $(\delta+n)\,k$. The second part of the assumption says that eventually the slope of $sf\,(k)$ becomes zero. Since the slope of $(\delta+n)\,k$ is $\delta+n$, this means that for sufficiently high k the slope of $sf\,(k)$ is lower than $\delta+n$, and therefore eventually $sf\,(k) < (\delta+n)\,k$.

What does this imply? Whenever $sf\,(k) > (\delta+n)\,k$, then equation (4.2.2) says that the capital stock per worker is growing. Conversely, when $sf\,(k) < (\delta+n)\,k$, equation (4.2.2) says that the capital stock per worker is shrinking. Economically, this means that an economy with a sufficiently low k will be accumulating capital while an economy with a sufficiently high k will tend to deplete its stock of capital. The reason for this

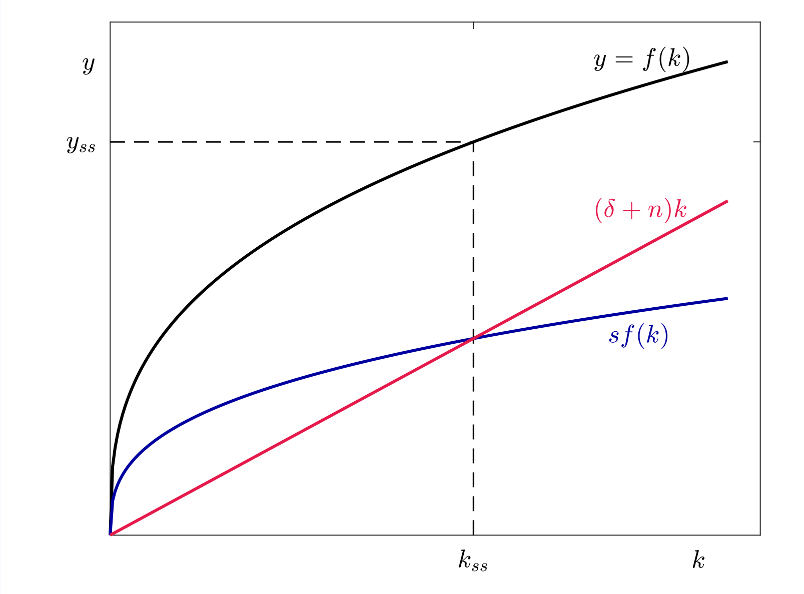


Fig. 4.2.1: *The forces that govern the evolution of the level of capital per worker in the Solow model.*

is that the relative magnitude of two forces pushing Δk in opposite directions changes with k. Depreciation (and dilution via population growth) is just proportional: the more capital there is, the more it depreciates. Investment is proportional to *output*, not to the capital stock. Due to the diminishing marginal product of capital, the increase in output and therefore investment that you get out of a higher capital stock is smaller and smaller as the capital stock increases. If Assumption 4.4 (Inada conditions) holds, eventually the extra investment is less than the extra depreciation, so the two lines cross.

The point k_{ss} is the level of capital-per-worker such that the two forces are exactly equal. If $k = k_{ss}$, then the capital stock per worker will remain constant from one period to the next. We refer to an economy where $k = k_{ss}$ as being in *steady state*. Notice that in a steady state output per worker also remains constant at its steady state level $y_{ss} = f(k_{ss})$.

For any $k < k_{ss}$ we have that $sf(k) > (\delta+n)k$, so k grows, and for any $k > k_{ss}$ we have that $sf(k) < (\delta+n)k$, so k shrinks. This implies that over time k moves closer and closer to the steady state.[4] Therefore over time the economy *converges* to the steady state. Mathematically:

$$\lim_{t\to\infty} k_t = k_{ss}$$

and therefore

$$\lim_{t\to\infty} y_t = f(k_{ss})$$

Therefore, over time, *the rate of growth of GDP per capita will slow down to zero.*

Notice one subtlety about terminology. When we say that in steady state the economy is not growing,

[4]This is not quite a rigorous mathematical argument. We haven't ruled out the possibility that k could jump across k_{ss} from one period to the next without really getting closer. But in fact it's straightforward to rule out this possibility.

what we mean is that k is not growing and y is not growing. However, L is growing. Therefore $K = kL$ and $Y = yL$ are also growing. Economically, this means that both GDP and the capital stock are growing but just enough too keep up with the growing population. GDP *per capita* is not growing.

In the case of a Cobb-Douglas production function, we can find an expression for k_{ss} explicitly:

$$\begin{aligned}
y_t &= k_t^{\alpha} && \text{(dividing 4.1.2 by } L_t\text{)} \\
\Delta k_{t+1} &= \frac{sk_t^{\alpha} - (\delta + n)\,k_t}{1+n} && \text{(replacing into 4.2.2)} \\
0 &= \frac{sk_{ss}^{\alpha} - (\delta + n)\,k_{ss}}{1+n} && \text{(applying the definition of a steady state)} \\
k_{ss} &= \left(\frac{s}{\delta + n}\right)^{\frac{1}{1-\alpha}} && \text{(rearranging)} \\
y_{ss} &= \left(\frac{s}{\delta + n}\right)^{\frac{\alpha}{1-\alpha}} && \text{(using the production function)}
\end{aligned}$$

Growth and Level Effects

Let's imagine that an economy is at its steady state and there is some change in its fundamental features. We can ask the model what the consequences of this will be.

Suppose first that there is an increase in the savings rate. It's often said that increasing investment (which in a closed economy is the same as saving) is desirable. Let's see what would happen in this model.

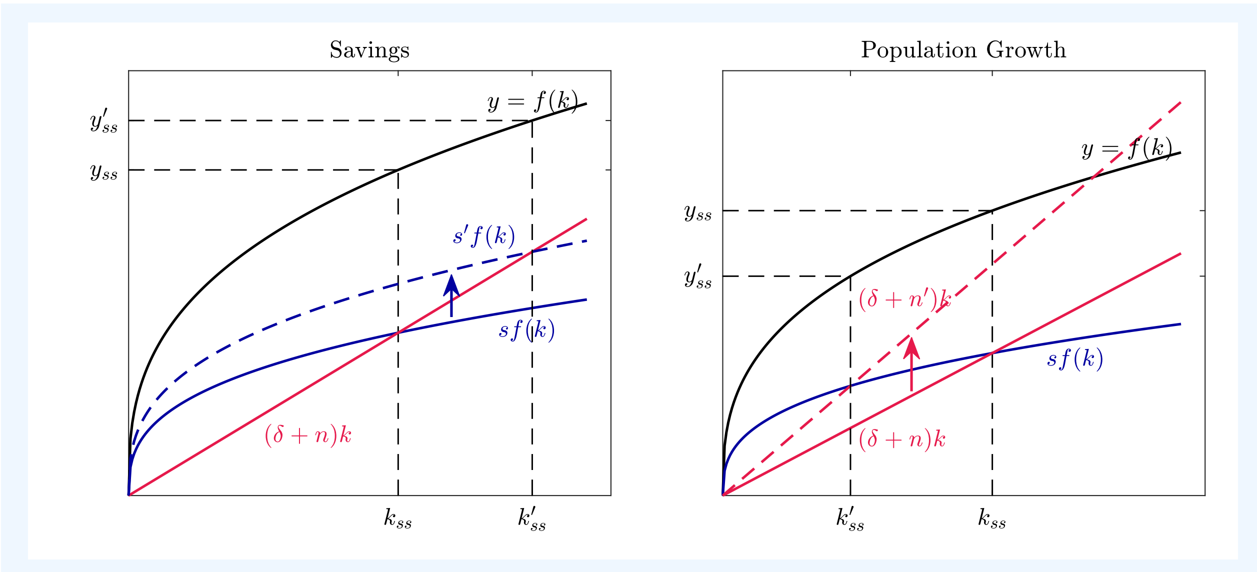


Fig. 4.2.2: *Increases in the savings rate and the rate of population growth.*

Graphically, we can represent an increase in the savings rate as a proportional upward shift in the $sf(k)$ curve, as shown in the left panel of Figure 4.2.2. Starting from the original k_{ss}, we now have that $sf(k) >$

$(\delta + n)\,k$, so the economy will begin to accumulate capital. Eventually, it will converge to a new steady state with a higher capital stock per worker k'_{ss} and higher output per worker y'_{ss}.

Suppose now that there is an increase in the rate of population growth. Graphically, this is represented by an increase in the slope of $(\delta + n)\,k$, as shown in the right panel of Figure 4.2.2. Starting from the original k_{ss}, we now have $sf\,(k) < (\delta + n)\,k$ so the stock of capital per worker will start to go down and eventually converge to a steady state with less capital per worker and lower output per worker.

Suppose now that some new technological discovery results in a change in the production function: we figure out a way to get more output out of the same amount of inputs. Graphically, we can represent this as an upward shift in $f\,(k)$ and therefore in $sf\,(k)$, as shown in Figure 4.2.3. Starting from the original k_{ss}, we now have that $sf\,(k) > (\delta + n)\,k$, so the economy will begin to accumulate capital. Eventually, it will converge to a new steady state with a higher capital stock per worker. Output per worker in the new steady state will be higher for two reasons: k is higher (a shift along the horizontal axis) and there is a vertical shift in the production function for any k.

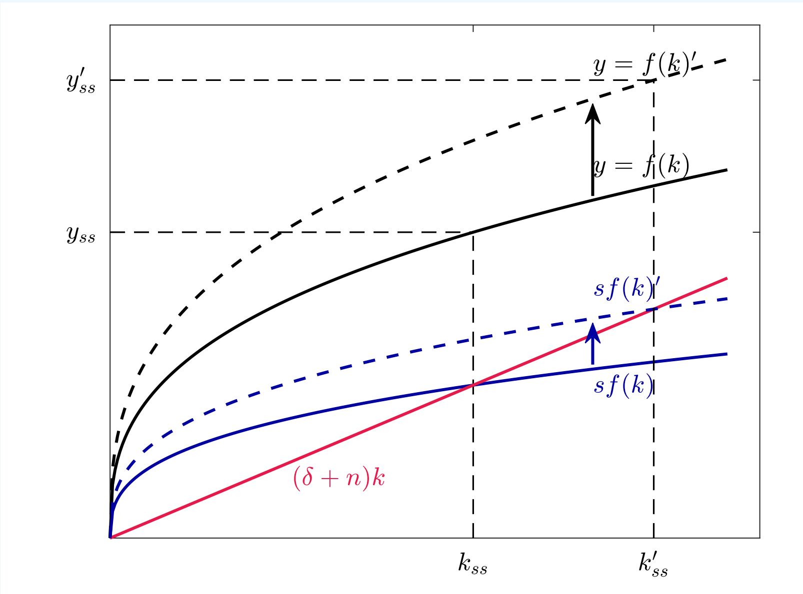


Fig. 4.2.3: *An upward shift in the production function.*

One thing that all these examples have in common is that none of these changes will result in sustained long-term growth. There will be growth (or negative growth) for a while as the economy moves towards a new steady state but this will be temporary: the long-term effect will be on the *level* of GDP per capita but not on the long term *growth rate* of GDP per capita, which is always zero.

4.3 The Golden Rule

Suppose we were to ask the question: how much *should* the economy save? This is a normative question and in order to answer it we have to have some sort of standard to make normative judgments. We'll consider one possible criterion here and revisit it later in Chapter 9. The so-called Golden Rule criterion is a very loose interpretation of the moral principle "one should treat others as one would like others to treat oneself". Applied to the question of the savings rate, it can be thought to mean that societies should save in such a way as to maximize the level of consumption in the steady state. Whether this is a good interpretation of the moral principle is more of a literary question than an economic one, but let's accept it for now. One justification for this objective is that if you were going to be born into a society that *is and will remain in steady state*, the Golden Rule society will be the one where you achieve the highest utility.

Finding the Golden Rule

If the economy is at a steady state, consumption will be

$$c_{ss} = (1 - s)\, y_{ss} \tag{4.3.1}$$

c_{ss} depends on s in two ways. First, there is a direct effect: the more you save, the less you consume. That's why s appears negatively in (4.3.1). Then, there is an indirect effect: the more you save, the higher the steady state capital stock, the higher the output out of which you can consume.

We are going to restate the question of the Golden Rule a little bit. Instead of thinking about choosing s, let's think about choosing k_{ss}. Why does this make sense? We know that changing s will change the amount of capital the economy ends up with in steady state (that's the point of Figure 4.2.2), so we can simply think about choosing some level of k_{ss} and then ask what s is needed to bring about this k_{ss}.

Let's start with this last step. Suppose we have decided on some level k_{ss} that we would like the economy to have in steady state. How much does the economy need to save to make this happen?

$$\begin{aligned} sf(k_{ss}) &= (\delta + n)\, k_{ss} && \text{(imposing steady state in (4.2.2))} \\ s &= \frac{(\delta + n)\, k_{ss}}{f(k_{ss})} && \text{(solving for } s\text{)} \end{aligned} \tag{4.3.2}$$

Equation (4.3.2) says that the savings rate that the economy needs to have to ensure a certain level of steady-state capital-per-worker is equal to the ratio of depreciation-plus-population-growth $(\delta + n)\, k_{ss}$ to total output $f(k_{ss})$. This has a simple interpretation: the economy needs to save at a rate that is sufficient to make up for the amount of depreciation and population growth that will take place at k_{ss}.

Using (4.3.2) and (4.3.1) we can obtain an expression for c_{ss} as a function of k_{ss}:

$$\begin{aligned} c_{ss} &= \left(1 - \frac{(\delta + n)\, k_{ss}}{f(k_{ss})}\right) f(k_{ss}) \\ &= f(k_{ss}) - (\delta + n)\, k_{ss} \end{aligned} \tag{4.3.3}$$

Equation (4.3.3) says that in steady state the economy will consume everything that is left over after making up for depreciation and population growth. Now we want to find the level of k_{ss} that will maximize this expression. Taking first order conditions, we obtain that the Golden Rule capital stock k_{gr} must satisfy:

$$f'(k_{gr}) = \delta + n \tag{4.3.4}$$

The Golden Rule capital stock is such that, at the margin, the additional output you get from having more capital exactly equals the extra investment that will be required to maintain it. Note that it's possible for an economy to be at a steady state with $k_{ss} > k_{gr}$. If this happens, this economy will have *higher* GDP per capita but *lower* consumption per capita than a Golden Rule economy .

For the Cobb-Douglas production function, we have

$$\alpha k_{gr}^{\alpha-1} = \delta + n$$

so

$$k_{gr} = \left(\frac{\delta+n}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

and replacing this in (4.3.2) simplifies to

$$\begin{aligned} s &= \frac{(\delta+n)\left(\frac{\delta+n}{\alpha}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{\delta+n}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}} \\ &= \alpha \end{aligned}$$

4.4 Markets

So far we've followed a mechanical approach: factors of production get inputed in the production function and output comes out. We haven't said anything about how this comes about: who makes what decisions and why. Now let's imagine that decisions are made by firms and workers that interact in markets.

Factor Markets

We are going to imagine that there are perfectly competitive markets for labor and capital. The labor market is straightforward to conceptualize (although we might not be persuaded that perfect competition is a good assumption). There is a wage w. This means that workers get paid w goods per unit of labor that they provide. Perfect competition means that a firm can hire as many units of labor as it wants at a wage w and workers get paid w per unit for however many units of labor they supply (we will maintain that they supply L inelastically). In equilibrium, w must be such that firms want to hire exactly the L units of labor that workers supply.

The way we are going to model the market for capital is a little bit less intuitive. We are going to assume that firms do not own the capital they use; instead they rent it from the households. This makes less of a difference than you might think: ultimately, the households own the firms so either directly or indirectly they own the capital. But it's conceptually useful to make the distinction between ownership and use of capital. Therefore we are going to assume that all capital is owned by households and rented by firms. A capital rental arrangement works as follows. The firm gets the right to use one unit of capital for one period. In exchange, the firm pays a rental rate r^K and, at the end of the period, returns the capital to its owner, with the understanding that it will have depreciated a little bit in the meantime. Again, we are going to assume perfect competition: firms can rent as much capital as they want at a rental rate r^K and r^K must be such that they choose to rent exactly the amount K that is available in the economy.

The Firm's Decision

Anyone in the economy can set up a firm, hire labor and capital and use the production function. We'll assume that the objective of firms is to maximize profits. Mathematically, this means that firm i solves the following problem:

$$\max_{K_i, L_i} F(K_i, L_i) - wL_i - r^K K_i$$

Profits are equal to the output the firm produces $F(K_i, L_i)$ minus the wL_i that it pays for L_i units of labor and the $r^K K_i$ that it pays for K_i units of capital. The first order conditions of this maximization problem are:

$$F_K(K_i, L_i) - r^K = 0 \tag{4.4.1}$$

$$F_L(K_i, L_i) - w = 0 \tag{4.4.2}$$

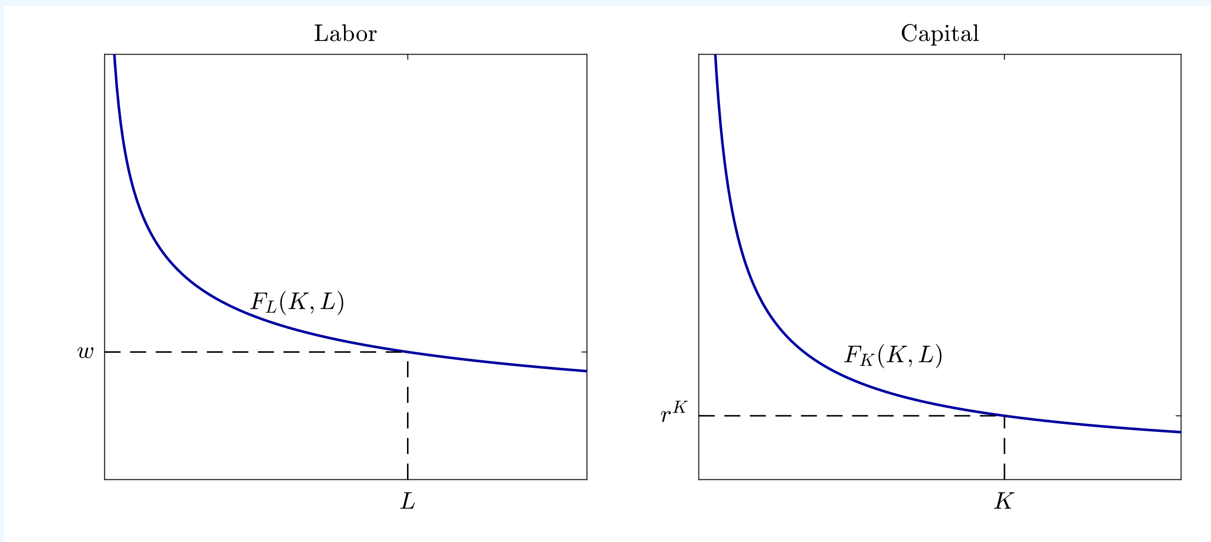


Fig. 4.4.1: *A firm's choice of capital and labor.*

Equation (4.4.1) has the following interpretation. Suppose the firm is considering whether to hire an additional unit of capital. If it does, this will produce extra output equal to the marginal product of capital $F_K(K_i, L_i)$ and it will cost the firm the rental rate r^K. If the difference $F_K(K_i, L_i) - r^K$ were positive, it would be profitable for the firm to hire more capital; if the difference were negative, the firm would increase profits by reducing the amount of capital it hires. Only if (4.4.1) holds is the firm satisfied with the amount of capital it hires. The interpretation of equation (4.4.2) is the same but with respect to labor: only if it holds is the firm satisfied with the number of workers it hired. Figure 4.4.1 illustrates this reasoning.

Market Clearing

In what follows, we'll rely on the following mathematical results.

Proposition 4.1 (Euler's Theorem)**.** The production function satisfies:

$$F_K(K,L)\,K + F_L(K,L)\,L = F(K,L) \tag{4.4.3}$$

Proof. Assumption 4.1 says:

$$F(\lambda K, \lambda L) = \lambda F(K,L)$$

Take the derivative with respect to λ on both sides:

$$F_K(\lambda K, \lambda L)\,K + F_L(\lambda K, \lambda L)\,L = F(K,L)$$

Setting $\lambda = 1$ gives (4.4.3) □

Proposition 4.1. The production function satisfies:

$$F_K(\lambda K, \lambda L) = F_K(K,L) \quad \text{for all } \lambda > 0 \tag{4.4.4}$$
$$F_L(\lambda K, \lambda L) = F_L(K,L) \quad \text{for all } \lambda > 0 \tag{4.4.5}$$

Proof. Assumption 4.1 says:

$$F(\lambda K, \lambda L) = \lambda F(K,L)$$

Take the derivative with respect to K on both sides:

$$F_K(\lambda K, \lambda L)\lambda = \lambda F_K(K,L)$$

which implies (4.4.4). Similarly, taking the derivative with respect to L on both sides leads to (4.4.5). □

Now we'll use Proposition 4.1 to show that all firms in the economy will use factors in the same proportions.

Setting $\lambda = \frac{1}{L_i}$ in (4.4.4) and $\lambda = \frac{1}{K_i}$ in (4.4.5) respectively we obtain:

$$F_K(K_i, L_i) = F_K\left(\frac{K_i}{L_i}, 1\right) \tag{4.4.6}$$

$$F_L(K_i, L_i) = F_L\left(1, \frac{L_i}{K_i}\right) \tag{4.4.7}$$

Equation (4.4.6) has the following interpretation. Suppose firm i chooses to hire K_i units of capital and L_i workers. The marginal product of capital will depend on the ratio $\frac{K_i}{L_i}$ but not on the absolute values of K_i and L_i. A firm with a lot of capital per worker will have a low marginal product of capital no matter how much of each factor it has in absolute terms. The key assumption that drives this result is Assumption 4.1: constant returns to scale. Equation (4.4.7) has the symmetric interpretation: the marginal product of labor also depends only on the ratio of factors of production. To put it in concrete terms, suppose we are running an orchard, which uses apple trees (a form of capital) and workers to produce apples. Equation (4.4.7) says that the number of extra apples we'll obtain if a worker spends an extra hour picking apples depends on how many hours per apple tree we are starting from but not on whether the farm is large or small.

The profit-maximization conditions (4.4.1) and (4.4.2) say that each firm is equating the marginal product to factor prices. Since they all face the same prices, they must all have the same marginal product. This in turn implies that all firms choose the same ratio of capital to labor. The only difference between different firms is their scale, but with constant returns to scale there is no difference between having many small firms or one large firm that operates the same technology. Therefore we can assume without loss of generality that there is just one representative firm that does all the production. Market clearing requires that the representative firm hire all the available labor L and all the available capital K. Replacing $K_i = K$ and $L_i = L$ in (4.4.1) and (4.4.2) implies:

$$r^K = F_K(K, L) \tag{4.4.8}$$

$$w = F_L(K, L) \tag{4.4.9}$$

The rental rate of capital will be equal to the marginal product of capital for a firm that hires all the capital and all the workers in the economy; the wage will be equal to the marginal product of labor for the same firm.

Using (4.4.6) and (4.4.7), this implies that the rental rate of capital and wages and will be:

$$r^K = F_K(k, 1) \tag{4.4.10}$$

$$w = F_L\left(1, \frac{1}{k}\right) \tag{4.4.11}$$

By Assumption 4.3, this means that, other things being equal, the rental rate of capital will be low and wages will be high in an economy with high k. If there is a low level of capital per worker, then the marginal product of capital will be low and so will the rental rate. Conversely, with a lot of capital to work with, the marginal product of labor will be high, so competition between firms will drive up wages.

Profits

We can now compute the profits of the representative firm

$$\text{Profit} = F(K, L) - wL - r^K K$$

Proposition 4.2. The representative firm earns zero profits

Proof. Replace $F_K(K, L)$ and $F_L(K, L)$ in (4.4.3) using (4.4.8) and (4.4.9):

$$r^K K + wL = F(K, L)$$

which gives the result □

Proposition (4.2) says that all the output that is produced gets paid either to the workers or to the owners of capital, with no profits left over for the owner of the firm. Even though they are trying to maximize profits, the maximum level of profits that the firms can attain is zero.

It's very important to remember that the definition of "profits" that we are using is different from the way the term is used in accounting. Let's see an example.

Example 4.2. A corporation called Plantain Monarchy owns a retail space on the ground floor of a building and runs a clothes shop in this retail space. In 2018, Plantain Monarchy sold clothes worth $1,000,000, paid $300,000 in wages and spent $500,000 buying clothes from manufacturers. The rent on a comparable retail space is $200,000 a year. What were the profits of Plantain Monarchy? Accountants would measure the profits of Plantain Monarchy this way:

Income Statement (Accounting)		
	Sales	1,000,000
-	Cost of Goods Sold	500,000
-	Expenses	300,000
=	Profits	200,000

Instead, the definition of pure economic profits would include the rental rate of capital as a cost. Even though Plantain Monarchy owns the retail space and does not need to pay rent, there is still an opportunity cost of not renting it out for $200,000. Therefore, under the definition of profits that we are using here, we have:

Income Statement (Economic)		
	Sales	$1,000,000$
-	Cost of Goods Sold	$500,000$
=	Value Added of Plantain Monarchy	$500,000$
-	Wages	$300,000$
-	Capital Rental	$200,000$
=	Profits	0

Even though from an accounting perspective it looks like Plantain Monarchy is profitable, the accounting profits it's earning are just the implicit rental from the capital that the firm owns.

In reality, of course, plenty of firms earn pure economic profits, i.e. profits beyond the implicit rental on the capital they own. Many firms make losses too. There are several possible reasons why firms might earn profits. One of the assumptions in our model is that there is perfect competition. If a firm has at least a little bit of monopoly power it can earn positive economic profits.[5] Another assumption in the model is that there is no risk: the output that will be produced is a perfectly predictable function of the inputs to the production process. In reality, firms face risk. It's possible that many firms actually earn zero *expected* profits but what we observe as nonzero profits is the result of either good or bad luck.

Interest Rates

Suppose that in addition to markets for hiring labor and renting capital there is a market for loans. A loan works as follows: the lender gives x goods to the borrower in period t and the borrower pays back $x(1+r_{t+1})$ goods to the lender in period $t+1$. r_{t+1} is the real interest rate on the loan. We say "real" interest to clarify that it's an interest rate in terms of goods, not in terms of dollars. If the loan was described in terms of dollars, we would need to convert dollars into goods by keeping track of how prices evolve.[6] For simplicity, we'll just describe loans in real terms directly: as exchanges of goods in one period for goods in the next period. We'll assume that no one ever defaults on their loans: they are always paid back. Also, we'll assume there is perfect competition: anyone can borrow or lend as much as they want at the interest rate r_{t+1}. Let's figure out what the interest rate is going to be in this economy.

Suppose someone wants to save x goods. They have two possibilities.

1. Physical investment. They convert their goods into x units of capital and rent them out in the following period. In the following period, this will give them:

	Rental Income	$r^K_{t+1}x$
$+$	Value of depreciated capital	$(1-\delta)x$
$=$	Total	$\left(1+r^K_{t+1}-\delta\right)x$

2. Lending. They lend their x goods in the loan market and get back $(1+r_{t+1})x$ goods.

[5] We'll think about models with monopoly power starting in Chapter 14.
[6] We'll look further at the distinction between real and nominal interest rates in Chapter 11.

They will be indifferent between the two options if and only if the following condition holds:[7]

$$r_{t+1} = r^K_{t+1} - \delta \tag{4.4.12}$$

We'll argue that condition (4.4.12) has to hold. Why? Suppose it were the case that $r_{t+1} < r^K_{t+1} - \delta$. Then one could make an infinite gain by borrowing at rate r_{t+1}, investing in physical capital, which earns $r^K_{t+1} - \delta$, using part of this to pay back the loans and keeping the difference. But if borrowing to invest is such a great deal, no one would be willing the lend and the loan market wouldn't clear. Conversely, if $r_{t+1} > r^K_{t+1} - \delta$ then everyone would want to lend and no one would want to borrow, so again the loan market wouldn't clear. Therefore (4.4.12) must hold. Condition (4.4.12) links the interest rate to the rental rate of capital and therefore to the marginal product of capital. Interest rates will be high in economies where the marginal product of capital is high (which, other things being equal, will be the case if capital per worker is low).

Response of Prices

Equations (4.4.10), (4.4.11) and (4.4.12) give us a way to think about what happens to the level of wages and interest rates in response to changes in the economy. Suppose there is an increase in the savings rate, as illustrated in Figure 4.2.2. In the new steady state, there is a higher level of k, which means wages will be higher and the rental rate of capital and therefore the interest rate will be lower. Similarly, faster population growth will lead to a shift to a steady state will lower wages and higher interest rates.

4.5 Technological Progress

If we want to understand the growth of GDP per capita in the US over the last 250 years the model we have studied so far doesn't have a lot of promise: it predicts that in the long run there will be no growth.[8] Now we are going to take the same economy and see what happens when there is technological progress.

We are going to represent technological progress as a change in the production function. Figure 4.2.3 shows what happens when there is a once-and-for-all change in the production function. Instead, we are now going to imagine that, due to technological progress, the production function moves up a little bit every period. We'll assume this upward shift takes a specific form, known as "labor augmenting" technological progress. This means that better technology is equivalent to having more workers. Mathematically, this means that we will assume that the production function takes the form:

$$Y = F(K, AL) \tag{4.5.1}$$

where the variable A represents the level of technology. (4.5.1) is a generalized version of the production function we had considered so far. Our original production function (4.1.1) is the special case where $A = 1$.

[7] We'll talk more about this equation in Chapter 9.

[8] One could still conjecture that 250 years is not long enough to test any predictions about "the long run". Strictly speaking, we never really reach the long run. Later we'll work on putting actual numbers on our model to see how fast it approaches the steady state. We'll see that after a few of decades we should expect almost no growth. See Exercise 5.1.

Define:

$$\tilde{L} \equiv AL$$

We'll refer to $\tilde{L}$ as "efficiency units of labor". If L workers are employed in a production process and the level of technology is A, then the output of the production process will be the same as if $\tilde{L} = AL$ workers were employed and the level of technology was 1.

Assumption 4.9. The level of technology grows at a constant, exogenous rate g: $A_{t+1} = (1+g)\,A_t$

Assumption 4.9 states that there is a stable, proportional rate of technological progress. We'll see that under this assumption the model will be consistent with a lot of facts about economic growth, in particular the steady rates of growth of advanced economies. On the other hand, it is rather disappointing to have to make this assumption. Ideally, one would like to have a deeper understanding of why there is technological progress and what determines how fast it takes place. We'll leave these important questions aside for now.

Mathematically, the model changes very little when we make Assumption 4.9. The key is to realize that instead of stating the production function in per capita terms, as in (4.2.1), we can instead write it in "per efficiency unit of labor" terms. Define:

$$\tilde{y}_t \equiv \frac{Y_t}{\tilde{L}_t}$$
$$\tilde{k}_t \equiv \frac{K_t}{\tilde{L}_t}$$

$\tilde{y}$ is output per efficiency unit of labor and $\tilde{k}$ is capital per efficiency unit of labor. These are not variables we are actually interested in but it's a convenient way to rescale the model.

As we did before, we can write down the production function in "per efficiency unit of labor" terms:

$$\begin{aligned} \tilde{y} &= \frac{F(K, AL)}{AL} \\ &= F\left(\frac{K}{AL}, 1\right) \\ &= f\left(\tilde{k}\right) \end{aligned} \tag{4.5.2}$$

The first step in (4.5.2) is just applying the definition of $\tilde{y}$ and using the production function (4.5.1) to replace Y. The second step uses Assumption 4.1 (constant returns to scale) and the last step just uses the definition of the function $f(k)$.

As we did before for k, we can derive an equation that describes how $\tilde{k}$ will evolve over time:

$$\begin{aligned} \Delta\tilde{k}_{t+1} &\equiv \tilde{k}_{t+1} - \tilde{k}_t && \text{(definition)} \\ &= \frac{K_{t+1}}{A_{t+1}L_{t+1}} - \tilde{k}_t && \left(\text{replacing } \tilde{k}_{t+1} \text{ with } \frac{K_{t+1}}{A_{t+1}L_{t+1}}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{(1-\delta)K_t + I_t}{A_{t+1}L_{t+1}} - \tilde{k}_t && \text{(using 4.1.4)}\\
&= \frac{(1-\delta)K_t + sY_t}{A_{t+1}L_{t+1}} - \tilde{k}_t && \text{(using 4.5.2)}\\
&= \left[\frac{(1-\delta)K_t + sY_t}{A_tL_t}\right]\frac{A_tL_t}{A_{t+1}L_{t+1}} - \tilde{k}_t && \text{(rearranging)}\\
&= \left[(1-\delta)\tilde{k}_t + s\tilde{y}_t\right]\frac{1}{(1+n)(1+g)} - \tilde{k}_t && \text{(using Assumptions 4.5 and 4.9: constant } n \text{ and } g)\\
&= \frac{s\tilde{y}_t - (\delta+n+g+ng)\tilde{k}_t}{1+n+g+ng} && \text{(rearranging)}\\
&\approx \frac{s\tilde{y}_t - (\delta+n+g)\tilde{k}_t}{1+n+g} && \text{(using that } ng \text{ is small)}\\
&= \frac{sf\left(\tilde{k}_t\right) - (\delta+n+g)\tilde{k}_t}{1+n+g} && \text{(using 4.5.2)} \quad (4.5.3)
\end{aligned}
$$

Formula (4.5.3) is a generalization of formula (4.2.2) and it has the same logic. The evolution of $\tilde{k}$ depends on the balance of investment pushing it up and depreciation, population growth and technological progress pushing it down. At first it may seem a little bit counterintuitive that technological progress pushes $\tilde{k}$ down: isn't technological progress supposed to help? The reason has to do with the way we define $\tilde{k}$: it's capital per efficiency unit of labor. If there is technological progress, this means that the number of efficiency units of labor is rising, just like it would from population growth. Therefore technological progress requires that we spread capital over more efficiency units of labor. For the purposes of calculating how $\tilde{k}$ evolves it doesn't matter whether $\tilde{L}$ rises because L rises or because A rises. It will matter very much once we convert back the per-efficiency-unit-of-labor measures into the per-person measures that we actually care about.

Given that formula (4.5.3) is so similar to formula (4.2.2) the dynamics that follow from it are also similar. The economy also has a steady state. The level of capital-per-efficiency-unit-of-labor in steady state, $\tilde{k}_{ss}$, is such that $sf\left(\tilde{k}_{ss}\right) = (\delta+n+g)\tilde{k}_{ss}$ so investment exactly balances out depreciation, population growth and technological progress. Graphically, it can also be represented by Figure 4.2.1, except that the straight line $(\delta+n)k$ is replaced by the slightly steeper line $(\delta+n+g)k$. As before, $\tilde{k}$ increases whenever it is below $\tilde{k}_{ss}$ and falls whenever it is above $\tilde{k}_{ss}$.

Does this mean that technological progress makes no difference for the predictions of the model? On the contrary! Let's see what happens when we convert per-efficiency-unit-of-labor variables into per-person variables. We know that in steady state, $\tilde{y}$ is constant. This implies that if the economy is in steady state:

$$
\begin{aligned}
\tilde{y}_{t+1} &= \tilde{y}_t && \text{(definition of steady state)}\\
\frac{Y_t}{A_tL_t} &= \frac{Y_{t+1}}{A_{t+1}L_{t+1}} && \text{(definition of } \tilde{y})\\
\frac{y_t}{A_t} &= \frac{y_{t+1}}{A_{t+1}} && \text{(definition of } y)\\
\frac{y_{t+1}}{y_t} &= \frac{A_{t+1}}{A_t} && \text{(rearranging)}
\end{aligned}
$$

$$\frac{y_{t+1}}{y_t} = 1 + g \qquad \text{(Assumption 4.9: constant } g\text{)} \tag{4.5.4}$$

This immediately implies the following:

Proposition 4.3.

1. In steady state, GDP per capita grows at the rate of technological progress.
2. In steady state, capital per worker grows at the rate of technological progress.
3. In steady state, the ratio $\frac{K}{Y}$ is constant.

Proof. Part 1 is just a restatement of (4.5.4). Part 2 follows from the same reasoning starting from $\tilde{k}_t = \tilde{k}_{t+1}$. Part 3 follows because:

$$\begin{aligned}\frac{K_{t+1}}{Y_{t+1}} &= \frac{\tilde{k}_{t+1}}{\tilde{y}_{t+1}} \\ &= \frac{\tilde{k}_t\,(1+g)}{\tilde{y}_t\,(1+g)} \\ &= \frac{K_t}{Y_t}\end{aligned}$$

□

Exercises

4.1 **An Earthquake**

Suppose an economy behaves according to the Solow growth model. It starts out at $t = 0$ at a steady state, with no technological progress and no population growth. Suppose an earthquake destroys half the capital stock. As a a consequence of this, what would happen in the short run (i.e. immediately) and in the long run (i.e. once the economy reaches steady state) to:

(a) GDP,

(b) GDP per capita,

(c) wages,

(d) interest rates.

4.2 **The Black Death**

In the middle of the 14th century, a plague killed about a third of the population of Europe. Assume that the economy of Europe was well described by the Solow model. How would the following variables change in response to the Black Death in the short run (i.e. immediately)?

(a) GDP,

(b) GDP per capita,

(c) wages,

(d) interest rates.

4.3 **Korean Unification**

Suppose both North and South Korea have the same technology level (!?), which doesn't grow, and also have constant population. Their respective populations and capital stocks are:

$$L^{South} = 10$$
$$K^{South} = 2430$$
$$L^{North} = 5$$
$$K^{North} = 160$$

The production function is:

$$Y = K^{\alpha} L^{1-\alpha}$$

with $\alpha = 0.4$. The depreciation rate is $\delta = 0.08$.

(a) Compute GDP per capita in each country.

(b) Compute wages and interest rates in each country.

(c) Suppose North and South Korea unify. What is the new $\frac{K}{L}$ in the unified country? What is GDP per capita in the unified country?

(d) Compute wages and interest rates in the unified country. Interpret

(e) Will people and/or machines physically move between the North and the South? In what direction?

4.4 **An AK Model**

Suppose the production function takes the form:

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

(a) Which of the assumptions that we made about F does this satisfy and which does it not satisfy?

(b) Suppose every other assumption we used in the Solow model holds, and there is no technological progress. Find an expression for $\frac{k_{t+1}}{k_t}$ (it may be useful to follow the steps that lead to expression (4.2.2)).

(c) Will this economy grow in the long run? Explain.

4.5 **A Malthusian Model**

Look up Thomas Malthus and read a little about his work. We are going to formalize a simple version of his ideas in a little model. The production function is:

$$Y_t = A_t L_t^{\alpha} \tag{4.5.5}$$

where A_t is technology, L is labor supply and $\alpha < 1$.

There is no capital, no saving and no investment, so consumption is given by

$$C_t = Y_t \tag{4.5.6}$$

The population evolves according to

$$L_{t+1} = L_t \left(1 + \gamma \left(\frac{C_t}{L_t} - \underline{c}\right)\right) \tag{4.5.7}$$

where γ and $\underline{c}$ are parameters.

(a) Malthus had in mind an agricultural economy where the total amount of land is fixed. What does that have to do with the way we wrote down the production function? What is the significance of assuming $\alpha < 1$?

(b) What does equation (4.5.7) mean? What is $\underline{c}$? What is γ? Why did Malthus think that something like equation (4.5.7) applied?

(c) Assume $A_t = A$ is constant over time. Using (4.5.5), (4.5.6) and (4.5.7), find an expression for the percentage rate of population *growth* $\frac{L_{t+1}}{L_t}$ as a function of the *level* of population L_t. Plot this function.

(d) Assume $A_t = A$ is constant over time. Will GDP per capita grow in the long run? Will the population grow in the long run?

(e) Suppose there is a one-time increase in A_t, from A to $A' > A$. What will happen to GDP per capita and the size of the population in the short run and in the long run? Show your reasoning graphically and/or algebraically.

(f) Suppose now that instead of being constant, technology improves at a constant rate, so that

$$A_{t+1} = A_t (1 + g)$$

Will GDP per capita grow in the long run? Will the population grow in the long run? Explain.

(g) When we have a constant rate of technological progress, how does the long-run *level* of consumption per capita depend on γ? Why?

CHAPTER 5

Confronting Theory and Evidence

In this chapter we'll take another look at the evidence on economic growth and test some of the predictions of the Solow model. We'll also use the model to suggest additional ways to look at the evidence.

5.1 The Kaldor Facts Again

Recall from Chapter 3 the so-called Kaldor Facts. Let's assume that advanced economies are well-described by the steady state of the Solow model with a constant rate of technological progress. Would that be consistent with the Kaldor Facts?

1. "The rate of growth of GDP per capita is constant". In the model, this is true by assumption. Proposition 4.3 says that GDP per capita grows at the rate of technological progress. Since we assume that technological progress takes place at a constant rate, then the rate of growth of GDP per capita is constant too. A success for the model, but not a huge one.

2. "The ratio of the total capital stock to GDP is constant". This is part 3 of Proposition 4.3, so another success for the model, this time with a result that is less obvious from the assumptions.[1]

3. "The shares of labor and capital income in GDP are constant". Let's see if this holds in our model. The share of labor in GDP is:[2]

$$\text{Labor share} = \frac{wL}{Y}$$

[1]Of course, all the results follow from the assumptions, but the usefulness of the model comes from telling us something that follows from the assumptions is not-so-obvious ways.

[2]When we write the production function with labor-augmenting technological progress as $F(K, AL)$, it's important to distinguish between (i) $\frac{\partial F(K,AL)}{\partial L}$ (the marginal product of labor) which corresponds to taking the partial derivative with respect to L while keeping A constant and (ii) $F_L(K, AL)$ (the derivative of F with respect to its second argument, evaluated at the point (K, AL)). For instance, with a Cobb-Douglas production function $Y = K^\alpha (AL)^{1-\alpha}$ we have

$$\frac{\partial F(K, AL)}{\partial L} = A(1-\alpha)\left(\frac{K}{AL}\right)^\alpha$$
$$F_L(K, AL) = (1-\alpha)\left(\frac{K}{AL}\right)^\alpha$$

From the labor demand decisions of firms we know that $w = \frac{\partial F(K,AL)}{\partial L}$, i.e. wages equal to the marginal product of labor. Instead, Proposition 4.1 is about F_L.

$$
\begin{aligned}
&= \frac{\frac{\partial F(K,AL)}{\partial L} L}{F(K,AL)} && \text{(using 4.4.9)} \\
&= \frac{A F_L(K,AL) L}{F(K,AL)} && \text{(taking the derivative)} \\
&= \frac{A F_L\left(\frac{K}{AL},1\right) L}{F\left(\frac{K}{AL},1\right) AL} && \text{(using Proposition 4.1 and Assumption 4.1)} \\
&= \frac{F_L\left(\frac{K}{AL},1\right)}{F\left(\frac{K}{AL},1\right)} && \text{(simplifying)}
\end{aligned}
$$

Since Proposition 4.3 implies that $\frac{K}{AL}$ is a constant, this implies that the labor share is a constant. The capital share is therefore also a constant because, by Proposition 4.1, they add up to 1.

We just proved that factor shares are constant if the economy is at a steady state with constant, labor-augmenting technological progress. If we assume that the production function takes the Cobb-Douglas form, then we can show that the labor share will be a constant even outside the steady state. Under the Cobb Douglas production function:

$$
\begin{aligned}
F(K,AL) &= K^{\alpha}(AL)^{1-\alpha} \\
w = \frac{\partial F_L(K,AL)}{\partial L} &= (1-\alpha) K^{\alpha} A^{1-\alpha} L^{-\alpha} \\
wL &= (1-\alpha) K^{\alpha} A^{1-\alpha} L^{1-\alpha} \\
\frac{wL}{Y} &= \frac{(1-\alpha) K^{\alpha} A^{1-\alpha} L^{1-\alpha}}{K^{\alpha}(AL)^{1-\alpha}} \\
&= 1-\alpha \qquad (5.1.1)
\end{aligned}
$$

so the labor share is $1-\alpha$ for any level of A, K and L. One reason we often assume that the production function takes the Cobb-Douglas form is that this immediately fits the finding of constant factor shares. Furthermore, this gives us an easy way to decide what is a reasonable value for α, since it's tied directly to factor shares.

4. "The average rate of return on capital is constant". As we saw in Chapter 3, this fact is an immediate implication of facts 2 and 3, and therefore it also holds in the Solow model. Just to check:

$$
\begin{aligned}
r^K &= F_K(K,AL) && \text{(using 4.4.8)} \\
&= F_K\left(\frac{K}{AL},1\right) && \text{(using Proposition 4.1)}
\end{aligned}
$$

and since Proposition 4.3 implies that $\frac{K}{AL}$ is a constant, this implies that r^K is a constant. r^K is the "gross" return on capital, meaning that it's what the owners of capital earn before depreciation. Once we subtract depreciation, we obtain the net return on capital $r^K-\delta$, which by equation (4.4.12) must equal the real interest rate. This will also be constant.

In terms of consistency with the Kaldor facts, the Solow model is doing quite well. Now let's see what other things we can do with it.

5.2 Putting Numbers on the Model

For some purposes it is useful to put concrete numbers on models. Let's see how we can use data to guide us towards reasonable numbers for the parameters of the Solow growth model. Let's imagine that we want our model to approximate the behavior of the US economy.

It is standard to assume that the production function takes the Cobb-Douglas form:[3]

$$F(K, AL) = K^{\alpha}(AL)^{1-\alpha}$$

The empirical basis for this assumption comes from equation (5.1.1): if the production function takes this form, factor shares will be constant even outside the steady state, which fits the Kaldor Facts. We just need to choose the right number for α, and again equation (5.1.1) gives us guidance. Recall from Figure 3.2.3 that the share of GDP that accrues to workers has averaged around 0.65 (although it's been lower recently). To be consistent with this fact, we should have $\alpha = 0.35$.

Proposition 4.3 says that GDP per capita grows at the rate of technological progress. GDP per capita in the US has grown at approximately 1.5% per year since 1800. To be consistent with this fact, we should use $g = 0.015$.

The US population has grown at a rate of approximately 1% per year since 1950. To be consistent with this, we should use $n = 0.01$.

Depreciation is a little bit harder, in part because different kinds of capital depreciate at different rates, in part because even for a specific kind of capital it's not so easy to determine how fast in depreciates. The Bureau of Economic Analysis uses the following values for some of the main types of capital: 0.02 for buildings, 0.15 for equipment, 0.3 for computers. We are not going to distinguish between different kinds of capital and we'll just use a single number for the overall capital stock. A plausible value is $\delta = 0.04$.

In choosing a number for s we need to use some judgment. The model we are using assumes that the economy is closed. As we saw in Chapter 4, this implies that savings equals investment. In reality, the US economy is not closed, and in the last couple of decades it has had higher investment than savings, as shown in Figure 5.2.1:[4]

If we want the model to replicate the investment rate that we have seen in the US in recent years, then we should set $s = 0.2$ approximately; if we wanted to replicated the savings rate we would set s a little bit lower.

Note that these figures for s, δ, n and g are consistent with the measured $\frac{K}{Y}$ ratio. Recall that by equation

[3] You will sometimes see this written as $AK^{\alpha}L^{1-\alpha}$. This doesn't make much difference. With the Cobb-Douglas function, the term $A^{1-\alpha}$ factors out anyway so it's just changing the units in which we measure "technology".

[4] The counterpart to this is a trade deficit. If you go back to equation (1.1.1) from Chapter 1:

$$\underbrace{Y - C - G}_{\text{Savings}} + \underbrace{M - X}_{\text{Trade Deficit}} = \underbrace{I}_{\text{Investment}}$$

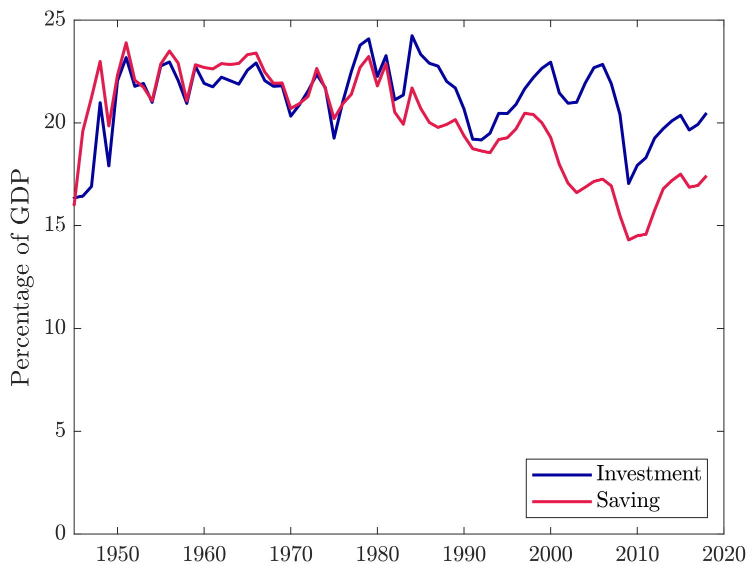


Fig. 5.2.1: *Investment (private plus public) and Saving as a percentage of GDP in the US. Source: NIPA.*

(4.5.3), in a steady state we must have:

$$\begin{aligned}\frac{K}{Y} &= \frac{\tilde{k}_{ss}}{\tilde{y}_{ss}} = \frac{\tilde{k}_{ss}}{f(\tilde{k}_{ss})} \\ &= \frac{s}{\delta + n + g} \\ &\approx 3.08\end{aligned} \tag{5.2.1}$$

which is close to what we saw in Figure 3.2.2.

The model predicts that the interest rate should be:[5]

$$\begin{aligned}r &= F_K(K, L) - \delta \\ &= \alpha \left(\frac{K}{AL}\right)^{\alpha - 1} - \delta \\ &= \alpha \frac{\delta + n + g}{s} - \delta \\ &\approx 0.074\end{aligned}$$

A 7.4% real interest rate is higher than has been observed in recent decades. Figure 5.2.2 shows the interest

[5]There are two ways to do the second step. One is to use equation (4.5.3) to solve for $k_{ss} = \left(\frac{s}{\delta+n+g}\right)^{\frac{1}{1-\alpha}}$ and plug it back into the formula. The other is to use the previously-shown result that the capital share of GDP is equal to α, set $\alpha = \frac{F_K(K,L)K}{Y}$, and use equation (5.2.1)

rate on 10-year inflation-indexed bonds (known as TIPS), which are the closest thing we have to market real interest rate, and these have been closer to 1% or 2%. These bonds have only been around for the past couple of decades but even before that we can reconstruct real interest rates from nominal interest rates and inflation measures, and on average they have been significantly below 7.4%. Clearly the model is in conflict with the data on this dimension. One likely source of this discrepancy is risk. In the model there is no risk, whereas in reality most real investment involves some risk. Riskier investments tend to earn higher returns than safe ones.[6] Maybe the more accurate comparison is between the model's prediction for interest rates and the return on risky investments like the stock market instead of a safe asset like inflation-indexed bonds. The average real return on the S&P500 index between 1929 and 2018 was 7.8%, much closer to the real interest rate in the model.

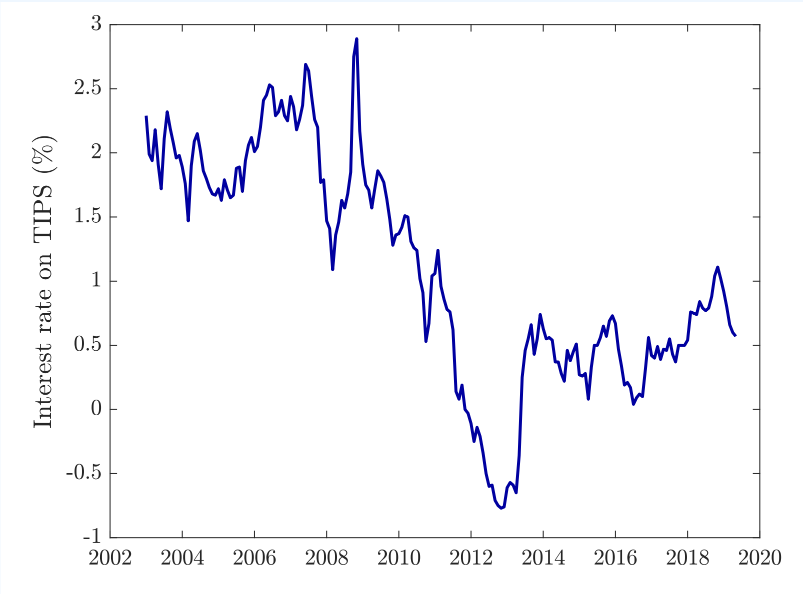


Fig. 5.2.2: *Interest rate on 10-year inflation-indexed bonds. Source: Board of Governors of the Federal Reserve.*

We saw in Chapter 4 that if an economy increases its savings rate it will reach a higher steady state level of capital per worker. With concrete numbers for the model parameters, we can ask quantitative questions like: how much higher? How long will it take to reach that level? Suppose the savings-and-investment rate rises from 0.2 to 0.25, starting from a steady state. We can use equation (4.5.3) to simulate the evolution of $\tilde{k}_t$. Figure 5.2.3 shows the results. In the new steady state, the capital stock (per efficiency unit of labor) is 41% higher and GDP (per efficiency unit of labor) is 13% higher. The interest rate falls from 7.4% to 5.1%. The transition to the new steady state takes time. Consumption only reaches its initial level after 17 years, and after 30 years, GDP has only risen 9%.

[6] In Chapter 8 we look at models of why that may be.

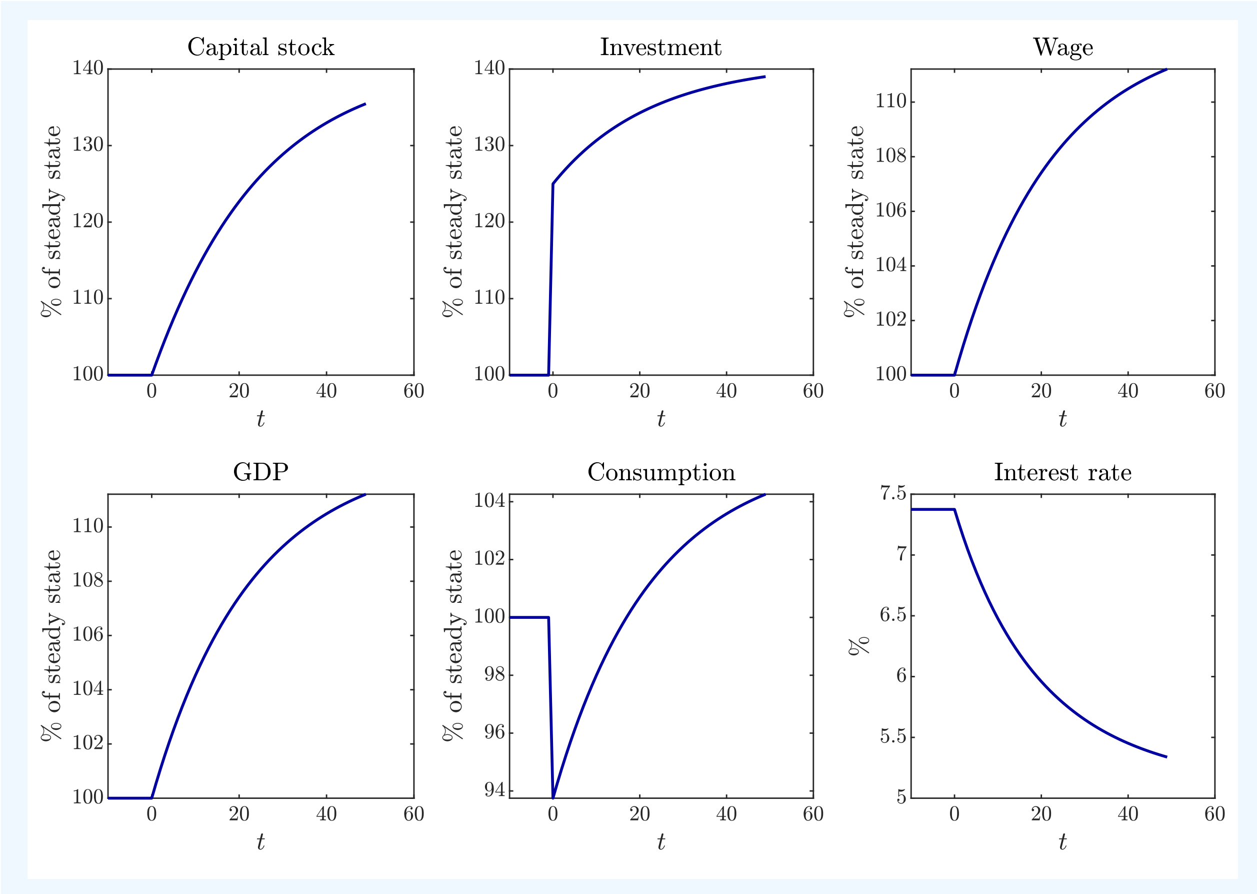


Fig. 5.2.3: *The economy's response to a higher saving rate. The capital stock, GDP, consumption, and investment are scaled by efficiency units of labor.*

5.3 The Capital Accumulation Hypothesis

One of the major questions in all of macroeconomics is why some countries are rich and others are poor. We can use the Solow model, together with data from different countries to test one possible answer to this question.

Conjecture 5.1. Technology levels are the same across countries and the differences in GDP per capita are the result of differences in $\frac{K}{L}$.

Conjecture 5.1 is a logically possible explanation of differences in GDP per capita across countries. Indeed, if two countries with the same production function had different levels of capital per worker, they would have different levels of GDP per capita. If the conjecture were true, it would have profound implications for economic development in poor countries: if the problem is low levels of capital, then capital accumulation is a good solution. However, we'll see that this conjecture is decisively rejected by the evidence. We'll address this in different ways.

Convergence

If two countries with the same production function have different levels of $\frac{K}{L}$, then we know that at least one of them, and maybe both, are not in steady state. Let's compute the growth rate of a country that is not at steady state. To avoid cluttering the algebra, let's assume that there is no technological progress and no population growth, but the argument holds regardless. Denote the growth rate of GDP per capita by g_y. Let's compute it:

$$\begin{aligned}
g_y &\equiv \frac{y_{t+1}}{y_t} - 1 && \text{(by definition)} \\
&= \frac{f\left(k_{t+1}\right) - f\left(k_t\right)}{f\left(k_t\right)} && \text{(replacing the production function 4.2.1)} \\
&\approx \frac{f'\left(k_t\right)\left[k_{t+1} - k_t\right]}{f\left(k_t\right)} && \text{(this is a first-order Taylor approximation)} \\
&= \frac{f'\left(k_t\right)\left[sf\left(k_t\right) - \delta k_t\right]}{f\left(k_t\right)} && \text{(using 4.2.2)} \\
&= sf'\left(k_t\right) - \delta \frac{f'\left(k_t\right) k_t}{f\left(k_t\right)} && \text{(rearranging)} \\
&= sf'\left(k_t\right) - \delta\alpha && \text{(letting } \alpha \text{ represent the capital share)} \qquad (5.3.1)
\end{aligned}$$

Let's see what equation (5.3.1) means. Suppose we compare two countries with different levels of GDP. The richer country, according to Conjecture 5.1, has a higher capital stock. By Assumption 4.3 (diminishing marginal product), a higher capital stock implies a lower marginal product of capital. Therefore equation (5.3.1) implies that the richer country will grow more slowly than the poorer country.

At some level, we sort of knew that already. By assumption, the countries are converging to the same steady state so it makes sense that the country that is already very close would grow more slowly than the one that is far behind. Equation (5.3.1) makes this precise: at any point in the path that leads to the steady state, the poorer country should be growing faster than the richer country.

This is something we can test directly with cross-country data. Is it the case that initially-poorer countries grow faster than initially-richer countries? Figure 3.3.1 from Chapter 3 gives us an answer. If Conjecture 5.1 were correct, we should see a strong negative correlation between initial GDP per capita and growth rates, and this is not what we see.

Figure 3.3.1 is not necessarily the end of the discussion. One thing that the theory does not clearly specify is whether a "country" is the right unit of observation. Maybe the lack of correlation between initial GDP and subsequent growth is the result of a lot of weird, small, possibly irrelevant countries? An alternative way to look at the cross-country data is to weigh each observation in proportion to the population of that country.[7]

[7] It's not entirely clear whether weighting by population is the right thing to do. If what we want to test is the universal validity of Conjecture 5.1, one could make the case that a small country provides an equally valuable experiment as a large country. On the other hand, not all important economic forces operate at the level of an entire country; perhaps each of India's 29 provinces should be considered a separate experiment.

Figure 5.3.1 shows what happens if we do that.

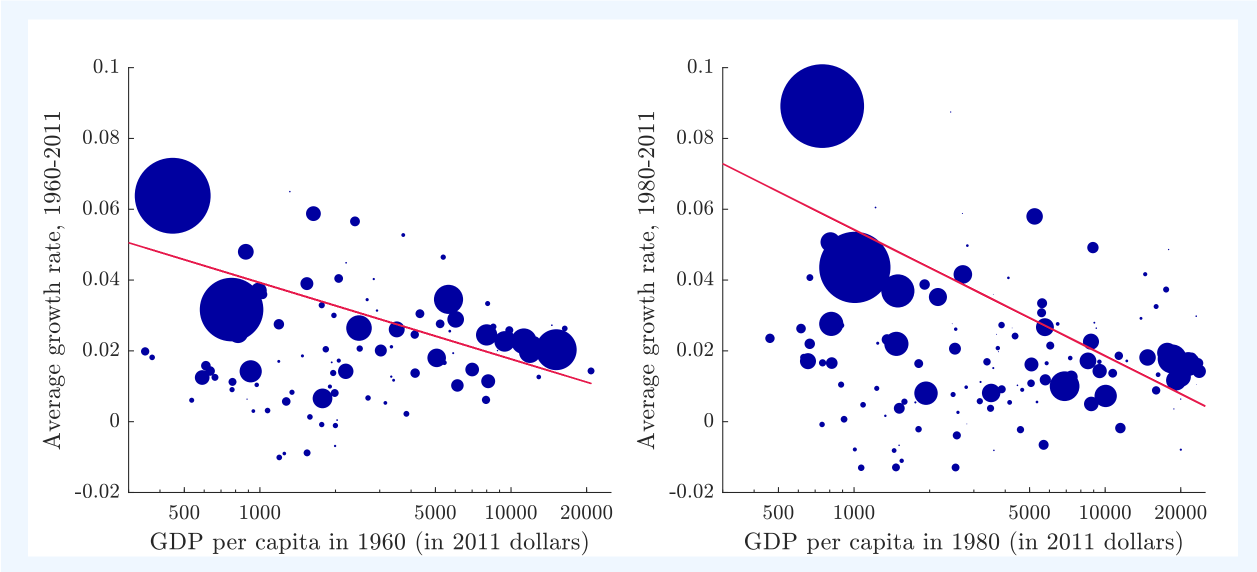


Fig. 5.3.1: *Growth across countries since 1960 and 1980, population-weighted. Source: Feenstra et al. (2015).*

Weighted this way, the data do show some convergence. If we focus on the period since 1980, the pattern is stronger. There is a simple fact behind this: China and India started this sample being very poor and have grown very fast, especially in the last few decades. Given their large population, these two observations largely determine the overall pattern.

Another question one can ask is whether Conjecture 5.1 might be true for specific groups of economies, even if it's not true for the world as a whole. For instance, different US states have different levels of GDP per capita. Maybe for US states it's true that they have the same production function but different levels of capital per worker? Capital abundance could be the main reason why Connecticut is richer than Louisiana even if it's not the main reason the US is richer than Paraguay. Figure 5.3.2 shows what happens if we repeat the exercise but focus, respectively, on US states and Western European countries. Here we do see strong evidence of convergence: US states and European countries that started out poor indeed grew faster than their richer counterparts. Note, however, that the US data in Figure 5.3.2 covers the period 1929-1988. In more recent decades, poorer US states have not grown faster than rich ones.

Note that from a logical point of view, finding evidence consistent with a conjecture is not the same as proving the conjecture. The evidence on convergence is consistent with a limited, intra-US or intra-Europe version of Conjecture 5.1 but it could also be that the conjecture is wrong and these economies are converging for other reasons (for instance, faster rates of technological progress in poor US states or European countries).

Direct Measurement

Another way to test Conjecture 5.1 is to:

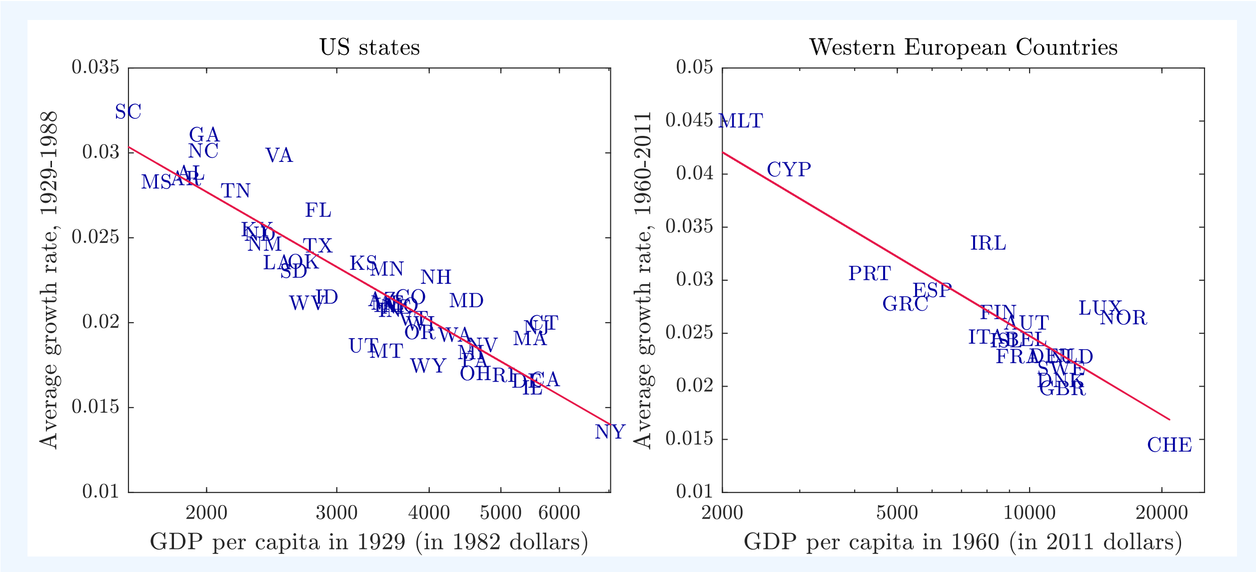


Fig. 5.3.2: *Growth across US states and Western European countries. Source: Barro and Sala-i-Martin (1991) and Feenstra et al. (2015).*

1. Directly measure the capital stock in each country
2. Assume a form for the production function (which Conjecture 5.1 assumes is the same across countries). For instance, assume $f(k) = Ak^\alpha$, which we know is a decent approximation to the production function for the United States if we set $\alpha = 0.35$.
3. Predict the GDP levels that you would get by just plugging in the measured level of capital per worker into the production function. For this, first solve for A using data for the United States:

$$A = \frac{y_{USA}}{k_{USA}^\alpha}$$

   Then compute predicted GDP per capita for country j (denoted $\hat{y}_j$), under Conjecture 5.1 by computing:

$$\hat{y}_j = Ak_j^\alpha$$

4. Compare predicted levels of GDP to measured levels of GDP

If the Conjecture 5.1 is correct, predicted and actual levels of GDP should look quite similar.

The main challenge in doing this is that measuring the capital stock in a way that is comparable across countries is actually quite hard (Exercise 4 asks you to think about some of the difficulties). So if the data don't line up with the prediction it's possible that mismeasurement of the capital stock is part of the answer.

Figure 5.3.3 shows the comparison between predicted and actual levels of GDP, and the patterns are very strong: in poor countries, actual GDP is consistently much lower than you would predict if you just knew their capital stock and assumed the production function was the same. Furthermore, the difference is greater the poorer the country. If the only difference was the capital stock, the poorest countries should have a GDP per capita of around \$10,000 instead of their actual GDP per capita, which is closer to \$1,000. This is more evidence against the common-production-function, same-technology-level assumption of Conjecture 5.1.

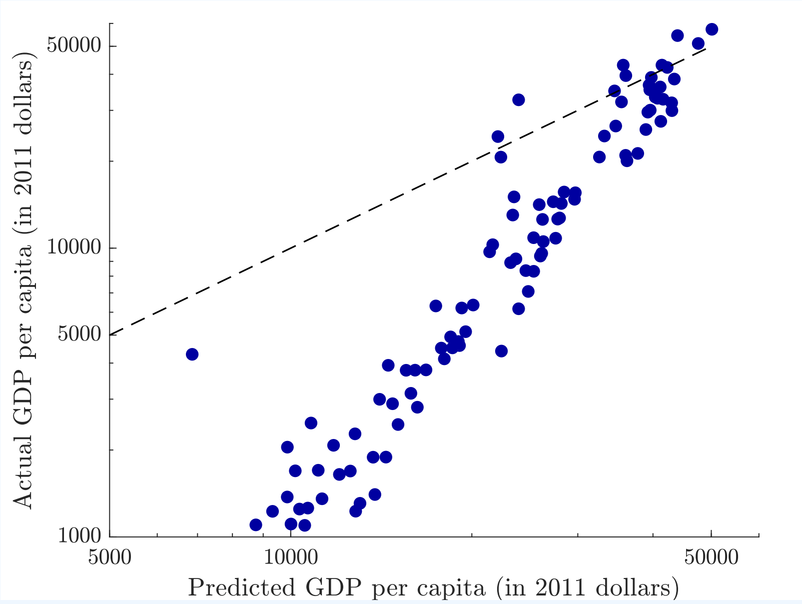


Fig. 5.3.3: Predicted GDP per capita on the basis of capital stock compared to actual GDP per capita. Source: Feenstra et al. (2015)

Evidence from Interest Rates and Capital Flows

Suppose we don't trust the data on capital levels across countries. Another things we can try to measure is interest rates in different countries. Why would these be informative? Recall from equations (4.4.10) and (4.4.12) in Chapter 4.4 that interest rates are equal to:

$$r = r^K - \delta = F_K\left(\frac{K}{L}, 1\right) - \delta$$

Other things being equal, an economy that has low levels of capital per worker will have a high marginal product of capital, a high rental rate of capital and high interest rates.

How much higher? This depends on the exact shape of the production function. Suppose that we assume a Cobb-Douglas production function and we compare two countries, A and B. We know that GDP per capita in country A is x times higher than in country B but we don't know the level of their respective capital stocks because we don't trust our measurements. If Conjecture 5.1 were true, how would the interest rates in the two countries compare?

Using the Cobb-Douglas production function we obtain:

$$r^K = \alpha k^{\alpha-1} \tag{5.3.2}$$

Equation (5.3.2) tells us that if we want to predict how rental rates of capital will differ across countries we need to find out how $k^{\alpha-1}$ will vary across countries. Let's compute this:

$$\begin{aligned} x &= \frac{y_A}{y_B} && \text{(by assumption)} \\ &= \left(\frac{k_A}{k_B}\right)^{\alpha} && \text{(using the Cobb-Doulgas production function)} \\ x^{\frac{\alpha-1}{\alpha}} &= \left(\frac{k_A}{k_B}\right)^{\alpha-1} && \text{(rearranging)} \end{aligned} \tag{5.3.3}$$

(5.3.3) tells us that the comparison of $k^{\alpha-1}$ across countries (and therefore the comparison of interest rates) is linked to the comparison of GDP per capita levels in a very specific way. Plugging in (5.3.3) into (5.3.2) we get:

$$\frac{r_A^K}{r_B^K} = x^{\frac{\alpha-1}{\alpha}} \tag{5.3.4}$$

Let's try to put some numbers on this to see what it means. Let's say we want to compare Mexico and the US. Mexican GDP per capita is approximately 0.3 times that of the US. Setting $\alpha = 0.35$, formula (5.3.4) then implies that

$$\frac{r_{US}^K}{r_{MEXICO}^K} = 0.3^{\frac{0.35-1}{0.35}} \approx 9.4$$

so the rental rate of capital in Mexico should be more than nine times higher than in the US. In order to account for why Mexico is so much poorer than the US under Conjecture 5.1 we must infer that the capital stock is *much* lower. If capital is so scarce, then it's marginal product must be very high and so must its rental rate.

We can link this back to interest rates by recalling that

$$r^K = r + \delta$$

Suppose the relevant interest rate in the US is 7.4% and the rate of depreciation is 4%, then we have that $r^K = 0.114$ in the US, which would mean that $r^K = 1.08$ in Mexico and therefore we should expect $r = r^K - \delta = 104\%$ in Mexico. Table 5.1 shows the result of repeating the calculation for several countries.

Table 5.1: *Interest rates implied by Conjecture 5.1 for different countries. Source for GDP data: Feenstra et al. (2015).*

Country	$x = \frac{y}{y_{US}}$	$\frac{k}{k_{US}} = x^{\frac{1}{\alpha}}$	$r = r_{US}^{K} x^{\frac{\alpha-1}{\alpha}} - \delta$
Switzerland	1.02	1.05	7.0%
USA	1	1	7.4%
Portugal	0.47	0.12	41%
Mexico	0.30	0.031	104%
China	0.25	0.019	146%
India	0.089	0.001	1,002%
Ethiopia	0.026	0.00003	9,819%

Table 5.1 shows that if Conjecture 5.1 were true, then we should observe *extremely* high interest rates in poor countries, because the marginal product of capital would be extremely high.

Lucas (1990) argued that if this were true, then the incentives for investors from rich countries to invest in poor countries would be huge. When we analyzed the Solow model we assumed that each country was a closed economy. But no country is a fully closed economy: cross-border investment is possible, even if not fully unrestricted. If the figures in Table 5.1 were correct, why invest in the US and earn an 7.4% return when you can invest in Ethiopia and earn a $9,819\%$ return? If one believed Conjecture 5.1, it would be surprising to observe *any* investment in rich countries at all.

5.4 Growth Accounting

Suppose that we observe that GDP in some country has increased and we want to understand why. Is is because they have invested and the capital stock has increased? Has there been technological progress? Or is it just that there are more people working? We can use a technique called "growth accounting" to measure the contribution of each of these factors. The basic idea is to:

1. Measure how much capital and labor have changed.

2. Figure out how much change in output we should expect from that. This is the key step; here we rely on combining theory and measurement.

3. Attribute all the changes in output that cannot be accounted for by changes in capital in labor to changes in productivity.

Start from the production function:

$$Y_t = F\left(K_t, L_t, A_t\right) \tag{5.4.1}$$

A couple of things to note about equation (5.4.1). First, we are including technology as a separate argument instead of assuming it just enters as labor-augmenting. The labor-augmenting case $F\left(K, AL\right)$ is a special case of (5.4.1) but we want to allow technological progress to possibly take other forms. Also, we are including time subscripts on all the variables because we want to think about how each of them changes over time.

Now define GDP growth as

$$g_Y \equiv \frac{Y_{t+1}}{Y_t} - 1$$

and compute:

$$\begin{aligned}
g_Y &= \frac{F(K_{t+1}, L_{t+1}, A_{t+1})}{F(K_t, L_t, A_t)} - 1 \\
&\approx \frac{F(K_t, L_t, A_t) + F_K(K_t, L_t, A_t)(K_{t+1} - K_t) + F_L(K_t, L_t, A_t)(L_{t+1} - L_t) + F_A(K_t, L_t, A_t)(A_{t+1} - A_t)}{F(K_t, L_t, A_t)} - 1 \\
&= \underbrace{\frac{F_K(K_t, L_t, A_t) K_t}{F(K_t, L_t, A_t)}}_{\text{Capital Share}} \underbrace{\frac{(K_{t+1} - K_t)}{K_t}}_{\text{Growth Rate of Capital}} + \underbrace{\frac{F_L(K_t, L_t, A_t) L_t}{F(K_t, L_t, A_t)}}_{\text{Labor Share}} \underbrace{\frac{(L_{t+1} - L_t)}{L_t}}_{\text{Growth Rate of Labor}} + \underbrace{\frac{F_A(K_t, L_t, A_t) A_t}{F(K_t, L_t, A_t)} \frac{(A_{t+1} - A_t)}{A_t}}_{\text{“Solow Residual”}}
\end{aligned} \tag{5.4.2}$$

The first step is a first-order Taylor approximation and the second is just rearranging.

What is equation (5.4.2) telling us? It's giving us a way to decompose the growth we observe into three terms. The first says how much the economy would have grown just as a result of capital accumulation, all else equal. The second tells us how much the economy would have grown just due to changes in the size of the labor force, all else equal. The last term measures all the growth that is *not* explained by changes in measured factors of production; we collect all of this and call it a “Solow residual”.

What is the logic behind equation (5.4.2)? Other things being equal, more capital will increase GDP by an amount equal to the marginal product of capital and more labor will increase GDP by an amount equal to the marginal product of labor. So if we want to determine how much extra production is the result of changes in capital and labor we need to know their respective marginal products. The key insight that we get from equation (5.4.2) is that looking at each factor's share of GDP is precisely what we need to do to infer what their respective marginal products are. Therefore the equation is telling us what things we need to measure. GDP growth and factor shares can be read off directly from GDP accounts. The growth rate of the capital stock is a little bit trickier because one needs to measure the original capital stock, estimate depreciation and measure investment, but each of these things can be done, even if perhaps not perfectly. Measuring the growth rate of labor is not just a question of measuring population growth (which is quite easy) because changes in labor force participation, unemployment or even hours worked per employed worker can be large. With sufficiently good labor market data, this is not hard to measure either.

This leaves us with the Solow residual as the only part of equation (5.4.2) that we can't measure directly. What do we do about it? The key is in the name: it's known as a Solow residual because you can measure it by measuring everything else in equation (5.4.2) and then solving for what the residual needs to be for the equation to hold, i.e.:

$$\text{Solow Residual} = \text{GDP growth} - \text{Capital share} \times \text{Capital Growth} - \text{Labor Share} \times \text{Labor Growth}$$

What do we do with the Solow residual? We interpret it as capturing the contribution of everything other than accumulation of capital and increases in labor to GDP growth. It includes everything from literal technological progress to changes in policies that lead to better (or worse) allocation of resources. It is sometimes referred to as “Total Factor Productivity” or TFP.

Example 5.1.

In the 1960s there was a lot of worry in the West about economic growth in the USSR. There was little reliable information about the performance of the Soviet economy but many observers had the impression that it was growing very fast. In the context of the Cold War, many in the West were panicking about this, for two main reasons. First, if the Soviet economy kept growing so fast, they would be able to afford more military expenditure, changing the balance of power. Second, if they were growing so fast, maybe they were on to something? Perhaps centralized direction of the economy was a superior technique for economic management? One of the more famous growth-accounting exercises was undertaken to try to answer this question. A big part of the effort was trying to reconstruct figures for GDP, the labor force and capital accumulation, which the USSR didn't publish. After that, applying formula (5.4.2) was relatively straightforward. Powell (1968) found that growth in the USSR between 1928 and 1966 could be decomposed as follows:[a]

GDP Growth	K Share	Capital Growth	L Share	Labor Growth	Residual
5.4%	0.4	6.5%	0.6	2.8%	1.1%

This decomposition teaches several lessons. First, the Soviet economy was indeed growing very fast, more than 2 percentage points faster than the US. Official Soviet figures claimed even faster growth, around 9% per year, but even 5.4% is very fast growth. Compounded over more than three decades, this was a huge transformation. Second, this growth was not mysterious. It was in large part a result of very high investment rates which led to fast capital accumulation and of increases in the labor force. Investment rose from 8% of GDP at the beginning of the period to 31% of GDP towards the end. The increase in the labor force was faster than population growth, especially due to the increased labor force participation of women. Third, the rate of productivity growth was unremarkable, close to US levels.

Fourth, and most controversially, the fast rate of growth should not be expected to continue. Cold War strategists should not panic. Does this conclusion follow directly from the accounting exercise? No. But suppose that we conclude that the combination of policies pursued by the USSR boils down to: high investment, an increase in labor force participation and mediocre TFP growth, and this is expected to continue. High investment leads to growth for a while but eventually runs into the diminishing marginal product of capital, as we saw in Chapter 4. Increases in labor force participation cannot continue: eventually almost everyone is working. So unless productivity growth somehow accelerates, a simple application of the Solow model would suggest that the USSR should be expected to slow down. To be fair, we don't have a great understanding of what determines productivity growth, so the prediction of a slowdown was conditional on the assumption that productivity growth would not suddenly accelerate. The performance of the Soviet economy in the 1970s and 1980s was a great vindication of this prediction.

[a]Factor Shares are harder to measure in an economy that does not rely on markets, as was the case in the USSR. The capital share of 0.4 is one point in the range of estimates that Powell (1968) considered. Also, for the period 1928-1937, the decision of whether to measure GDP at base-year prices or final-year prices makes a big difference. The reason is that the manufacturing sector was expanding more than the agricultural sector at the same time that the prices of manufactured goods were falling. The 5.4% figure below corresponds to using final-year prices; using base-year prices gives 6.7%.

5.5 Where do TFP differences come from?

We have quite good evidence that production functions are not the same in all countries, that some countries can simply extract more output from the same amount of capital and labor. Furthermore, the Solow growth model predicts that growth in productivity is the only way sustain economic growth in the long run. What do we know about what explains TFP differences across countries? There are several conjectures, not necessarily incompatible with each other.

Human Capital

So far we've been treating all labor as though it was the same, but different people's labor may contribute differently to total output. In fact, we have good evidence that this is the case, because different people get paid very different wages. Economists use the term "human capital" to describe the productivity differences that are embedded in people's knowledge, skills or talent. Calling this human capital hints at the idea that this is something that can be accumulated, for instance through education. One possible explanation for why different countries have different productivity is that their labor is not actually the same because in some countries labor has a lot more "human capital" in it.

Hall and Jones (1999) compute differences in productivity across countries by directly measuring the capital stock and computing a human-capital-adjusted measure of the labor supply. Their basic idea is to use a worker's years of education as a measure of human capital. But how do you choose the units? Is a worker with 12 years of education equivalent to 2 workers with no education? Or more? We can use evidence from the labor market to answer this. Suppose we measure wages for workers with different levels of education. Under the assumption that markets are competitive, each worker gets paid their marginal product so wages are the correct measure of human capital. If we measure an empirical relation that says:

$$\text{Wage} = w\,(\text{Years of Education})$$

then we can use the function w to convert years of education into units of human capital. Then we use education data across countries to convert the labor force into human-capital-adjusted labor force.

Do differences in human capital explain the cross-country differences in productivity that we find when we use unadjusted labor? In part, but large differences remain unexplained.

Figure 5.5.1 repeats the exercise from Figure 5.3.3, but using information on both human capital and physical capital to derive the predicted level of GDP per capita. Comparing the two figures we can see that differences in human capital help to close some of the gap between predicted GDP per capita and actual GDP per capita. However, most countries still have GDP per capita that is much lower than predicted. In other words, low human and physical capital are not enough to account for the relative poverty of poor countries.

Geography

Figure 5.5.2 shows a map of the world with colors representing levels of GDP per capita.

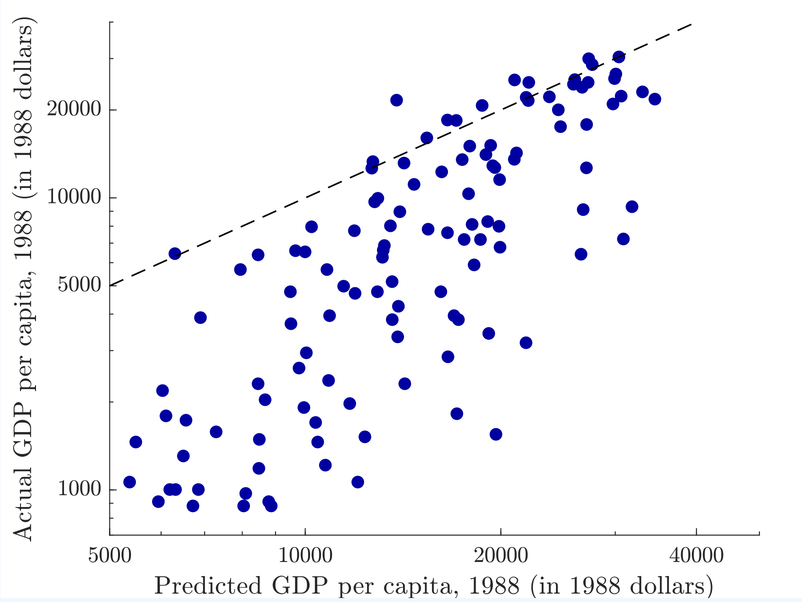


Fig. 5.5.1: *Predicted GDP per capita on the basis of capital stock human capital compared to actual GDP per capita. Source: Hall and Jones (1999).*

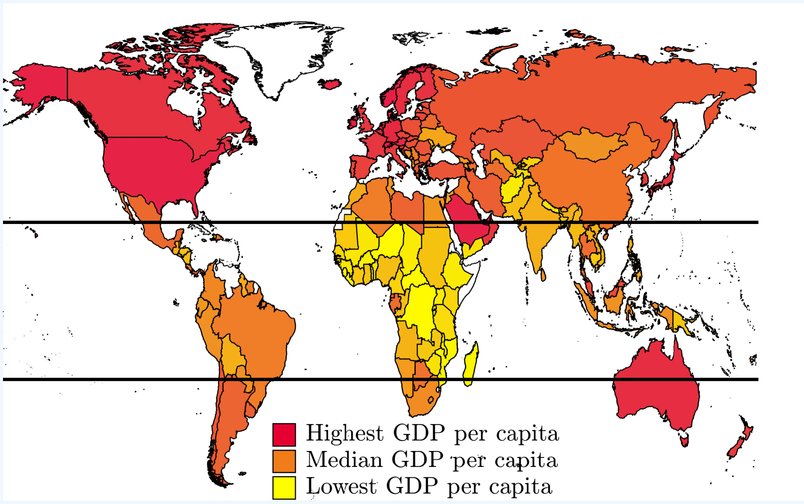


Fig. 5.5.2: *GDP per capita at PPP in 2017. Source: World Bank.*

There is a very strong pattern. Countries in colder climates tend to be much richer than countries in hotter climates. Most of the poorest countries in the world are between the tropics. There are a number of reasons why that might be. Sachs (2001) argued that one reason is that climate itself matters.

Crop yields are typically lower in very hot weather, higher in temperate weather and lower again in very cold weather. This might be a fixed aspect of the "technology" that affects productivity directly. On the other hand, agriculture is not such a large part of the economy, especially in rich countries, so this is unlikely to be the full explanation. In other words, the reason Sweden is rich is not that crop yields are higher than in

Kenya.

Another channel that links geography to productivity is disease. Many diseases, like malaria, are endemic to hot climates. In addition to killing millions of people, these diseases may affect productivity by hindering the physical development of children, by keeping children away from school and hurting their learning, by keeping adults away from jobs, etc.

Institutions

The term "institutions" is used to refer to a whole number of political and social conditions that vary across countries: democracy, the adherence to the rule of law, political transparency, respect for property rights, etc. One conjecture is that these institutions could have a large effect on productivity. You'll see an example of why that might be in Exercise 5.5.

At a basic level, we do observe a strong correlation between institutions and TFP. The same countries that are richer are more democratic. Whether political and social institutions are the reason for their prosperity is less clear. Perhaps it's just that as countries become richer (for other reasons), the political balance shifts toward more democracy and openness.

One famous study (Acemoglu et al., 2001) made the case that the relationship is indeed causal, at least in part. It focused on countries that had been at some point colonized by European powers. The study measured the mortality rate of European settlers in these colonies and found a strong correlation between this and present-day institutions. The measure of "institutions" is a "risk of expropriation" index developed by the consultancy Political Risk Services that attempts to measure how likely it is that an investor's property will be taken from them (a high score means a low chance of expropriation). The left panel of Figure 5.5.3 shows how the mortality of European settlers in colonized countries correlates with this risk-of-expropriation measure. The correlation is strong and negative: countries where settlers had high mortality now have higher risk of expropriation.

The authors of the study argued that this is because of the different types of colonization undertaken in different places. In low-mortality places, Europeans had higher hopes of settling, so they brought with them the relatively more open institutions that were developing in Europe. In high-mortality places, they had less hope of settling so they set up what the authors call "extractive" institutions, which used political and military authority to extract natural and other resources. These extractive institutions have persisted into the post-colonial period and, the argument goes, are what explains today's low productivity. The right panel of Figure 5.5.3 shows how how the mortality of European settlers correlates with GDP per capita. Countries where settlers had high mortality now have lower GDP per capita.

There is some debate as to whether Acemoglu et al. (2001) have really nailed down the causal argument. Could there be other explanations of why the places where Europeans had lower mortality in the colonial era are wealthier today? At least two alternative explanations have been proposed. One is that the same factors that contributed to high mortality (in particular, tropical diseases) are also directly responsible for today's low productivity. Another is that in places of low mortality, European settlers brought other things besides more open institutions: technology, human capital, trade links, etc.

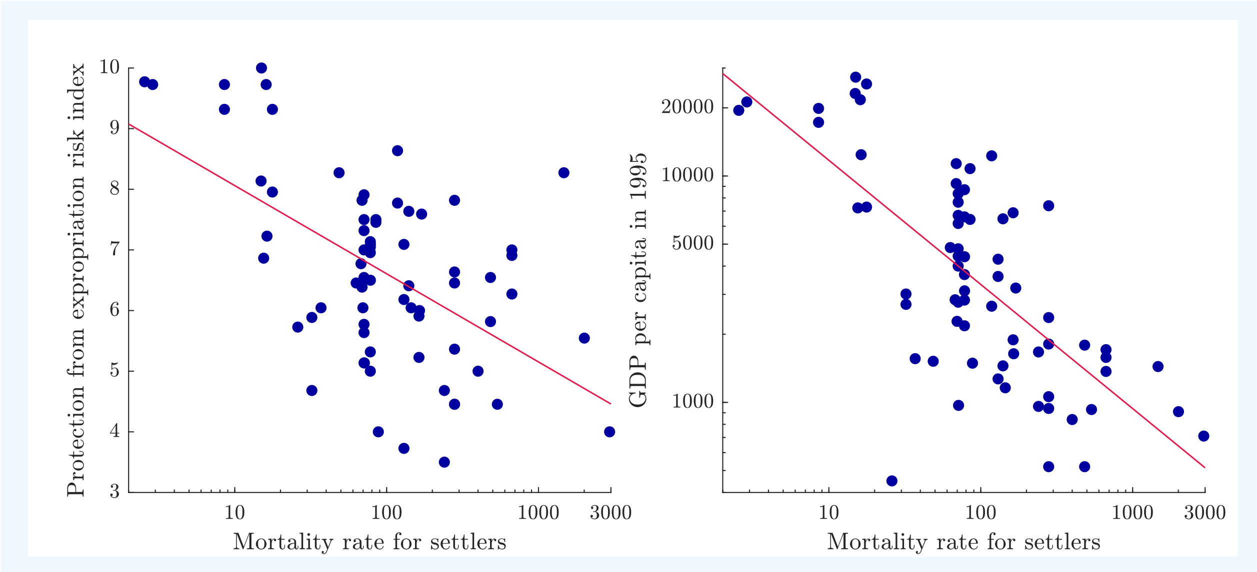


Fig. 5.5.3: *Mortality of colonial settlers, present-day institutions and present-day GDP per capita. Source: Acemoglu et al. (2001).*

Resource Allocation and Misallocation

Maybe poor countries get less output out of capital and labor because they don't put them to good use. Suppose that there are many different firms in each country. To maximize output, each unit of capital and each worker should be working for the firms where their marginal product is highest, which implies that the marginal products are equated across firms. There might be several factors preventing this from happening, and these could vary by country.

One possible factor is entry regulation. The World Bank measures all the requirements for starting a business in different countries: permits, registration delays, etc. These vary greatly by country. One possibility is that many potential businesses don't even get started because of these barriers to entry, so capital and labor get allocated to less productive uses. Figure 5.5.4 shows the correlation between the number of days it takes to start a new business and GDP per capita, which is negative. One possible interpretation is that the barriers to entry for new firms, of which delay is just one example, lead to misallocation of resources and therefore have a *causal* effect on productivity. Alternatively, it could be that countries that are richer *for other reasons* have speedier procedures for registering new firms.

Similar effects could result from lack of credit, restrictions on foreign investment, monopoly rights or taxation or regulation that is uneven across firms. All of these could lead capital and labor away from their most productive uses.

One interesting piece of evidence on this comes from Bloom et al. (2013). They surveyed a sample of textile firms near Mumbai, India, and found that they differed greatly in their productivity and management practices. Why were the most efficient firms not expanding and taking business away from the less efficient firms? Bloom et al. (2013) found that one of the main determinants of a firm's size is the number of male

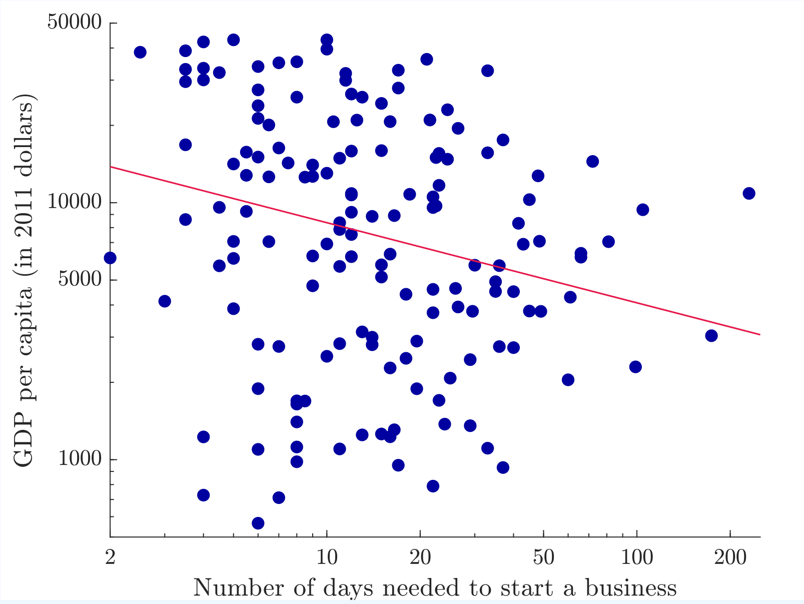


Fig. 5.5.4: Time to start a new business and GDP per capita. Each observation represents a country. Source: World Bank (2011).

family members the owner had. Their interpretation of this finding is that because enforcing contracts is not easy in India, it is very hard for the owner of a firm to delegate management to outsiders who are not family members.[8] This limits the extent to which the more productive firms can expand, so the less productive firms end up employing an inefficiently large amount of capital and labor.

Hsieh and Klenow (2009) estimated the dispersion in the marginal product of capital and labor in manufacturing firms in China, India and the US. Based on these estimates, they calculated that if China and India could reduce the degree of misallocation of capital and labor across firms to the levels observed in the US, their TFPs could increase by $30\% - 50\%$ and $40\% - 60\%$ respectively.

Exercises

5.1 **Quantifying the Solow Model**

Suppose the economy is described by the Solow growth model with the parameter values we used in Section 5.2.

(a) Suppose the economy starts out with a level of GDP per efficiency unit of labor that is only 10% of its steady state level, i.e. $\tilde{y}_0 = 0.1\tilde{y}_{ss}$. What is the initial capital stock $\tilde{k}_0$? Use a spreadsheet to compute $\tilde{k}_t$ and $\tilde{y}_t$ for $t = 1, 2, \ldots, 100$. How many years does it take for $\frac{\tilde{y}_t}{\tilde{y}_{ss}} > 0.95$?

(b) Suppose the economy begins in steady state but there is a fall in the rate of population growth from $n = 0.01$ to $n = 0$. How much higher will GDP per efficiency unit of labor be in the new steady

[8] Also: sexism. Only male family members were associated with firm size.

state? Use a spreadsheet to compute $\tilde{k}_t$ and $\tilde{y}_t$ for $t = 1, 2, \ldots, 100$. How many years does it take for $\frac{\tilde{y}_t}{\tilde{y}_{ss}} > 0.95$?

5.2 **The Speed of Convergence**

Suppose the economy is described by the Solow growth model with a Cobb-Douglas production function, no population growth and no technological progress. The economy is not in steady state. GDP capita in period t is $y_t = \theta y_{ss}$, with $\theta \in (0, 1)$.

(a) How far away from steady state is the capital stock? Find an expression for the level of $\frac{k_t}{k_{ss}}$ that would be consistent with $y_t = \theta y_{ss}$, in terms of θ and parameters.

(b) Find an expression for k_{ss}, the steady state level of capital, in terms of parameters.

(c) Use your results from parts (a) and (b) and equation (5.3.1) to find an expression for g_y, the rate of growth of GDP, as a function of θ and parameters.

(d) Let $\alpha = 0.35$ and $\delta = 0.04$. Compute g_y for an economy whose GDP is $\theta = 0.8$ times its steady-state level? Compute $\frac{g_y}{1-\theta}$. What fraction of the gap between y_t and y_{ss} does the economy close in one year? Repeat this calculation for $\theta = 0.9$ and $\theta = 0.99$.

5.3 **Causes of Growth and Growth Accounting**

Suppose an economy is well described by the steady state of the Solow Growth Model with constant technological progress and no population growth. Imagine we take data from this economy and do a growth-accounting exercise. How much growth will we attribute to capital accumulation and how much to technology? How does this relate to the result that says that without technological progress there would be no growth in the steady state?

5.4 **Measurement of the Capital Stock**

The capital stock is not easy to measure. A common way to try to do this is to:

- guess a value for K_0 at some date in the past (maybe when statistics were first collected),
- measure investment every year thereafter (which is also not that easy), and
- use the equation

$$K_{t+1} = (1 - \delta) K_t + I_t \tag{5.5.1}$$

to compute the capital stock for every year thereafter.

Consider an economy that is in a steady state without technological progress (and has been there for an infinitely long time). Every year the level of investment is:

$$I_t = 0.2$$

and the depreciation rate is:

$$\delta = 0.1$$

(a) What is the capital stock?

(b) The statistical office only started measuring investment in year 0. By then the economy was already in steady state. They guessed (incorrectly) that the capital stock in year 0 was $K_0 = 1$. Use equation (5.5.1) to compute how the estimate of the capital stock changed over time after year 0 and plot your answer. (You don't need to provide a closed form solution, a nice graph made with a spreadsheet is enough). Compute the ratio:

$$\frac{K_{ESTIMATE}}{K_{TRUE}}$$

for the years 5, 10, and 50. (Again, you can just take the numbers from a spreadsheet).

(c) Suppose now that the statistics office started collecting statistics a long time ago (an infinite time ago) so we don't have the problem of making the wrong initial guess. However, instead of using the correct value of δ, the statistical office incorrectly believes that $\delta = 0.05$. What is their estimate of the capital stock? Why does it differ from the truth?

(d) The production function in this economy is

$$Y = AK^{\alpha}L^{1-\alpha} \tag{5.5.2}$$

and there are accurate data on GDP and on total hours of labor supply. Using the inaccurate estimate of K from part (c), plus accurate data on total labor and GDP, an economist is trying to measure the economy's productivity, i.e. to solve for A in equation (5.5.2). Find an expression for

$$\frac{A_{ESTIMATE}}{A_{TRUE}}$$

as a function of

$$\frac{K_{ESTIMATE}}{K_{TRUE}}$$

Given the numbers from part (c) and using $\alpha = 0.35$, how far would the economist's estimate be from the truth?

5.5 **Sources of TFP**

Suppose that the true production function in Gotham is

$$Y = AK^{\alpha}L^{1-\alpha} \tag{5.5.3}$$

Unfortunately, crime is a huge problem in Gotham, so that for each worker doing actual work firms need to hire γ security guards just to protect their products from being stolen. The security guards will of course describe their activity as "work" even though they are not actually producing anything. Use the notation N to refer to the *total* labor force (including workers and guards) and denote the number of actual production workers by L.

(a) Find an expression for total output as a function of A, K, N, γ and α.

(b) Write down the problem of a firm that has to choose capital and labor to maximize profits. Notice that the firm will have to pay a wage to the security guards even though they will not produce anything.

(c) If the representative firm hires all the workers and rents all the capital, what will be the wage and the rental rate of capital? Express it as a function of A, K, N, γ and α.

(d) Suppose an economist studying Gotham is trying to estimate A using equation (5.5.3). The economist has accurate data on K, N and Y. However, the economist doesn't really know whether workers are involved in production or in security services: in national statistics they all look "employed". Therefore the economist will plug in the value of N instead of the value of L into the estimate of A. What will the economist's estimate of A be? How does it compare to the true value of A?

(e) How does this relate to the findings that link GDP levels to social and political institutions?

(f) Suppose that, in a economy that didn't have the crime problem of Gotham, the government attempted to "create jobs" by mandating that firms hire γ "assistants" for every production worker. The job of assistants is to look at production workers all day long and not do anything. Using the analysis above, what would be the effects of such a policy?

5.6 **Interest Rates**

Suppose we observe that Usuria (a closed economy) grows at approximately 6% a year, and we are trying to understand why. We know the labor force has been constant.

- Conjecture 1: The economy has been at a steady-state-with-technological-progress. There has been capital accumulation just to maintain $\frac{K}{AL}$ constant but the cause of growth has been TFP growth.
- Conjecture 2: The economy started from a very low level of capital stock (below steady state) and has been growing because it is converging to the steady state, but TFP has been constant.

Ideally, if we wanted to distinguish between Conjecture 1 and Conjecture 2 we could do a growth-accounting exercise. Unfortunately, Usuria does not collect reliable statistics on capital accumulation that would enable us to do this. We do, however, have data on interest rates in Usuria. How would one use these data to distinguish between Conjecture 1 and Conjecture 2? Be as mathematically precise as possible.

5.7 **GDP Accounting, Interest Rates and Growth Accounting**

Proletaria is a mythical country in Central Asia, where the currency is the ruble. The exchange rate is 10 rubles per euro.

Suppose the production function in Proletaria is given by:

$$Y = K^{\alpha} L^{1-\alpha}$$

(a) Find expressions for the output-to-capital ratio $\frac{Y}{K}$ and the marginal product of capital F_K as functions of K and L.

During 2014, the following events took place in Proletaria.

- Kapitas, the main manufacturer in the country, imported an industrial welder made in Germany, for which it paid 1,000 euro.
- By the end of the year, the industrial welder was no longer new. Its estimated value in the resale market was 960 euro.
- 100 workers worked for Kapitas the entire year making screws and nails, using the new welder. Each of them received wages for 350 rubles.
- Each of the Kapitas workers paid 100 rubles in income taxes.
- The total output of Kapitas consisted of 1 million screws and 1 million nails. All of it was exported to Austria for a total of 4,000 euro.
- The government of Proletaria employed one of the 100 workers as Chief of the Secret Police (in addition to his factory job) to maintain law and order, and paid her a salary of 10,000 rubles.
- The workers ate beef imported from France, which cost a total of 2,500 euro.

(b) Construct GDP accounts (in rubles) for Proletaria by production, income and expenditure.

(c) What were the income shares of labor and capital?

(d) Suppose that we know that the capital stock in Proletaria is 100,000 rubles and that the depreciation rate of the industrial welder is typical for this country's capital stock. What interest rate should we expect to observe?

(e) During 2015, additional investment in Proletaria has exceeded depreciation so that the capital stock now stands at 120,000 rubles. The labor force also grew thanks to immigration, and now consists of 105 workers instead of 100. GDP during 2015 was 55,000 rubles (prices were constant). How much of the growth in GDP between 2014 and 2015 can be attributed to growth in Total Factor Productivity?

5.8 National Accounts and The Golden Rule

In 2018, the following events happened in Aurum, which is a closed economy that uses the Denarius (plural: Denarii) as its currency.

- Grano, Inc. hired Cornelia to plant and harvest 500 tons of wheat on a large plot of land using a tractor. Both the land and the tractor are owned by Grano, Inc.
  - It paid Cornelia 200 Denarii for her work on the harvest.
  - It sold the wheat to Panem, Inc. for 1,000 Denarii.
- The tractor the Cornelia used:
  - Was new at the beginning of the year. It was worth 2,000 Denarii.
  - Was no longer new at the end of the year. By then it was worth 1,800 Denarii.
- Panem, Inc. hired Gaius to grind the wheat in its mill and produce a million loaves of bread.
  - It paid Gaius 1,400 Denarii for his work.
  - It sold the bread for 4,000 Denarii to the hungry citizens, who eat it.

- Fabrica, Inc. hired Livia to build a combine harvester.
  - It paid Livia 400 Denarii for her work.
  - It sold the combine harvester to Grano, Inc. for 1,000 Denarii.

(a) Construct GDP accounts (in Denarii) for Aurum by production, income and expenditure.

(b) What were the income shares of labor and capital?

(c) What was the savings rate?

(d) Assume the following:

- the savings rate is constant,
- the production function has the form:

$$Y = AK^{\alpha}L^{1-\alpha}$$

  and there is no technological progress,
- the depreciation rate for the tractor is typical for the capital stock as a whole,
- the population is constant.

In the long run, would the level of consumption be higher if the economy slightly increased its saving rate?

5.9 **Disease and TFP**

The production function in Influenzistan is:

$$Y = K^{\alpha}(AH)^{1-\alpha} \tag{5.5.4}$$

where H is the total number of hours of work.

(a) Solve equation (5.5.4) for A, so that you end up with an expression for A in terms of all the other variables and parameters.

(b) Suppose an economist wants to measure A. The economist has official data from the Influenzistan Statistics Office on:

- total hours worked by all workers,
- the total level of investment, which has been constant for many many years,
- an estimate of the average rate of depreciation of the capital stock,
- GDP and its subcomponents, measured by the income method.

How can he use this data to get an estimate of A? Don't be vague: provide exact formulas that describe each step precisely.

(c) Suppose now that the economist has the same data as above except he doesn't know total hours worked. Instead, the Influenzistan Statistics Office only keeps records on the total number of employed workers (call this number L), but not on how many hours each of them works. In order to get an estimate of A, the economist assumes that the average worker works N hours per year, which is the number of hours worked by the average worker in neighboring Healthistan. Write down an expression for the economist's estimate of A (call this number $\hat{A}$).

(d) Unfortunately, the residents of Influenzistan keep getting sick, so they have to spend a significant portion of their time recovering at home instead of working. As a result, they each work xN hours instead of N hours, where $x < 1$. Derive an expression for $\frac{\hat{A}}{A}$, i.e. for the ratio of the economists' estimate of A to its true value. Interpret your finding.

PART III

Microeconomic Foundations

This part of the book looks at the microeconomics that is the basis of modern macroeconomics. The Solow growth model assumes that the economy saves an exogenous fraction of its output and, because everyone works, the size of the labor force is exogenous. In this part of the book, we analyze the economic forces that shape individuals' decisions to consume, save, work and invest and how they all fit together.

In Chapter 6, we think about intertemporal decision making. Consumption and savings decisions involve the relationship between the present and the future. We study a simple two-period model of these decisions and use it to think about the evidence on how people make these decisions. We then extend it to think about more than two periods.

In Chapter 7, we think about the labor market. First we go over some of the statistics that are used to measure the labor market. We then study a model of how people divide their time between the labor market and everything else, and how labor market equilibrium is determined. Finally, we think of reasons why the labor market might not clear, in order to think about unemployment.

In Chapter 8, we look at investment decisions and how they relate to expectations about the future, prices of future goods, uncertainty, and asset prices.

In Chapter 9, we study general equilibrium: how all the decisions by households and firms fit together to determine economic outcomes. We also study the First Welfare Theorem, which provides conditions under which a market economy is Pareto Efficient.

CHAPTER 6

Consumption and Saving

6.1 The Keynesian View of Consumption

Keynes (1936) made the following observation about consumption decisions:

> "The fundamental psychological law, upon which we are entitled to depend with great confidence both a priori from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as their income increases but not by as much as the increase in the income."

What does this mean? Is this theory correct? Let's first translate this theory into mathematical language and then try to assess it. First of all, it is stating that how much people consume depends on their income, which seems reasonable enough. Let's use C to denote consumption and Y to denote income. Keynes says that there is a function $c(\cdot)$ (sometimes known as a consumption function) that relates consumption to income:

$$C = c\left(Y\right) \tag{6.1.1}$$

Furthermore, he is saying that consumption depends on income in a specific way. Here Keynes' language is more ambiguous but one possible interpretation is that he is saying that:

$$c'\left(Y\right) < 1$$

so when income rises by one dollar, consumption rises but by less than one dollar. The quantity $c'(Y)$ is known as the "marginal propensity to consume". It measures how much of an extra dollar of income is dedicated to consumption.

Possibly, depending on how one interprets his language, Keynes is saying something more: that if income rises 1%, consumption rises, but by less than 1%, so the *elasticity* of consumption with respect to income is positive but less than 1. Mathematically:

$$\frac{c'(Y)Y}{c(Y)} = \frac{\partial \log(c(Y))}{\partial \log(Y)} < 1$$

One way to test this conjecture is to take a sample of households, measure their income, measure their consumption and see whether the best fit of equation (6.1.1) has $c'(Y) < 1$ and/or $\frac{\partial \log(c(Y))}{\partial \log(Y)} < 1$. Figure 6.1.1 shows the result of doing precisely that. The Consumer Expenditure Survey asks a sample of households to report their income and their consumption (among other things). The figure shows scatterplots of consumption against income for these households, both in absolute terms and in logarithmic scale, to measure $c'(Y)$ and $\frac{c'(Y)Y}{c(Y)}$ respectively.[1] The evidence seems consistent with both interpretations of Keynes's statement: both the best fit estimate of $c'(Y)$ and $\frac{c'(Y)Y}{c(Y)}$ are lower than 1. $c'(Y)$ is approximately 0.25, so households whose income is one more dollar spend an additional 25 cents. $\frac{c'(Y)Y}{c(Y)}$ is approximately 0.55, so households with 1% higher income spend approximately 0.55% more.

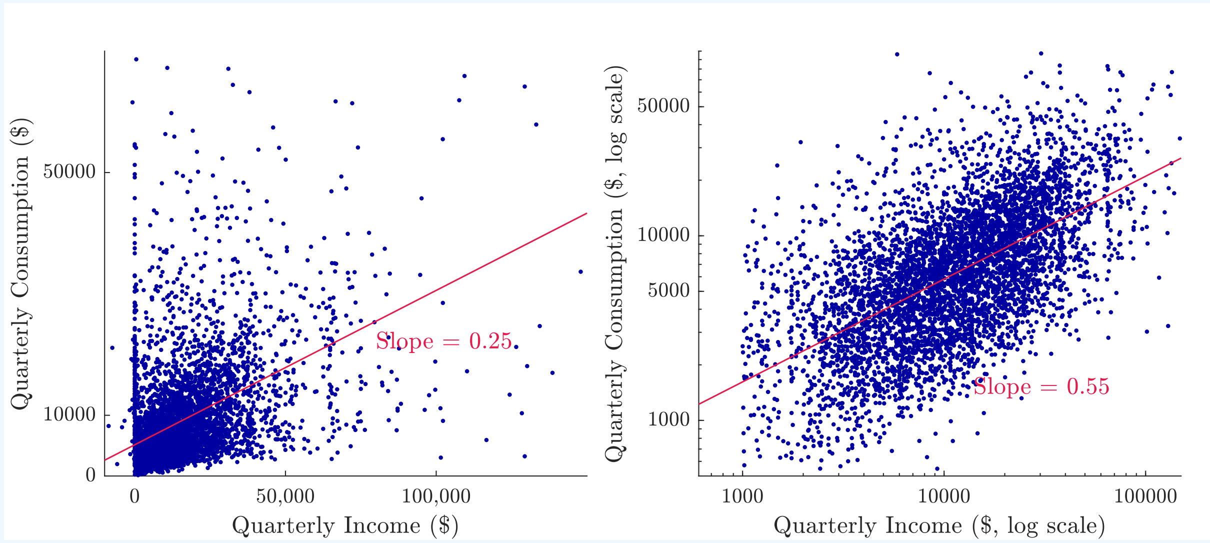


Fig. 6.1.1: *Evidence on the Keynesian consumption function. Each dot represents a household. Source: Consumer Expenditure Survey, 2014.*

For some time, around the mid-20th century, this type of evidence was considered quite conclusive, leading to a firm belief in the Keynesian consumption function as a good description of consumption behavior. This led to following kind of speculation: what is going to happen as the economy's productive capacity expands over time? If the elasticity of consumption with respect to income is less than 1, this implies that over time, as income increases, the ratio $\frac{C}{Y}$ will fall. Is the economy going to produce more and more goods that nobody wants to consume? What are we going to do with all these goods?[2] Is there going to be massive unemployment because nobody wants all the stuff that we'd produce if everyone was working?

Aggregate data gives us a way to test this conjecture. Figure 6.1.2 shows the relationship between aggregate consumption and aggregate income from national accounts. In the left panel we see the relationship in the United States, where each dot represents a different year. The best fit estimate of $\frac{c'(Y)Y}{c(Y)}$ is 0.97. There is a

[1] The logarithmic scale graph only includes households with quarterly income of at least $1,000.

[2] You can see echoes of this preoccupation in Orwell's famous novel *1984*. It was not uncommon to interpret war, and the huge destruction that is brings about, as a "solution" to the "problem" of over-production.

simple explanation for this: consumption has been close to a constant fraction of GDP, approximately 65%. If $c(Y) = 0.65Y$ then $\frac{c'(Y)Y}{c(Y)} = 1$. The right hand panel shows the relationship across countries, where each dot represents a different country for the year 2011. In this case, the best-fit estimate of $\frac{c'(Y)Y}{c(Y)}$ is 0.85. In both cases the estimate is much closer to 1 than in the individual household data. Overall, it does *not* seem to be the case that countries consume a lower fraction of their income as they grow rich.[3]

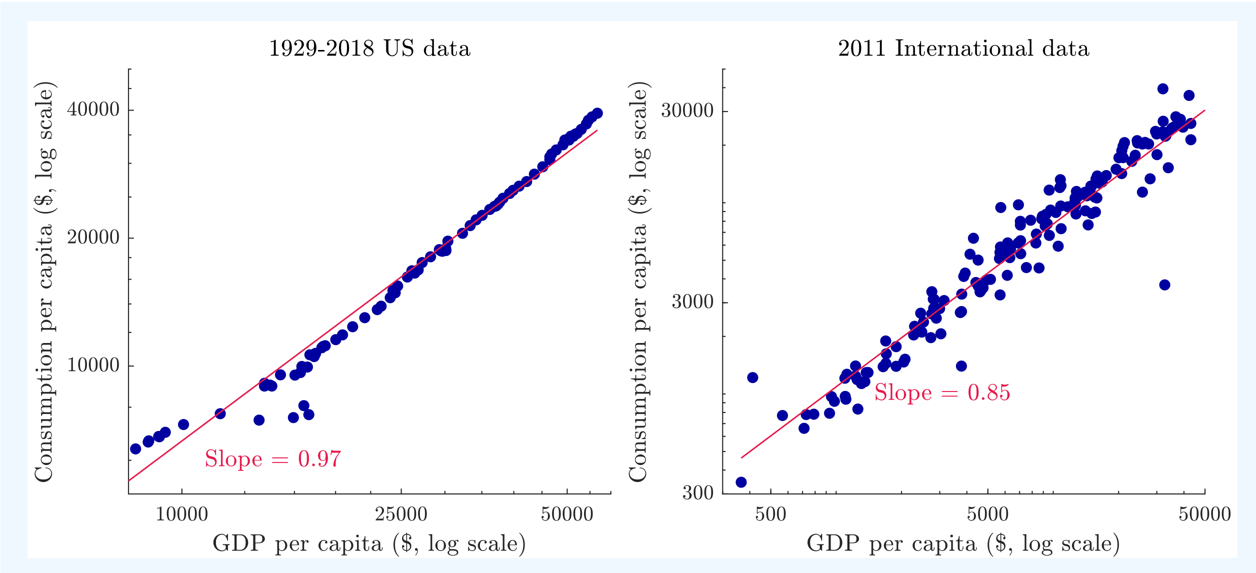


Fig. 6.1.2: *Evidence on the Keynesian consumption function from aggregate data. The left panel is US time-series evidence; the right panel is cross-country evidence. Sources: NIPA and Feenstra et al. (2015)*

In the aggregate data over time we don't see the pattern that we see in the cross-sectional data. The preoccupation about decreasing consumption rates over time seems to be unwarranted. What is going on? Why do the two kinds of data look so different?

6.2 A Two-Period Model of Consumption

Let's take a step back and try to develop a theory of how households make consumption decisions and see whether this can help us understand some of the patterns we just saw. We'll start from a very simple example and then think about more features.

Let's imagine that this household is going to live for two periods. In period 1 they will obtain income y_1 and in period 2 they will obtain income y_2. They have to decide how much they are going to consume in period 1. The advantage of consuming is that they like to consume; the advantage of not consuming is that by saving they can afford to consume more in period 2, which they also like. Let's assume that their preferences

[3]Note the possibility of reverse causality in the cross-country data. The Solow model predicts that, other things being equal, countries that choose to save more and consume less will have higher GDP. This will produce an estimate of $\frac{c'(Y)Y}{c(Y)}$ lower than 1 even if the true elasticity is equal to 1.

are described by the following utility function:

$$U(c_1, c_2) = u(c_1) + \beta u(c_2) \tag{6.2.1}$$

This way of thinking about the consumption decision makes it mathematically equivalent to the kind of two-good consumption problem studied in microeconomics: here the two goods are c_1 ("consumption in period 1") and c_2 ("consumption in period 2"). In addition, for simplicity, we are assuming that the utility function is additively separable in the two goods and that the only difference in how much they care about each of them is the term β. β is just some number, typically assumed to be less than 1 to represent impatience: the same level of consumption gives the household more utility if it comes now than if it comes in the next period.

Now that we have preferences, we need to think about the household's budget. How much of each of the two goods can the household afford? What is their relative price? Imagine that the household consumes c_1 in the first period. This means it can save:

$$a = y_1 - c_1 \tag{6.2.2}$$

The household earns interest on these savings, so by saving a, in period 2 it can afford to consume:

$$c_2 = y_2 + (1 + r)a \tag{6.2.3}$$

where r is the real interest rate between periods 1 and 2. Replacing a from (6.2.2) into (6.2.3) and rearranging, we get

$$c_1 + \frac{1}{1+r}c_2 = y_1 + \frac{1}{1+r}y_2 \tag{6.2.4}$$

Equation (6.2.4) is a standard budget constraint for a two-good consumption problem.

The term $\frac{1}{1+r}$ is the price of period-2 goods in terms of period-1 goods. Why does this make sense? If you sell one period-1 good to the market, the market is willing to give you $(1 + r)$ period-2 goods. This is exactly what a price means: at what rate is the market willing to exchange one good for another. High interest rates mean that period-1 goods are expensive relative to period-2 goods: the market is willing to provide a lot of period-2 goods in exchange for period-1 goods. Conversely, low interest rates mean that period-1 goods are cheap.

The right hand side of (6.2.4) is the household's total budget. Why? The household's income consists of y_1 period-1 goods and y_2 period-2 goods. Adding them up at their respective market prices tells us how much the household can afford in total. The expression $y_1 + \frac{1}{1+r}y_2$ is called the "present value" of income and likewise the expression $c_1 + \frac{1}{1+r}c_2$ is called the "present value" of consumption. The idea of a "present value" is to express future quantities in terms of the amount of the present goods that the are equivalent to at the relevant market prices. Here the price of future goods is $\frac{1}{1+r}$ so we multiply by this term in order to add them to present goods.

The term a in equations (6.2.2) and (6.2.3) represents savings, but we haven't said anything yet about whether a needs to be a positive number. Do equations (6.2.2)-(6.2.4) also apply when $a < 0$? It depends on what we think about the household's ability to borrow. If we assume that the household can borrow as much as it wants at the interest rate r (and must always pay back its debts), then it is OK to allow for negative values of a, and the budget constraint (6.2.4) still applies. $a < 0$ simply means that the household is borrowing in order to pay for $c_1 > y_1$. For now we'll make this assumption; later on we'll think about what happens when the household cannot borrow.

We are going to imagine that the household takes as given its current and future income y_1 and y_2 and the interest rate and simply solves a standard consumer optimization problem:[4]

$$\begin{gathered} \max u(c_1) + \beta u(c_2) \\ s.t. \\ c_1 + \frac{1}{1+r}c_2 \leq y_1 + \frac{1}{1+r}y_2 \end{gathered} \tag{6.2.5}$$

Figure 6.2.1 shows the solution to problem (6.2.5). As is standard in microeconomics, the household will choose the highest indifference curve it can afford, which implies that it will pick a point where the indifference curve is tangent to the budget constraint. Notice two properties of the budget constraint. First, its slope is $-(1+r)$. As usual, the slope of the budget constraint is the relative price. Higher interest rates mean a steeper budget constraint. Second, the budget constraint goes through the point (y_1, y_2) since the household has the option to just consume its income each period.

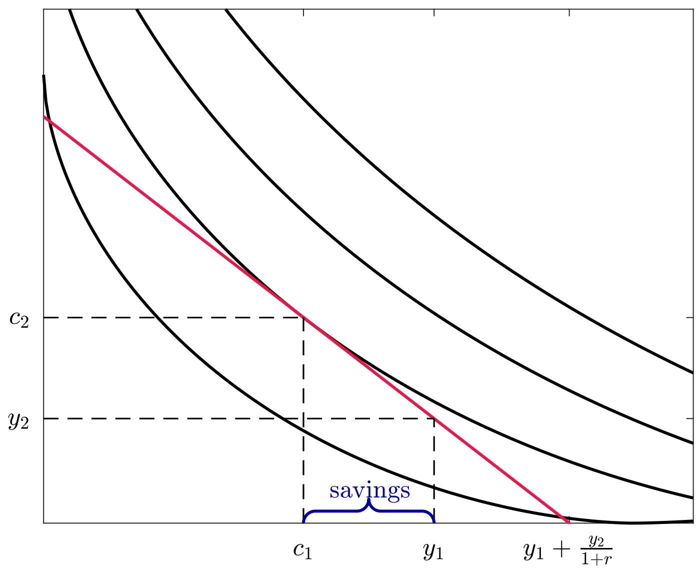


Fig. 6.2.1: *The consumption-savings decision as a two-good consumption problem.*

We can also find the solution to problem (6.2.5) from its first order conditions. The Lagrangian is:[5]

$$L(c_1, c_2, \lambda) = u(c_1) + \beta u(c_2) - \lambda \left[c_1 + \frac{1}{1+r}c_2 - y_1 - \frac{1}{1+r}y_2 \right]$$

[4] There is some disagreement about whether budget constraints should be written as equalities or as weak inequalities. I like the version with weak inequality because it says that the household could, in principle, not spend all its income. Since this never happens anyway, it's not a big deal which way we write it.

[5] This problem is sufficiently simple that we don't need to use a Lagrangian to solve it. We could just as easily replace

λ is the Lagrange multiplier of the budget constraint. It has the usual interpretation of the marginal utility of a unit of wealth. The first order conditions of the problem are:

$$u'(c_1) - \lambda = 0 \tag{6.2.6}$$

$$\beta u'(c_2) - \lambda \frac{1}{1+r} = 0 \tag{6.2.7}$$

Solving equation (6.2.6) for λ, replacing in equation (6.2.7) and rearranging we obtain:

$$u'(c_1) = \beta(1+r)\,u'(c_2) \tag{6.2.8}$$

Equation (6.2.8) is known as an "Euler equation" and plays a central role in modern macroeconomics. It describes how households trade off the present against the future, and has the following interpretation. Suppose a household is deciding whether to allocate one unit of wealth to consumption or to save it for the future. If it consumes it, it will obtain the marginal utility of present consumption, $u'(c_1)$. This gives us the left hand side of (6.2.8). If instead the household saves, it obtains $(1+r)$ units of future wealth because the market pays interest. Each unit of future wealth gives the household the marginal utility of future consumption, which is $u'(c_2)$, multiplied by β to account for the household's impatience. This gives us the right hand side of (6.2.8). At the margin, the household must be indifferent between allocating the last unit of wealth between these two alternatives, so (6.2.8) must hold. (6.2.8) is also the algebraic representation of the tangency condition shown in Figure 6.2.1. The slope of the household's indifference curve is given by the marginal rate of substitution between period-1 consumption and period-2 consumption: $\frac{u'(c_1)}{\beta u'(c_2)}$. The slope of the budget constraint, as we saw, is $1+r$, so (6.2.8) says that the two are equated.

Some Examples

Figure 6.2.2 shows two possible patterns of income over time and the household's consumption decision in each case. The left panel shows a household for whom period 1 represents their working age and period 2 represents a time in which they are planning to retire, so $y_2 = 0$. Understanding that their income will be low in the future, they choose $a = y_1 - c_1 > 0$ in order to be able to consume while they are retired. The right panel represents the opposite case. Here the household has low y_1 and much higher y_2, so they are optimistic about future income compared to current income. In this example the household chooses $a = y_1 - c_1 < 0$, so they borrow to consume more than their income.

Both examples have some features in common. In both cases expectations about the future, not just current income, affect consumption decisions. In both cases the household is trying to even out or "smooth" consumption over time, i.e. they are using borrowing and saving to prevent their consumption from moving up and down as much as income does.

$c_2 = y_2 + (1+r)(y_1 - c_1)$ into the objective function and solve:

$$\max_{c_1} u(c_1) + \beta u(y_2 + (1+r)(y_1 - c_1))$$

which gives of the same solution.

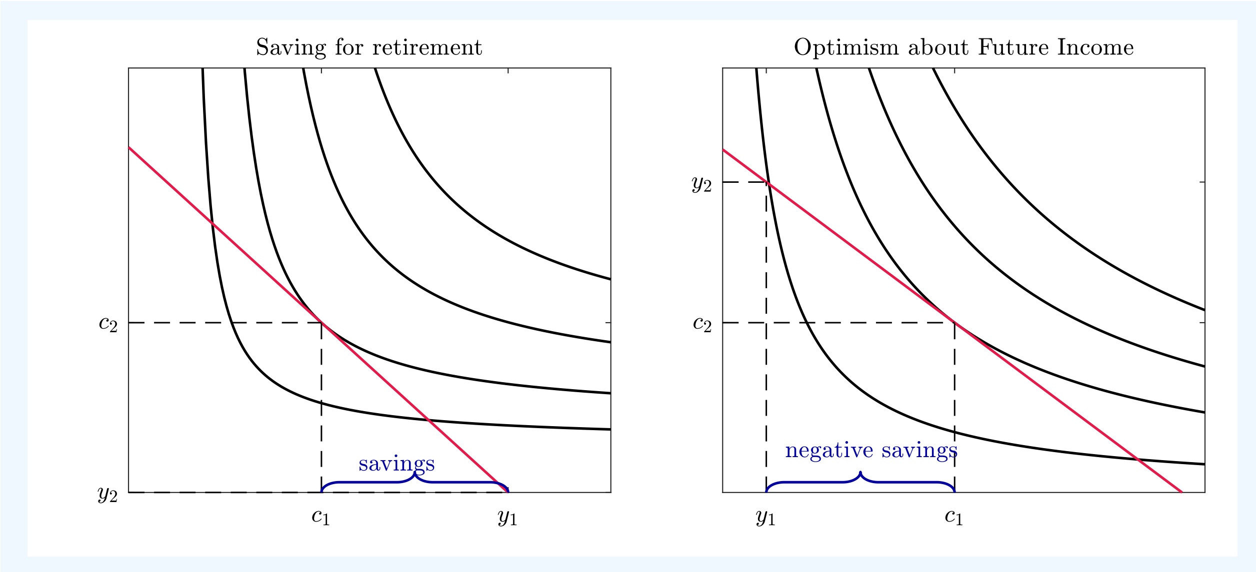


Fig. 6.2.2: *Consumption decisions in two examples.*

The Effect of Interest Rates

Let's imagine that interest rates change. How do households change their consumption? The answer to this question is going to play an important role in some of the models of the entire economy that we'll analyze later. For now, we are going to study the question in isolation, just looking at the response of an individual household to an exogenous change in the interest rate. For concreteness, let's imagine that the interest rate rises.

Let's first take a look at this question graphically. A change in interest rates can be represented by a change in the budget constraint, as in Figure 6.2.3. The new budget constraint still crosses the point (y_1, y_2) because the household can afford this no matter what the interest rate is, but the slope of the budget constraint is different. With higher interest rates, it becomes steeper. As with any change in prices, this can have both income and substitution effects.

The substitution effect is straightforward: as we saw before, a higher interest rate means that present goods have become more expensive relative to future goods. Other things being equal, this would make the household substitute away from present goods towards future goods, i.e. save more and consume less.

The income effect is a little bit more subtle. Do higher interest rates help or hurt the household? That depends on whether the household is borrowing or saving to begin with. If the household is saving, then higher interest rates mean that it is earning more on its savings, which can only help them attain higher utility. This is the case depicted in Figure 6.2.3. Instead, if the household was borrowing, then higher interest rates means that it's paying more interest on its loans, which hurts them.[6]

Graphically, it's possible to decompose income and substitution effects in the following way. First, imagine

[6]There is an additional possibility, which is that the household was choosing to borrow when interest rates were low but saves instead when the interest rate rises. In this case the income effect could go either way.

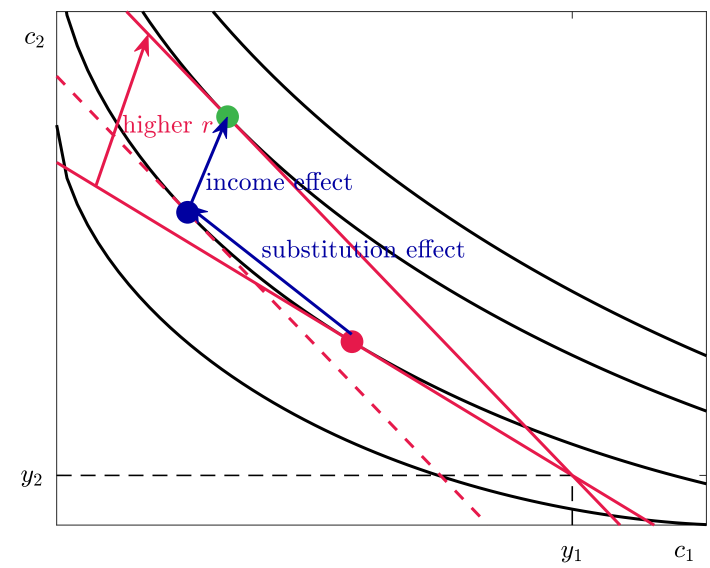


Fig. 6.2.3: *Consumption response to higher interest rates.*

changing the interest rates (and therefore the slope of the budget constraint) but adjusting the position of the budget constraint so that the household can attain the original indifference curve and ask how much of each good the household consumes. This is a way of isolating the substitution effect: how much the household rebalances between present and future consumption due to the new prices while holding utility constant. Second, move the budget constraint from the adjusted line to the actual new budget constraint. The difference between the household's consumption at the adjusted budget and the true new budget measures the income effect: at the same prices, how much more or less can the household afford.[7]

Let's go back to the question of how consumption reacts to a rise in the interest rate. We know that the substitution effect would make consumption go down and the income effect could go either way. When the income effect is negative, then both income and substitution effects go in the same direction and we know that consumption falls when interest rates rise. When the income effect is positive, then income and substitution effects are pushing in opposite directions and the net effect could go either way. Figures 6.2.4 shows examples where each of these things happen. On the left panel is the "saving for retirement" example. Here the household is saving so the income effect of higher interest rates is positive, and in fact stronger than the substitution effect, so the household increases its consumption. On the left panel is the "optimism" example, where the household was borrowing against its high future income. Here the income effect of higher interest rates is negative, and reinforces the substitution effect, leading to lower consumption.

[7]This way of decomposing income and substitution effects is known as the Hicks decomposition. An alternative is the Slutsky decomposition, where the substitution effect is measured at the budget such that the original *consumption plan* is affordable instead of the original *utility level*.

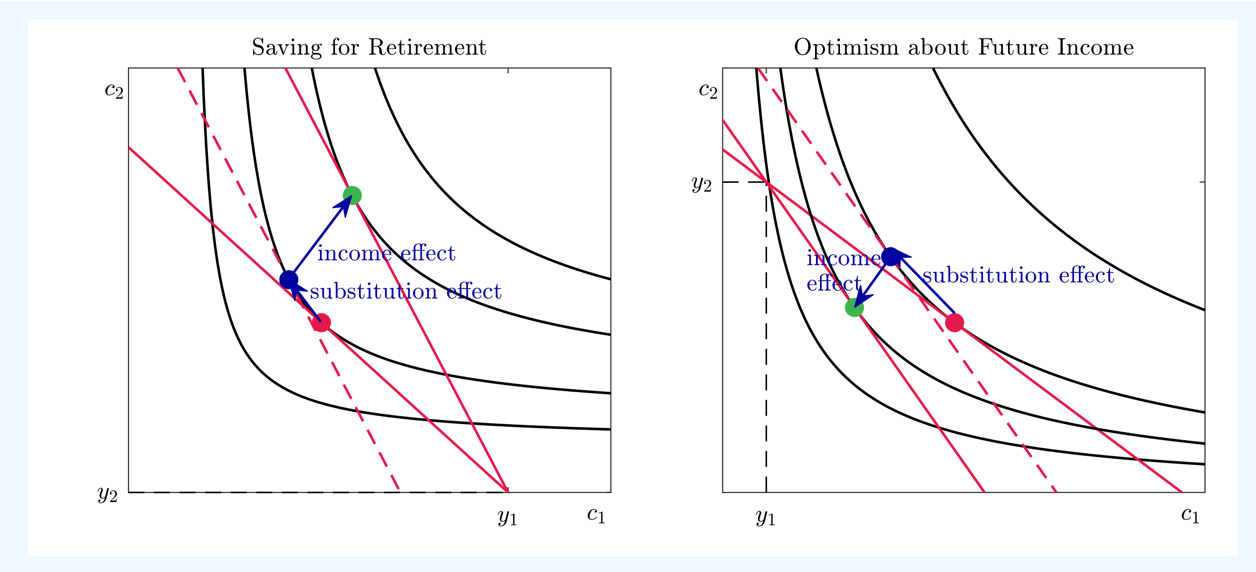


Fig. 6.2.4: *Higher interest rates in two examples.*

An Explicit Example

If preferences take the CRRA form we can go beyond equation (6.2.8) and get an explicit formula for how much the household is going to consume.[8] CRRA utility takes the form:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

so marginal utility is:

$$u'(c) = c^{-\sigma}$$

Replacing this in equation (6.2.8) gives us:

$$c_1^{-\sigma} = \beta(1+r)\,c_2^{-\sigma}$$
$$\Rightarrow c_2 = [\beta(1+r)]^{\frac{1}{\sigma}}\,c_1$$

Now replace c_2 into the budget constraint (6.2.4):

$$c_1 + \frac{1}{1+r}[\beta(1+r)]^{\frac{1}{\sigma}}\,c_1 = y_1 + \frac{1}{1+r}y_2$$

and solve for c_1:

$$c_1 = \frac{y_1 + \frac{1}{1+r}y_2}{1+\beta^{\frac{1}{\sigma}}(1+r)^{\frac{1}{\sigma}-1}} \tag{6.2.9}$$

[8]CRRA stands for "constant relative risk aversion". We first encountered this functional form in Chapter 2.

Equation (6.2.9) gives us an explicit formula for how consumption depends on present income, future income and interest rates. The numerator is the present value of income: consumption is proportional to this. The denominator captures the effects of the household's impatience (measured by β), the relative price of period-1 consumption $(1+r)$ and the household's willingness to substitute consumption in one period for consumption in the other (measured by σ).

The Permanent Income Hypothesis

The model we have been analyzing gives an alternative hypothesis to Keynes's view of how households make consumption decisions. In this model, consumption does not really depend on current income. Instead, it depends on the total value of the household's income over time. We can see this directly in the budget constraint (6.2.4). What determines how much consumption the household can afford is not y_1 but rather $y_1 + \frac{1}{1+r} y_2$. That same expression shows up in the numerator of equation (6.2.9).

One way of putting this into words is to say that consumption depends on "permanent income", i.e. some sort of average level of income over time. The budget constraint (6.2.4) tells us exactly what's the right way to take the average: by weighting each period's income by the relevant price and thus computing a present value. But this is a detail. The broader point is that what matters is average income and not any one period's income. This way of thinking about consumption became known as the "permanent income hypothesis" (the term is due to Friedman (1957)). Let's see what this means.

Suppose we ask: how much does the household's consumption increase if it finds out that its income has increased? In this model, the answer is "it depends". In particular, "it depends on whether the increase in income is perceived as temporary or permanent". Formula (6.2.9) makes this clear.

Suppose that y_1 increases but the household does not change its expectation of what y_2 is going to be. In other words, this is perceived as a temporary increase, for instance because this is a worker who just received a one-time bonus. How much is consumption going to respond?

$$\frac{\partial c_1}{\partial y_1} = \frac{1}{1 + \beta^{\frac{1}{\sigma}} (1+r)^{\frac{1}{\sigma}-1}} < 1 \tag{6.2.10}$$

This is the marginal propensity to consume, the answer to the question: how much extra consumption does the household choose when it gets an additional unit of (temporary) income? Equation (6.2.9) tells us that the marginal propensity to consume is less than 1. In other words, the household will increase its consumption by less than the increase in income, as Keynes believed.

Suppose instead that the household perceives that the increase in income is permanent, so that y_2 increases by the same amount as y_1, for instance because this is a worker who has just received a permanent raise. Let's say the permanent raise is by some amount Δ. How much is consumption going to respond?

$$\frac{dc_1}{d\Delta} = \frac{1 + \frac{1}{1+r}}{1 + \beta^{\frac{1}{\sigma}} (1+r)^{\frac{1}{\sigma}-1}} \tag{6.2.11}$$

Comparing formulas (6.2.10) and (6.2.11) immediately tells us that the marginal propensity to consume out of permanent income is higher than the marginal propensity to consume out of temporary income. If we make the further simplification that $\beta(1+r)=1$, then formula (6.2.11) simplifies and we get:[9]

$$\frac{dc_1}{d\Delta}=1$$

In other words, when the household perceives an increase in income as being permanent, it increases its consumption by the same amount of the increase in income.

Reconciling with the Data

Let's go back to Figures 6.1.1 and 6.1.2. The permanent income hypothesis gives us a way to reconcile this seemingly contradictory evidence.

If the permanent income hypothesis is correct, then the households with *temporarily* low income will have relatively high $\frac{c}{y}$ because their permanent income is higher than their current income. Conversely, the households with *temporarily* high income will have relatively low $\frac{c}{y}$ because their permanent income is lower than their current income.

The data in Figure 6.1.1 shows plots consumption and and *total* income in a single year, without making the distinction between temporary and permanent. Total income probably reflects a mixture of of temporary and permanent factors. Suppose we look at households who had high income this year. Some of them will be in this group because they are always in this group (i.e. they have high permanent income). Others will be in this group because they had a particularly good year, but their permanent income is not that high. Conversely, if we look at households who had low income this year, this will include households with low permanent income and households with not-so-low permanent income who had a bad year. On average, households with higher *total* income are more likely to have had a better-than-usual year. According to the permanent income hypothesis, these are precisely the households that should have lower $\frac{c}{Y}$, resulting in the lower-than-one slopes that we observe.

Once we add up over many people and long periods of time, then the temporary components average out. Some people in the sample will have had a good year but others will have had a bad year, so measured average income becomes closer to permanent average income. The permanent income hypothesis tells us that consumption should be proportional to permanent income, so we expect $\frac{c'(Y)Y}{c(Y)}=1$, which is indeed close to what we observe.

Taxes and Ricardian Equivalence

One of the things we are going to be interested in is how the economy reacts to changes in policies in general and taxes in particular. One ingredient in answering that question is to figure out how household consumption will respond to taxes.

We are going to assume that the government wants to make purchases of goods and services equal to G_1 and G_2 in periods 1 and 2 respectively. For now we are not going to ask why or how the government chooses G_1 and G_2, we'll just take them as given. In order to pay for this spending, the government is going to collect

[9] $\beta(1+r)=1$ means that the household's impatience and the market interest rates exactly offset each other. If we go back to the Euler equation (6.2.8), this implies that $c_1=c_2$.

taxes τ_1 and τ_2 from the household in periods 1 and 2 respectively. We'll assume that these taxes are "lump sum", meaning that nothing that the household does affects how much tax it owes. The government need not exactly balance its budget in each period. Just like the household, it can borrow and save as much as it wants at the interest rate r, but must pay its debts.

The government budget constraint can be written as:

$$B = G_1 - \tau_1 \tag{6.2.12}$$

where B is the amount that the government borrows, equal to the difference between the amount it spends in period 1 and the taxes it collects. If $B < 0$, this means the government is saving. In period 2, the budget is:

$$\tau_2 = G_2 + (1 + r)\, B \tag{6.2.13}$$

The government has to collect enough taxes to pay for spending and also pay back its debt with interest. Solving (6.2.12) for B, replacing in (6.2.13) and rearranging we get:

$$G_1 + \frac{1}{1+r} G_2 = \tau_1 + \frac{1}{1+r} \tau_2 \tag{6.2.14}$$

Equation (6.2.14) has the same interpretation as the household's budget (6.2.4). Total government revenue (in present value) is on the right hand side and must be equal to total government spending (in present value), which is on the left hand side.

Now that the household has to pay taxes, its budget changes because it can only use after-tax income to pay for consumption. Instead of (6.2.4), the household's budget is now:

$$c_1 + \frac{1}{1+r} c_2 = y_1 - \tau_1 + \frac{1}{1+r} (y_2 - \tau_2) \tag{6.2.15}$$

Adding up (6.2.14) and (6.2.15) we get that the household's budget is:

$$\underbrace{c_1 + \frac{1}{1+r} c_2}_{\text{Present value of consumption}} = \underbrace{y_1 + \frac{1}{1+r} y_2}_{\text{Present value of income}} - \underbrace{\left(G_1 + \frac{1}{1+r} G_2\right)}_{\text{Present value of government spending}} \tag{6.2.16}$$

The logic behind equation (6.2.16) is as follows: the budget of the household is equal to the present value of income minus the present value of taxes. But the government budget implies that the present value of taxes equals the present value of spending. Therefore the household's budget must equal the present value of income minus the present value of government spending.

Equation (6.2.16) has one important implication because of what's *not* in it. τ_1 and τ_2 don't appear in the equation. Of course, τ_1 and τ_2 have to be such that the government's budget constraint (6.2.14) is satisfied but any combination of τ_1 and τ_2 that satisfies (6.2.14) is equivalent from the point of view of the household, i.e. the timing of taxes does not matter. This property is known as "Ricardian equivalence", after David Ricardo, a 19th century economist who first discussed it.

How does Ricardian equivalence come about? Imagine that the government announces that it is going to lower taxes τ_1 but leave spending G_1 and G_2 unchanged. Upon hearing this announcement, people immediately calculate the implications for the government budget and come to the (correct) conclusion that the government is going to have to raise taxes τ_2 in order to pay for the debts that it will incur in period 1. They realize that their after-tax income in period 2 is going to be lower and therefore want to save now in order to pay for those future taxes. Therefore they do not alter their consumption at all and just save all the extra after-tax income that they get from the lower τ_1.

Note that a lot of assumptions have to be satisfied for Ricardian equivalence to hold, and we might have good reason to doubt them. People have to be perfectly rational and understand the government's budget. The change in taxes cannot change their expectations of future government spending. The interest rate at which the government borrows and lends must be the same as the one faced by the households, and everyone must be able to borrow and lend at that rate. The taxes that the government is charging must be lump-sum so that households cannot change their tax obligations by changing their behavior. Relaxing any of those assumptions can lead to Ricardian equivalence not holding. You'll see an example of this in Exercise 6.5.

It's also important to remember what the Ricardian equivalence result, even when it holds, does *not* say, because people sometimes get this wrong. Ricardian equivalence does not say that anything the government does is irrelevant. It also doesn't say anything about what happens if the government changes G_1 or G_2. The only thing that is irrelevant is the timing of taxes, everything else held equal.

Precautionary Savings

So far we have been assuming that the household faces no uncertainty. In particular, it knows exactly how much income it's going to have in the future. How would the household's decisions change if it faced uncertainty? We are going to compare two households:

- Household A. Its period-1 income is y_1 and it knows that its period-2 income will be y_2.

- Household B. Its period-1 income is also y_1 but is it is uncertain about its period-2 income. It can be $y_2 + \epsilon$ (with 50% probability) or $y_2 - \epsilon$ (with 50% probability); ϵ is some positive number.

On average (across the possible "states of the world") both households make the same lifetime income. Does this mean they are going to make the same consumption choices? Let's see.

We have already solved household A's problem, so let's look at household B. Equation (6.2.2) still applies: if the household consumes c_1 its savings will be $a = y_1 - c_1$. This level of savings will result in two possible levels of period-2 consumption, depending on whether it ends up having high or low income. Its consumption will be:

$$c_2^H = y_2 + \epsilon + (1 + r)(y_1 - c_1)$$

if it has high income or:

$$c_2^L = y_2 - \epsilon + (1 + r)(y_1 - c_1)$$

if it has low income. The household will therefore solve:

$$\max_{c_1} u(c_1) + \beta\left[\frac{1}{2}u(y_2 + \epsilon + (1+r)(y_1 - c_1)) + \frac{1}{2}u(y_2 - \epsilon + (1+r)(y_1 - c_1))\right]$$

As we did in Chapter 2, we are assuming that the way the household deals with uncertainty is by maximizing average (or "expected") utility.

The first order condition for this problem is

$$u'(c_1) - \beta(1+r)\left[\frac{1}{2}u'(y_2 + \epsilon + (1+r)(y_1 - c_1)) + \frac{1}{2}u'(y_2 - \epsilon + (1+r)(y_1 - c_1))\right] = 0$$

$$\Rightarrow u'(c_1) = \beta(1+r)\left[\frac{1}{2}u'\left(c_2^H\right) + \frac{1}{2}u'\left(c_2^L\right)\right] \tag{6.2.17}$$

Equation (6.2.17) is the generalization of equation (6.2.8) to the case where the household faces uncertainty. The difference comes from the fact that the household doesn't know what its marginal utility of consumption in period 2 is going to be. Therefore it chooses savings on the basis of the *expected* marginal utility of period-2 consumption. Will this uncertainty make the household save more or save less? Let's look at this mathematically first and then try to make sense of what it means.

Proposition 6.1. If $u'(c)$ is a strictly convex function, then household B saves more than household A.

Proof. Assume the contrary, i.e. $c_1^B \geq c_1^A$. Then:

$$\begin{aligned}
u'\left(c_1^B\right) &\leq u'\left(c_1^A\right) && (u'(c) \text{ is a decreasing function})\\
\beta(1+r)\left[\frac{1}{2}u'\left(c_2^H\right) + \frac{1}{2}u'\left(c_2^L\right)\right] &\leq \beta(1+r)\,u'\left(c_2^A\right) && (\text{using (6.2.8) and (6.2.17)})\\
\left[\frac{1}{2}u'\left(c_2^H\right) + \frac{1}{2}u'\left(c_2^L\right)\right] &\leq u'\left(c_2^A\right) && (\text{simplifying})\\
\frac{1}{2}c_2^H + \frac{1}{2}c_2^L &> c_2^A && (u'(c) \text{ is decreasing and strictly convex})\\
y_2 + (1+r)(y_1 - c_1^B) &> y_2 + (1+r)(y_1 - c_1^A) && (\text{using the budget constraints})\\
c_1^B &< c_1^A && (\text{rearranging})
\end{aligned}$$

which is a contradiction. □

Figure 6.2.5 shows the reasoning graphically. If the household had no uncertainty, it would consume $\mathbb{E}(c)$ in period 2 and have marginal utility $u'(\mathbb{E}(c))$. Introducing uncertainty means it may consume either c_H or c_L. If $u'(c)$ is convex, as in the figure, then $\mathbb{E}[u'(c)] > u'(\mathbb{E}[c])$. Introducing uncertainty increases the expected period-2 marginal utility and makes saving more attractive.

The consumption-reducing, savings-increasing effect of future uncertainty on consumption behavior is known as "precautionary savings". If the future is uncertain, households reduce their consumption as a precaution. Proposition 6.1 says that if marginal utility is convex then households will have precautionary savings behavior. Do we have any reason to believe that $u'(c)$ is indeed convex? Probably the best reason to believe

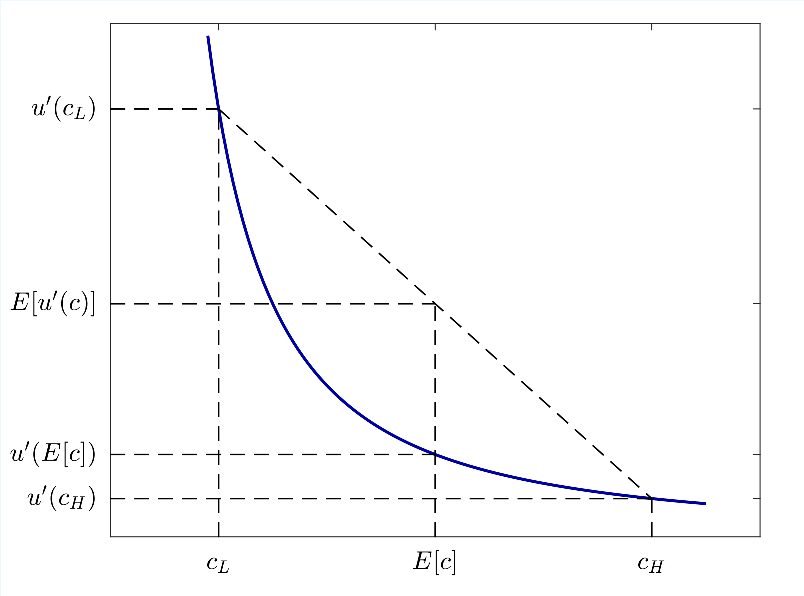


Fig. 6.2.5: *Convex marginal utility and precautionary savings.*

this comes from reasoning in the opposite direction. If we believe, as many economists do, that precautionary savings are an empirically important phenomenon, then a utility function with convex marginal utility is probably the right way to represent preferences. For what it's worth, the commonly-used CRRA function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ satisfies this, since:

$$\begin{aligned} u'(c) &= c^{-\sigma} \\ u''(c) &= -\sigma c^{-(1+\sigma)} \\ u'''(c) &= \sigma(1+\sigma)c^{-(2+\sigma)} > 0 \end{aligned}$$

and $u'''(c) > 0$ means $u'(c)$ is convex.

6.3 Extension to Many Periods

For many questions, the simplification of only considering two periods is good enough. For others, explicitly taking into account that there are more periods can be useful. We'll see examples of this later on. For now, we'll just look at how to analyze mathematically a many-period household savings problem. This turns out to be very similar to analyzing a two-period problem.

The household's preferences are:

$$\sum_{t=0}^{T} \beta^t u(c_t) \tag{6.3.1}$$

This extends the idea of (6.2.1) to T periods. The consumption of each of the future periods affects the household utility, but since $\beta < 1$, future periods matter less the further away they are.

One common assumption is to set $T = \infty$. The usual justification for this assumption is that, even though people die, they take into account their effect of their decisions on the money they leave to their children, on their children's children, etc.

Let's now think about the household's budget. Let a_t denote the level of savings that the household has at the beginning of period t. At the beginning of the following period, the household will have savings of:

$$a_{t+1} = (1+r)\,a_t + y_t - c_t \tag{6.3.2}$$

The idea behind (6.3.2) is to keep track of everything that adds or subtracts from the household's savings. The household increases savings by earning interest and by getting income and reduces them by consuming. Note that household faces an infinite number of constraints like (6.3.2), each linking the savings in two consecutive periods.

How do we analyze a maximization problem with an infinite number of constraints? There is more than one way. Here what we'll do is collapse them all into one by substituting one constraint into the next over and over again. Start from the period-0 and period-1 constraints:

$$a_1 = y_0 - c_0 + (1+r)\,a_0 \tag{6.3.3}$$

$$a_2 = y_1 - c_1 + (1+r)\,a_1 \tag{6.3.4}$$

Replace a_1 from (6.3.3) into (6.3.4)

$$a_2 = y_1 - c_1 + (1+r)(y_0 - c_0) + (1+r)^2\,a_0$$

Similarly, the period-2 constraint can be written as:

$$\begin{aligned} a_3 &= y_2 - c_2 + (1+r)\,a_2 \\ &= y_2 - c_2 + (1+r)(y_1 - c_1) + (1+r)^2 (y_0 - c_0) + (1+r)^3\,a_0 \end{aligned}$$

Generalizing, we can write the period-T constraint as:

$$a_{T+1} = \sum_{t=0}^{T} y_t (1+r)^{T-t} - \sum_{t=0}^{T} c_t (1+r)^{T-t} + (1+r)^{T+1}\,a_0$$

or, rearranging:

$$\frac{a_{T+1}}{(1+r)^T} + \sum_{t=0}^{T} \frac{c_t}{(1+r)^t} = \sum_{t=0}^{T} \frac{y_t}{(1+r)^t} + (1+r)\,a_0 \tag{6.3.5}$$

Equation (6.3.5) is conceptually very similar to equation (6.2.4). On the left we have the term $\sum_{t=0}^{T} \frac{c_t}{(1+r)^t}$, which is the present value of all the consumption over T periods and on the right we have the term $\sum_{t=0}^{T} \frac{y_t}{(1+r)^t}$, the present value of income over T periods. In equation (6.2.4) we had those same terms for the special case of $T = 2$.

There is an extra $(1+r)\,a_0$ term on the right. This is the value (including interest) of any savings that the household was born with. In the two-period example we were implicitly assuming that $a_0 = 0$ so this term didn't appear.

Also, there is an extra $\frac{a_{T+1}}{(1+r)^T}$ term on the left. This is the present value of any savings that the household has after the final period. If the household plans to have $a_{T+1} > 0$ then all the savings it has left over after period T cut into what it can afford to consume. Conversely, if the household chooses $a_{T+1} < 0$ this means that the household is planning to leave debts behind after period T. If the household could choose any value of a_{T+1} it wanted, then the best plan is clear: choose $a_{T+1} = -\infty$, i.e. leave infinite debts behind in order to be able to afford infinite consumption. Clearly this is not a reasonable model. A standard assumption, which we implicitly made in the two-period model, is that the household must pay back any debts by period T, i.e. $a_{T+1} \geq 0$. Imposing this leads to constraint:

$$\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} \leq \sum_{t=0}^{T} \frac{y_t}{(1+r)^t} + (1+r)\, a_0 \tag{6.3.6}$$

How about when $T = \infty$? In this case there is no last period where we can say: you need to pay your debts by this date. On the other hand, there should be some limit over how much debt you can accumulate: otherwise the household can afford vast amounts of consumption by simply running up ever-greater debts. A standard way to model the limits on the household's debts is to impose:

$$\lim_{T\to\infty} \frac{a_T}{(1+r)^T} \geq 0 \tag{6.3.7}$$

Constraint (6.3.7) allows the household to have large amounts of debt, as long as those debts don't grow to infinity too fast over time. It's sometimes known as a no-Ponzi condition, i.e. it says that household cannot run a Ponzi scheme.[10] If we impose the no-Ponzi condition, then (6.3.5) reduces to:

$$\sum_{t=0}^{\infty} \frac{c_t}{(1+r)^t} \leq \sum_{t=0}^{\infty} \frac{y_t}{(1+r)^t} + (1+r)\, a_0 \tag{6.3.8}$$

which is just like equation (6.3.6) with $T = \infty$.

Putting everything together, we can write the household's maximization problem as:

$$\sum_{t=0}^{T} \beta^t u(c_t)$$

$$s.t$$

$$\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} \leq \sum_{t=0}^{T} \frac{y_t}{(1+r)^t} + (1+r)\, a_0$$

[10] Ponzi schemes are names after Charles Ponzi. Ponzi was an Italian immigrant in Boston in the 1920s. He offered to pay 50% interest in 45 days and attracted a lot of money from investors. He had no way to deliver those huge interest rates but as long as more investors kept coming in sufficiently fast he was able to pay old investors with the money he got from new ones. The scheme collapsed spectacularly after a few months.

Now let's set up a Lagrangian:

$$L(c_0, c_1, \dots) = \sum_{t=0}^{T} \beta^t u(c_t) - \lambda \left[\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} - \sum_{t=0}^{T} \frac{y_t}{(1+r)^t} - (1+r) a_0 \right]$$

and take first order conditions with respect to c_t for some generic periods t and $t+1$:

$$\beta^t u'(c_t) - \lambda \frac{1}{(1+r)^t} = 0$$

$$\beta^{t+1} u'(c_{t+1}) - \lambda \frac{1}{(1+r)^{t+1}} = 0$$

Solving for λ and rearranging we get:

$$u'(c_t) = \beta(1+r) u'(c_{t+1}) \tag{6.3.9}$$

Equation (6.3.9) is the Euler equation (6.2.8) again. Now it describes the tradeoff for consuming in any two consecutive periods instead of just periods 1 and 2, but the interpretation is the same as before. In a sense, going through the many-period case doesn't tell us all that much that we didn't know already from the two-period case. We'll see some of the uses of the infinite-period model later on.

6.4 Behavioral Theories

In everything we've done so far the assumption has been that consumption-savings decisions result from a rational calculation. If you have ever met actual people, you might have doubts about this assumption. The challenge for macroeconomic theory is that there is one way to behave rationally and many, many ways to behave not-quite-rationally. Which of the many possible departures from perfect rationality is important enough to take into account when we think about the macroeconomy? This is very much an open question.

One of the best-known pieces of evidence showing that something other than full rationality is at play comes from 401(k) default options. 401(k) plans (named after the section of the tax code that governs them) are tax-advantaged investment plans that are offered by some firms to their employees. The employee contributes a fraction of their salary to their individual account to be withdrawn upon retirement. In the meantime, all the returns on investment earned within the account are not taxed. Typically, employees can choose how much of their salary to contribute to their account. If they don't make a decision of how much they want to contribute, then their contribution is set to some default option. In the world of perfect rationality, employees would calculate the optimal level of 401(k) savings based on preferences, interest rates, etc. and set their contribution accordingly. The default option should have no effect on how much they contribute. In practice, researchers have found that default options tend to have very large effects on what people end up doing. Choi et al. (2004) studied three companies that switched the default option from contributing zero to contributing between 2% and 3% of the employees' salary. As a result of this, the percentage of employees who saved in 401(k) funds rose by more than 40 percentage points, even though their actual options had not changed at all.

There are several types of not-quite rational models of consumption behavior. One type of model is based on the idea that people don't pay attention to all the relevant factors all the time: they just follow some approximate rule that works sort-of-OK for them, such as "keep 3 months of salary in the bank account", "save \$200 per month" or "neither save nor borrow".[11] Under this sort of model, our calculations of how households' consumption react to interest rates might be all wrong: it's possible that households don't pay attention to the interest rate at all.

Another class of models is based on the idea that people have poor self-control: whenever they encounter a good that they like, they buy it (as long as they can pay for it either using their savings or by borrowing). Under the extreme version of this model, nothing about the future affects consumption decisions because people are just not thinking about the future.[12] A variant of this idea says that people have poor self-control but in their calmer moments they understand this and try to arrange things to avoid falling into temptation, for instance by limiting their access to their own savings.

In the rest of this book we are going to stick with the simple rational model, but it's useful to keep in the back of our mind that consumption behavior could depart from rationality in all sorts of interesting ways.

Exercises

6.1 **Two-Period Problem with Taxes and Initial Wealth**

Suppose a household solves the following two-period consumption-savings problem with taxes:

$$\max_{c_1,a,c_2} u(c_1) + \beta u(c_2)$$

$$s.t.$$

$$a = a_0 + y_1 - \tau_1 - c_1$$

$$c_2 = y_2 - \tau_2 + (1+r)\,a$$

with $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, where: c_1 is consumption at time 1, c_2 is consumption at time 2, y_1 is household income at time 1, y_2 is household income at time 2, τ_1 are taxes at time 1, τ_2 are taxes at time 2, and a_0 is initial wealth.

(a) Solve for the household's choice of c_1, c_2 and a in closed form.[13]

(b) How does $\frac{c_1}{y_1}$ depend on y_2? What would happen if households suddenly became optimistic about the future?

(c) How does $\frac{c_1}{y_1}$ depend on a_0? Interpret.

(d) How does $\frac{c_1}{y_1}$ depend on β? Interpret.

[11] One variant of this idea, known as "rational inattention", imagines that paying attention is costly and households rationally choose what things they are going to pay attention to and which things they are going to ignore.

[12] There is also a rational version of this model. It's not that people have poor self-control, it's that they have a very low β so they strongly prefer to enjoy the present at the expense of the future. In this version, it may be rational to consume as much as possible in the present, even in the full understanding that it will mean consuming less in the future.

[13] "In closed form" means that you have an explicit expression for something. Suppose I ask you to determine some variable x. Then something like $x = y^2 - z$ is a solution "in closed form"; something like $e^x + x - b = 0$ also implicitly tells what x should be but is not in closed form.

(e) Suppose $y_2 = \tau_1 = \tau_2 = 0$ and compute $\frac{\partial c_1}{\partial r}$. How does the answer depend on σ? Interpret the answer. [Hint: this is a hard question, not the maths but the interpretation. Think about what σ means for the relative importance of income and substitution effects]

6.2 Savings Rates for Different People

An economist has data on the occupation, this year's income (denoted y) and this year's consumption (denoted c) of a sample of 26-year-olds. Within this sample, some are top professional athletes and others are medical doctors in their first year of residence. [Make whatever assumptions you think are reasonable and, if you wish, refer back to Exercise 6.1]

(a) Suppose one computed $s = \frac{y-c}{y}$ for each individual in the sample. Should we expect s to be higher for athletes or doctors?

(b) Suppose interest rates go down. How should we expect the response of s to differ between the two groups?

6.3 Consumption and Interest Rates

A household solves a special case of the problem in Exercise 6.1, with $\tau_1 = \tau_2 = a_0 = 0$. Suppose that, by coincidence, the values of β, r, y_1 and y_2 are such that it is optimal for the household to consume:

$$c_1 = y_1$$
$$c_2 = y_2$$

(a) What will happen to c_1 if interest rates increase? [A graph will be helpful. Make sure you draw it carefully, it's probably useful to make it large]

(b) Does the overall utility achieved by the household increase, decrease or stay the same?

(c) Suppose this household was the only household in the economy and Jones & Klenow, using this year's data, applied their measure of welfare to this economy. How would their measure of welfare change with the change in interest rates?

(d) Explain the relationship between the answer to part (b) and the answer to part (c).

6.4 Consumption and Income

Suppose there are two households in the economy. Each of them solves a special case of Exercise 6.1, where $\tau_1 = \tau_2 = a_0 = 0$, $\beta = 1$ and $r = 0$.

(a) Solve for c_1 as a function of y_1 and y_2.

(b) Suppose the incomes of each of the households are:

	y_1	y_2
Household A	2	4
Household B	6	4

Compute c_1 and c_2 for each of the households.

(c) Suppose an economist is trying to decide what is a reasonable model for consumption behavior and only has data for period 1. Is the data supportive of the Keynesian view of the consumption function? Explain.

(d) Suppose the same economist now looks at data for period 2 in addition to data for period 1. Is the data supportive of the Keynesian view of the consumption function? Explain.

6.5 **Credit Constraints and Ricardian Equivalence**

Suppose a household solves the following variant of the problem in Exercise 6.1:

$$\max_{c_1,a,c_2} u(c_1) + \beta u(c_2)$$
$$s.t.$$
$$a = y_1 - \tau_1 - c_1$$
$$c_2 = y_2 - \tau_2 + (1+r)\,a$$
$$a \geq -b \tag{6.4.1}$$

(a) What does equation (6.4.1) mean? What does b represent?

(b) Plot the household's budget constraint. In the same graph, plot constraint (6.4.1).

(c) Solve for c_1, c_2 and a.

[Hint: notice that the constraint (6.4.1) is a weak inequality, not an equality, so it may or may not be binding. If it is not binding, then you can use the answer from Exercise 6.1. Then think about what happens if it is indeed binding. Then figure out whether or not it will be binding]

(d) Show that, other things being equal, constraint (6.4.1) is more likely to be binding if

i. $y_2 - \tau_2$ is high,

ii. $y_1 - \tau_1$ is low,

iii. b is low.

Interpret each of these conditions.

(e) Suppose that the government announces a "stimulus package" of size Δ. This involves lowering τ_1 by Δ and increasing τ_2 by $\Delta(1+r)$ so that the present value of taxes is unchanged. How does c_1 respond to the stimulus package if we start from a situation where constraint (6.4.1) is NOT binding? How does c_1 respond to the stimulus package if we start from a situation where constraint (6.4.1) is binding? Explain.

(f) Suppose that instead of announcing a stimulus package, the government announces that it will allow households to borrow Δ from the government and repay it back (with interest) at $t = 2$. How do the effects of this policy compare with the effects of the stimulus package? Explain.

6.6 **A Tax on Savings**

A household solves a special case of the problem from Exercise 6.1, with $\tau_1 = \tau_2 = a_0 = 0$. Suppose now that the government introduces a tax on interest income, so that a household that saves a (and therefore earns interest ra) will have to pay τra in taxes. (If the household borrows instead of saving it pays no tax).

(a) Plot the new budget constraint.

(b) Show graphically:

i. an example where the new policy persuades households to save more,

ii. an example where the new policy persuades the household to save less,

iii. an example where the household does not change its decision in response to the new policy.

Explain.

6.7 One Rational Household

Consider an economy that is well described by the Solow growth model. The production function is:

$$Y = K^{\alpha} L^{1-\alpha}$$

The population is constant an equal to 1 and there is no technological progress. The saving rate is s and the depreciation rate is δ.

(a) What will the capital stock be in the long run?

(b) What will be the real interest rate in the long run?

For the rest of the question, assume that $s = 0.4$, $\alpha = 0.35$ and $\delta = 0.1$ and the economy is initially at a steady state

(c) If the savings rate increases from $s = 0.4$ to $s = 0.5$:

i. Will GDP increase in the long run?

ii. Will consumption increase in the long run?

(d) Suppose a single household in the economy (the Friedmans) decides that, instead of just saving an exogenous fraction s of their income, they are going to start choosing consumption and saving to maximize the following standard preferences:

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

The Friedmans can borrow or lend at the market interest rate. Since it's just the Friedmans who start acting this way, and they are small relative to the economy, we are going to assume that the aggregate economy (aggregate quantities, prices, etc.) remains unchanged. Will the Friedmans' consumption be high initially and then fall over time, be low initially and then rise over time, or remain constant?

6.8 The Marginal Propensity to Consume

Suppose a household has preferences given by (6.3.1) over consumption across infinite periods. Its discount factor and the real interest rate satisfy:

$$1 + r = \frac{1}{\beta}$$

The household starts with no wealth, will receive income y_0 in period $t = 0$ and income $\bar{y}$ in every subsequent period, from $t = 1$ onwards.

(a) Use the Euler equation (6.3.9) and the budget constraint (6.3.8) to find an expression for initial-period consumption c_0.

(b) Compute $\frac{\partial c_0}{\partial y_0}$

(c) Suppose $r = 0.04$. What is the marginal propensity to consume out of a purely temporary increase in income? Describe in words what a household does with a temporary increase in income.

6.9 The Marginal Propensity to Consume in Proportional Terms

A household solves a special case of the problem in Exercise 6.1, with $\tau_1 = \tau_2 = a_0 = 0$ and $y_2 = my_1$, where m is some number.

(a) Suppose that both y_1 and y_2 increase by $x\%$. By what percentage do c_1 and c_2 increase?

(b) Suppose y_1 increases by $x\%$ but y_2 remains unchanged. By what percentage does c_1 increase? How does this depend on m and r? Explain.

CHAPTER 7

Labor and Leisure

When we looked at the Solow model we assumed that everyone in the population was working. In this chapter we are going to think about the labor market in a little bit more detail. We'll start by looking at some of the ways we measure what's going on in the labor market. Then we'll think about the incentives that govern the decision of how much to work. At first, we'll maintain the assumption that the labor market is perfectly competitive, so a worker can supply as much labor as they want at the equilibrium wage w. Finally we'll think about non-perfectly-competitive models of the labor market.

7.1 Measuring the Labor Market

The usual way to measure the state of the labor market is by surveys of both individuals and employers. In the US, the main survey of individuals is conducted by the Bureau of Labor Statistics (BLS) and is called the Current Population Survey (CPS). Other countries conduct similar surveys although the exact questions they ask vary slightly from one country to another.

Individuals are classified into three main groups:

- **Employed**: if they have worked (including as employees or self-employed) in the past week.
- **Unemployed**: if they did not work during the past week but actively looked for a job.
- **Out of the labor force**: if they did not work and did not look for a job in the past week.

Based on this classification we define:

$$\begin{aligned}
\text{Labor force} &\equiv \text{Employed} + \text{Unemployed} \\
\text{Participation rate} &\equiv \tfrac{\text{Labor force}}{\text{Population}} \\
\text{Employment rate} &\equiv \tfrac{\text{Employed}}{\text{Population}} \\
\text{Unemployment rate} &\equiv \tfrac{\text{Unemployed}}{\text{Labor force}}
\end{aligned}$$

Figure 7.1.1 shows the movement over time in these quantities. The labor force participation rate increased for most of the second half of the twentieth century, peaking around 67% in 2000, and has since fallen to about

63%. It moves more smoothly than the employment rate, which has higher-frequency fluctuations, which correspond to movements in unemployment. The unemployment rate is quite volatile, moving up and down between about 3% and 10%.

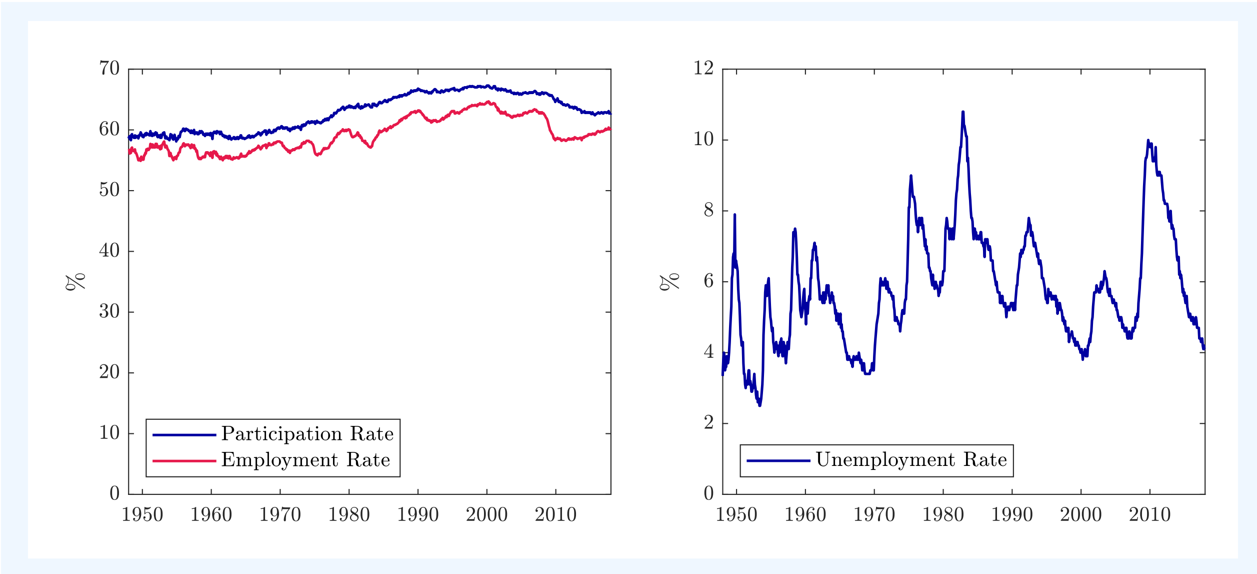


Fig. 7.1.1: *Labor market indicators in the United States. Source: CPS.*

Flows Across Employment Status

The CPS also keeps track of how people shift between employment, unemployment and out of the labor force. At any point in time, there are large numbers of people who change status in every direction. Figure 7.1.2 shows the magnitude of these flows for the month of October, 2018.

Using the data on stocks and flows we can compute the rates at which people transition from one status to another. The (monthly) job finding rate is defined as the number of workers who shift from unemployment to employment, expressed as a fraction of the pool of unemployed workers. The (monthly) job loss rate is defined as the number of workers who shift from employment to unemployment, expressed as a fraction of the pool of employed workers. Figure 7.1.3 shows the evolution over time of these rates. The job loss rate oscillates between 1% and 2% per month, while the job finding rate oscillates around 30% per month.[1]

Vacancies and the Beveridge Curve

Across the market from workers looking for jobs are firms looking for workers. Starting in 2001, the BLS has conducted a survey called the Job Openings and Labor Turnover Survey (JOLTS) which asks firms, among other things, how many job openings (sometimes called "vacancies") they currently have. For earlier periods,

[1] Both of these rates are somewhat underestimated. If a worker switches status back and forth within the same month, the monthly survey will not detect this and will record no transition. This is especially important for the job finding rate since the denominator is smaller.

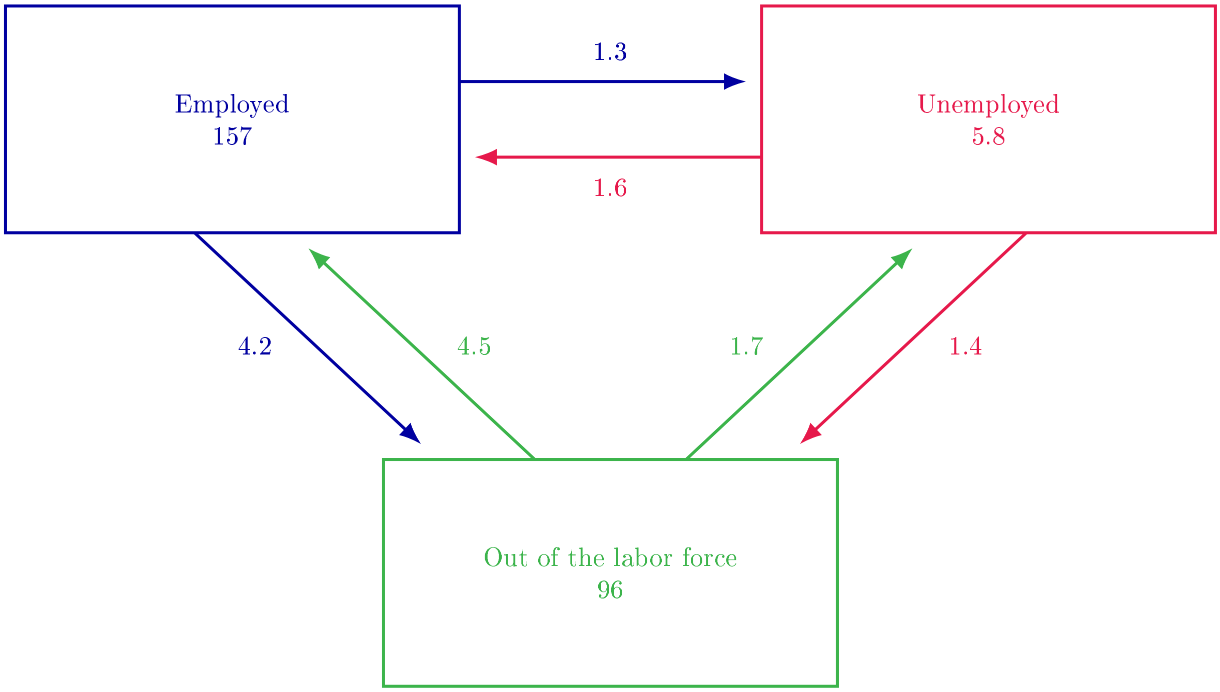


Fig. 7.1.2: *Stocks and flows of workers across labor market status in October, 2018. Figures in millions of workers. Source: CPS.*

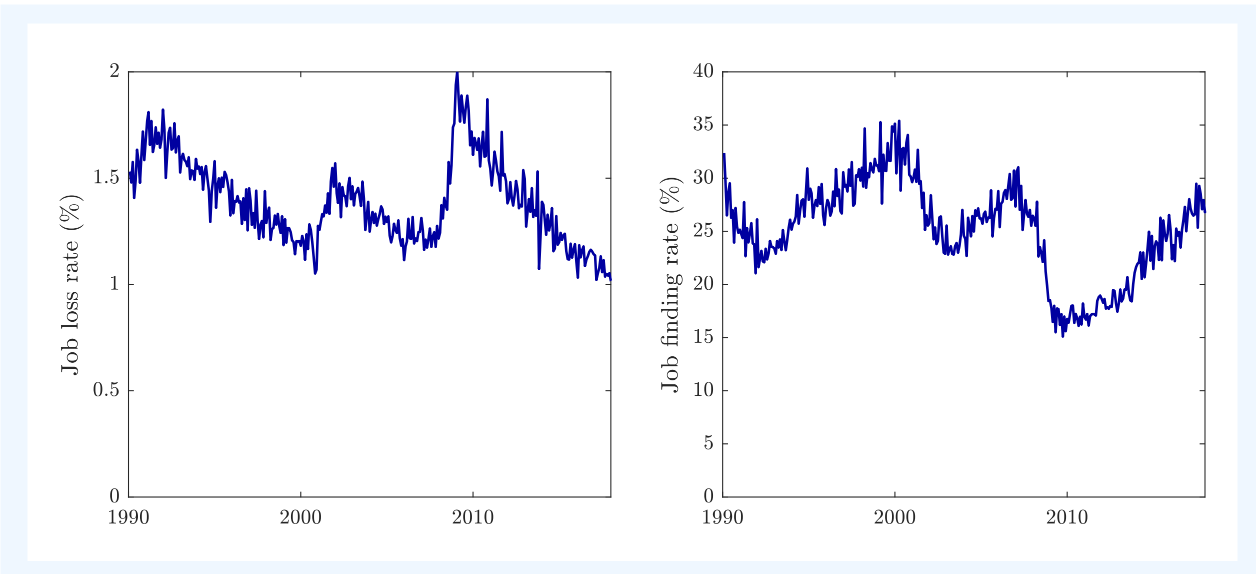


Fig. 7.1.3: *Monthly job loss rate and job finding rate. Source: CPS.*

there are measures of job openings based on sources like help wanted ads in newspapers. The vacancy rate is defined as the ratio of vacancies to the total labor force.

Figure 7.1.4 shows the relationship between the vacancy rate and the unemployment rate in the US economy. There is a strong negative relationship. High vacancy rates have tended to coincide with low unemployment rates. This negative relationship between vacancies and unemployment is known as the Beveridge Curve. When the ratio of vacancies to unemployed workers is high, the labor market is said to be "tight".

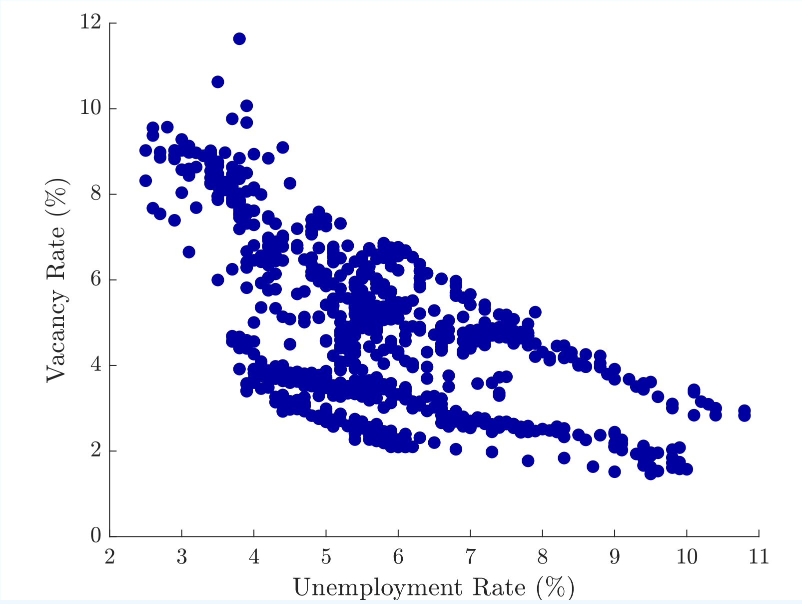


Fig. 7.1.4: *The US Beveridge Curve, 1948-2018. Each dot is one month. Sources: Unemployment from CPS. Vacancies from NBER Macrohistory Database, Barnichon (2010) and JOLTS.*

What do the Measures Tell Us?

Let's start with the most widely reported statistic: the unemployment rate. A high unemployment rate is typically viewed as a problem while a low unemployment rate is viewed as a success, and with good reason. By definition, people who are unemployed would like to be employed but have not been able to achieve this. However, just looking at the unemployment rate does not give a full account of what is going on in the labor market.

First, searching for a job is a productive use of somebody's time. We often, including in this book, treat all workers and all jobs as being identical, but it's obvious that this is not literally true. Finding a job requires search effort because workers are trying to find jobs that suit them and employers are trying to find workers that suit them. Looking at help wanted ads, writing resumes, contacting potential employers, etc., are part of the process of getting the right person into the right job. Unemployment is partly a reflection of the fact that this whole process is time-consuming.

On the other hand, people are counted as unemployed only if they took active steps to try to find a job. There are plenty of people who would like a job but have not taken active steps within the past week to find one. We can see evidence for this directly from Figure 7.1.2: there is a large flow of people from "out of the labor force" into "employed" every month: these are workers who were not actively looking for a job but nevertheless found one and took it. One reason why people who want a job might not be looking for one is that they might believe that it's very unlikely that they will find one. These are sometimes known as "discouraged workers." If a large fraction of the people counted as "out of the labor force" are discouraged workers, then a low unemployment rate need not mean that the outcomes in the labor market are good.

An alternative approach is to avoid making distinctions between people who are actively looking for work and those who are not. Notice that the denominator in the employment rate and the unemployment rate is different. The employment rate looks at how many people are employed as a fraction of the population rather that as a fraction of those in the labor force. This measure treats those who don't work by choice, discouraged workers, and the unemployed in the same way. If unemployed workers become discouraged and leave the labor force, then the unemployment rate goes down but the employment rate is unchanged.

Is a high employment rate the best indicator of good outcomes in the labor market? Not necessarily. There are many reasons why some people choose not to work: they retire, they take care of their families, they study full time, etc. A low employment rate could be a symptom of changes in the extent to which people are choosing these alternative uses of their time and not necessarily a problem with the functioning of the labor market or the overall economy.

Despite their limitations, these measures do tell us something useful about the economy. Looking at how these variables behave will be one way to assess various theories about how the economy works.

7.2 A Static Model of the Labor Market

We are going to imagine that a worker has the preferences

$$U(c,l) = u(c) + v(l) \tag{7.2.1}$$

where c stands for consumption and l stands for leisure. The function $u(c)$ describes how much the worker enjoys consumption and the function $v(l)$ describes how much the worker enjoys dedicating time to non-market activities (we call these "leisure" but they could include non-market production such as doing laundry). We are going to imagine that both $u(c)$ and $v(l)$ are concave functions. This means that the household experiences diminishing marginal utility of both leisure and consumption.

The worker has a total of one unit of time, so the amount of time he spends working is given by

$$L = 1 - l$$

You'll sometimes see preferences over consumption and leisure expressed in terms of *dis*utility from working rather utility from leisure, with a function of the form:

$$U(c,L) = u(c) - z(L)$$

Setting $z(L) = -v(1-L)$ makes the two formulations exactly equivalent. We'll stick to expression (7.2.1).

The worker has to decide how much of his time to dedicate to market work and how much to dedicate to leisure. One way to interpret this decision is literally: imagine that the worker has a job that allows him to choose how many hours to work (for instance, the worker is an Uber driver) and think about how the worker makes this choice. More broadly, there are many decisions that involve trading off higher income against less leisure: choosing between a full-time job and a part-time job; choosing between a high-stress, highly paid job and a lower-paid, more relaxed job; choosing at what age to retire; choosing how many members of a many-person household will be working in the market sector, etc. We can think about the choice of "leisure" as summarizing all of these decisions.

The worker gets paid a wage w per unit of time. Below we'll think about where this wage level comes from but for now we are just thinking about the worker's decision problem, which takes the wage as given. The total amount the worker can spend on consumption goods is given by the budget:

$$c \leq w(1-l)$$

The worker solves the following problem:

$$\begin{aligned} &\max_{c,l} u(c) + v(l) \\ &s.t. \\ &c \leq w(1-l) \end{aligned} \tag{7.2.2}$$

This is a two-good consumption problem. The two goods here are time and consumption goods. The only thing to keep in mind is that the household is initially endowed with one unit of time, and it has to choose how much of it to sell in order to buy consumption.

Figure 7.2.1 shows the solution to problem (7.2.2). The worker will choose the highest indifference curve he can afford, which implies that he will pick a point where the indifference curve is tangent to the budget constraint. For any wage, the budget constraint always goes through the point $(1,0)$: the worker can always choose to enjoy his entire endowment of time in the form of leisure and consume zero. The slope of the budget constraint is $-w$: w is the relative price of time in terms of consumption goods. When w is high, time is expensive relative to goods, so the budget constraint becomes steeper.

We can also find the solution to problem (7.2.2) through its first order conditions. The Lagrangian is:[2]

$$L(c,l,\lambda) = u(c) + v(l) - \lambda[c - w(1-l)]$$

The first order conditions are:

$$\begin{aligned} u'(c) - \lambda &= 0 \\ v'(l) - \lambda w &= 0 \\ \Rightarrow \frac{v'(l)}{u'(c)} &= w \end{aligned} \tag{7.2.3}$$

[2]This problem is sufficiently simple that we don't need to use a Lagrangian to solve it. We could just as easily replace $c = w(1-l)$ into the objective function and solve:

$$\max_{l} u(w(1-l)) + v(l)$$

which gives us the same solution.

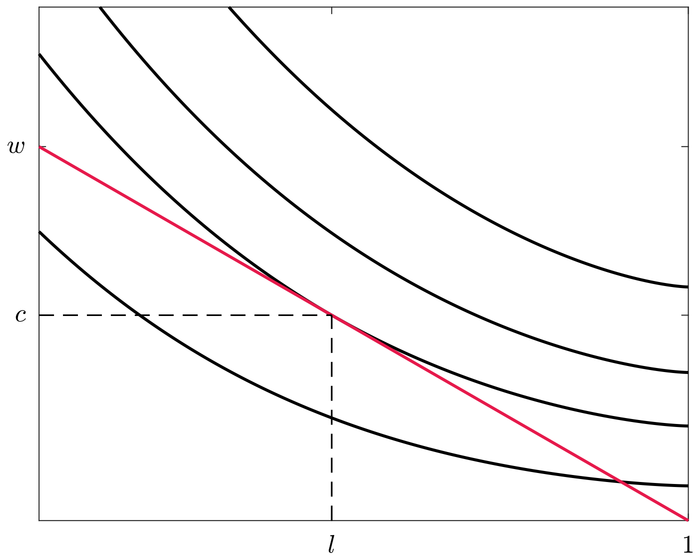


Fig. 7.2.1: *The consumption-leisure decision as a two-good consumption problem.*

Equation (7.2.3) describes how the worker trades off dedicating time to market work or to leisure. If the worker allocates a marginal unit of time to leisure, he simply enjoys the marginal utility of leisure $v'(l)$. If instead the worker spends that time at work, he earns w, so he is able to increase his consumption by w; this gives him w times the marginal utility of consumption $u'(c)$. At the margin, the worker must be indifferent between allocating the last (infinitesimal) unit of time between these two alternatives, so (7.2.3) must hold. (7.2.3) is also an algebraic representation of the tangency condition shown in Figure 7.2.1. The slope of the indifference curve is given by the marginal rate of substitution between leisure and consumption: $\frac{v'(l)}{u'(c)}$. The slope of the budget constraint is w, so (7.2.3) says that the two are equated.

The Effect of Wage Changes

Let's imagine the wage w changes. How does the worker change his choice of leisure and consumption? The answer to this question is going to play an important role in some of the models of the entire economy that we'll analyze later. For now, we are going to study the question in isolation, just looking at the response of an individual worker to an exogenous change in the wage. For concreteness, let's imagine that the wage rises.

Let's take a first look at this question graphically. A change in wages can be represented by a change in the budget constraint, as in Figure 7.2.2. The new budget constraint still crosses the point $(1, 0)$, but the slope of the budget constraint is steeper. As with any change in prices, this can have both income and substitution effects.

The substitution effect is straightforward: a higher wage means that time is more expensive. Other things being equal, this would make the worker substitute away from leisure (which has become relatively expensive) towards consumption (which has become relatively cheap). This makes the worker work more.

In addition, the higher wage unambiguously helps the worker: the worker is selling his time so a higher

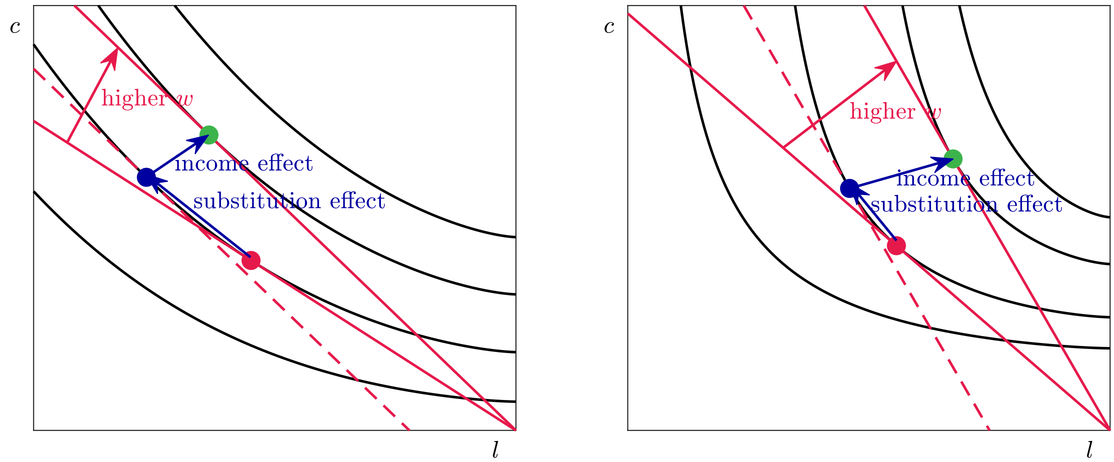


Fig. 7.2.2: *Consumption and leisure response to higher wages. Income and substitution effects.*

price is good for him. In other words, there is a positive income effect. For the *consumption* choice, the income effect reinforces the substitution effect since both push the worker to consume more. For the *leisure* choice, this goes in the opposite direction as the substitution effect: as the worker becomes richer, he wants more of everything, including leisure. In the example depicted on the left panel of Figure 7.2.2, the substitution effect dominates. The worker ends up at a point to the left of where he started, showing he has decided to work more (and get less leisure) when wages rise. The right panel shows an example where the income effect dominates so the worker decides to work less (enjoy more leisure) when the wage rises.

An Explicit Example

Suppose that the utility function takes the following form:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$
$$v(l) = -\frac{\theta\epsilon}{1+\epsilon}(1-l)^{\frac{1+\epsilon}{\epsilon}}$$

where θ and ϵ are parameters. For this case, we can get an explicit formula for how much consumption and leisure the worker will choose. Marginal utility of consumption and leisure are, respectively:

$$u'(c) = c^{-\sigma}$$
$$v'(l) = \theta(1-l)^{\frac{1}{\epsilon}}$$

so replacing in equation (7.2.3) we get

$$\frac{\theta (1-l)^{\frac{1}{\epsilon}}}{c^{-\sigma}} = w$$

Using the budget constraint to replace c we get:

$$\frac{\theta (1-l)^{\frac{1}{\epsilon}}}{(w(1-l))^{-\sigma}} = w$$

$$1-l = \theta^{-\frac{1}{\frac{1}{\epsilon}+\sigma}} w^{\frac{1-\sigma}{\frac{1}{\epsilon}+\sigma}} \tag{7.2.4}$$

Equation (7.2.4) gives us an explicit formula for how much the worker will choose to work depending on the wage and the parameters in the utility function. Will this worker work more or less when the wage is higher? Mathematically, this depends on whether the exponent on w is positive or negative. If $\sigma < 1$ then the exponent is positive and the worker will work more when the wage is higher, i.e. the substitution effect dominates; if $\sigma > 1$ then the income effect dominates and the worker works less when the wage is higher. As we've seen before, in this particular utility function, the parameter σ tells us about curvature. In economic terms, it tells us how fast the marginal utility of consumption goes down when consumption increases. Why is this the relevant aspect of preferences that governs income and substitution effects? If the marginal utility of consumption falls rapidly as consumption increases, the worker is unwilling to substitute towards even more consumption when wages rise and he chooses to enjoy more leisure instead.

Adding Taxes and Transfers

Let's try to figure out how the worker's decisions will respond to changes in tax policy. We'll represent tax policy in a highly simplified way, with just two numbers: τ and T. τ is the tax rate on the worker's income: the government collects a fraction τ of the worker's income as taxes. T is a transfer that the worker gets from the government. This is intended to represent the multiple income-support policies that many countries have in place: unemployment insurance, food assistance, pensions, public healthcare, public education, etc. The worker's budget constraint is:[3]

$$c \leq w(1-l)(1-\tau) + T$$

If we were looking at this from the government's perspective, we would have to think about how τ and T are linked: the government must set τ to collect enough revenue to afford T. For now, we'll look at this from the worker's perspective, taking τ and T as given.

Figure 7.2.3 shows how taxes and transfers affect the worker's budget constraint. There are two effects, one from the taxes and one from the transfers. The effect of transfers is to simply shift the budget constraint up: for any given level of leisure, the worker can consume T more. In particular, he can consume T without

[3] T and τ are taken as constants for simplicity, but this is less restrictive than you might think. For instance, suppose a housing program offers subsidized rents to low income households. The poorest households receive \$3,000 per year in subsidies but this benefit is phased out depending on the household's income until a household making \$30,000 a year receives no benefit at all. We could represent this policy as $T = 3,000$ and $\tau = 0.1$ because the loss of housing subsidies is, effectively, a tax on the household's labor earnings. What we miss by having constant values for T and τ is that a lot of policies are non-linear in complicated ways. Even the simple program described above has an implicit tax rate of 0.1 on incomes below \$30,0000 but no taxes on marginal income above \$30,000.

working at all. The effect of taxes is to lower the slope of the budget constraint: from the point of view of the worker, the price at which he can sell time to obtain consumption is the *after-tax* wage $w(1-\tau)$.

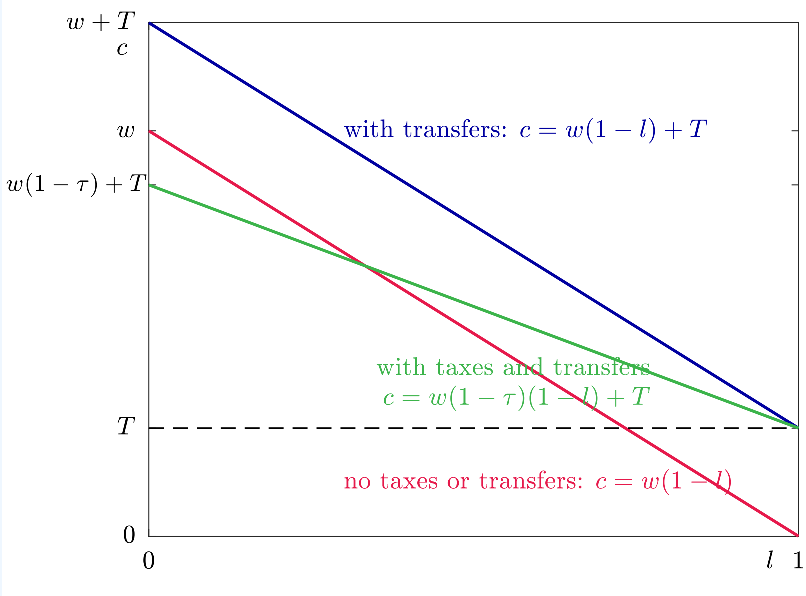


Fig. 7.2.3: *The worker's budget constraint when there are taxes and transfers.*

The effect of higher transfers is a pure income effect. Prices have not changed but the worker is richer, so he chooses higher consumption and higher leisure. This is illustrated in Figure 7.2.4.

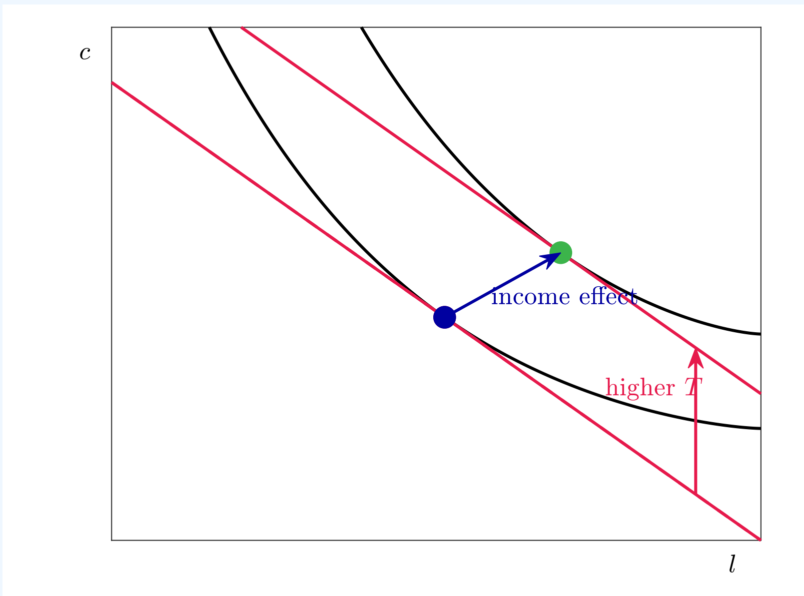


Fig. 7.2.4: *Consumption and leisure response to higher transfers.*

The effect of higher tax rates is just like the effect of lower wages. There is a substitution effect (the after-tax price of time has gone down, so the worker chooses higher leisure) and an income effect (the worker is poorer, so he chooses less leisure), and they push the leisure choice in opposite directions. The first-order condition for the household's problem becomes:

$$\frac{v'(l)}{u'(c)} = w(1-\tau) \tag{7.2.5}$$

The worker equates the marginal rate of substitution to the after-tax wage $w(1-\tau)$ rather than the full wage w.

7.3 Some Evidence on Labor and Leisure Decisions

Long-Run Evidence on Income and Substitution Effects

We have seen that in general higher wages could make workers choose to work more or less. Which way does it go in practice? International and historical data offers us the chance to see how workers' choices vary across settings with very different wages.

Ramey and Francis (2009) piece together data from time-use surveys for the US for the period 1900-2006 to try to determine whether the amount of time spent on leisure has gone up or down. We know that wages have increased a lot over the last 100 years (by a factor of 9 approximately). If the income effect dominates, we should see that over time people are choosing more leisure; if the substitution effect dominates, we should see that over time people are choosing less leisure. Figure 7.3.1 shows the trends in hours per week spent on leisure, broken down by age. Overall, there seems to be a very slight upward trend in leisure, especially from 14-17 year olds and over-65-year olds.[4] This suggests that income and substitution effects almost cancel each other out, and perhaps income effects dominate slightly.

Bick et al. (2018) do a similar measurement, but instead of looking at variation in wages and leisure over time they look at variation in wages (or more exactly, in GDP per capita) and leisure across countries at a given point in time. Figure 7.3.2 shows how hours worked correlate with GDP per capita across countries. There is a negative correlation, suggesting that overall the income effect tends to dominate and people work less as wages increase. However the relationship is not very strong, suggesting that income and substitution effects are not so far from canceling each other out.

Comparing the US and Europe

Figure 7.3.3 shows how hours worked per employed person in the US and some European countries have evolved over the last few decades.[5] Up until the early 1970s or so, the US and Western Europe looked very similar but then they started to diverge so that today Europeans work less than Americans. How come?

[4] The study separately measures hours spent on market work, nonmarket work, schoolwork and "pure" leisure. The figure shows trends in pure leisure time.

[5] The figure only looks at employed persons and says nothing about employment rates, but these tell the same story: the US and Europe look similar until the 1970s and then employment rates are higher for the US.

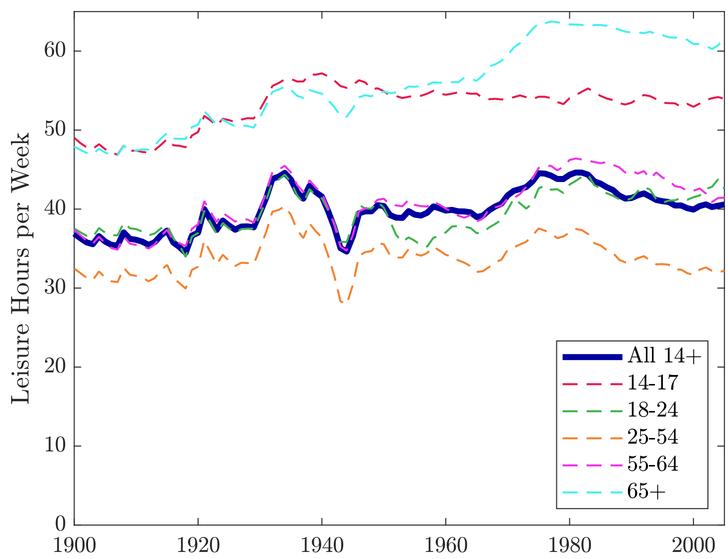


Fig. 7.3.1: Average hours of leisure per week for everyone ages 14+ in the US. Source: Ramey and Francis (2009).

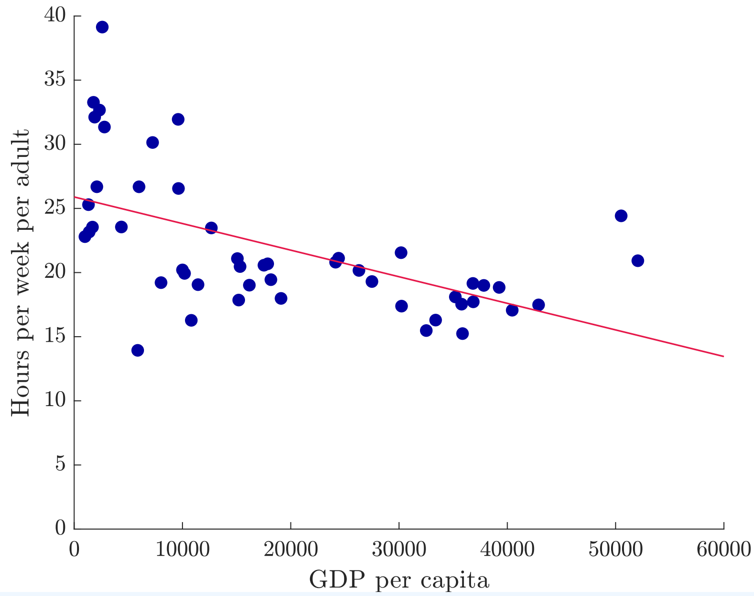


Fig. 7.3.2: Average hours of work per week across countries. Source: Bick et al. (2018).

One hypothesis, put forward by Prescott (2004), is that the reason is differences in tax and social security policy. Europe has higher tax rates and social spending than the US. These, the argument goes, discourage Europeans from working as hard as Americans. Prescott proposes a simple version of the model in this chapter and argues that it can explain the magnitude of the differences between US and European labor markets. Exercise 7.5 asks you to go through the details of Prescott's calculation and to think about the role

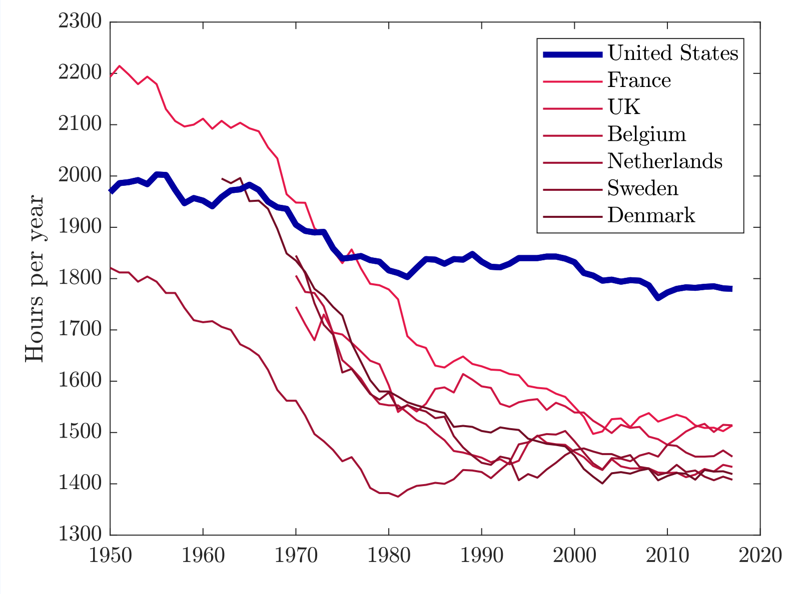


Fig. 7.3.3: *Hours worked per year per employed person in the US and selected European countries. Source: OECD.*

of income and substitution effects.

There is no consensus among economists about whether Prescott's hypothesis is correct. It has been criticized from a few different angles. One criticism focuses on the elasticity of labor supply. Exercise 7.5 asks you to compute the elasticity of labor supply that is implicit in Prescott's calculations. This matters because it governs *how much* labor supply responds to changes in incentives. Most microeconomic estimates of this elasticity are quite a bit lower than Prescott's value, though there is some debate as to how to translate microeconomic estimates into macroeconomic calculations. A second line of criticism focuses on timing. Policies in the US and Europe became different in the 1960s and 1970s but the differences in labor supply continued to widen well after that, suggesting something else was going on (or that policies take a very long time to have an effect).

Some other explanations for the US-Europe difference have been proposed. Blanchard (2004) argues that differences in preferences may be a large part of the reason: maybe Europeans place a higher value on leisure than Americans. Economists tend to be a little bit uncomfortable with explanations based on differences in preferences. We cannot observe preferences directly so these theories are very hard to test (but this does not necessarily mean they are wrong). Furthermore, we need to explain why Europeans work less than Americans now but this was not the case in the 1950s. Perhaps cultural differences only become relevant once society reaches a certain level of income.

Alesina et al. (2006) argue that a large part of the explanation may have to do with the different role of labor unions in the US and Europe. Unions tend to be stronger in Europe and union contracts tend to specify shorter hours and longer holidays than non-union contracts.[6] One possibility is that Europeans work less than

[6] A separate, interesting, question is why union contracts look different than non-union contracts. A naive answer would be to say that unions have more bargaining power with respect to employers than individual workers (which is probably true) so they get better terms. But these better terms could be in the form of higher wages or less work. Why do unions prioritize leisure over

Americans because they live in a more unionized labor market.

7.4 A Dynamic Model

So far we have looked at consumption-savings decisions and consumption-leisure decisions as separate problems. In reality, households are making both decisions: how much to work and how much to consume. Do those decisions interact or is it OK to look at them in isolation? Let's see how a household would behave if they had to make both decisions at once.

The household solves the following problem:

$$\begin{gathered}\max_{c_1,l_1,c_2,l_2} u\left(c_1\right)+v\left(l_1\right)+\beta\left[u\left(c_2\right)+v\left(l_2\right)\right]\\ s.t.\\ c_1+\frac{1}{1+r}c_2\leq w_1\left(1-l_1\right)+\frac{1}{1+r}w_2\left(1-l_2\right)\end{gathered} \tag{7.4.1}$$

The household now has four goods to choose from: consumption in each period and leisure in in each period. As in the consumption-savings problem from Chapter 6, β represents how much it discounts the future. The budget constraint just says that the present value of consumption must be less or equal than the present value of income. Income in period t is given by the wage w_t times the amount of time the household dedicates to market work $1-l_t$.

The Lagrangian for this problem is:

$$L\left(c_1,l_1,c_2,l_2,\lambda\right)=u\left(c_1\right)+v\left(l_1\right)+\beta\left[u\left(c_2\right)+v\left(l_2\right)\right]-\lambda\left[c_1+\frac{1}{1+r}c_2-w_1\left(1-l_1\right)-\frac{1}{1+r}w_2\left(1-l_2\right)\right]$$

and the first-order conditions are:

$$\begin{aligned}u'\left(c_1\right)-\lambda&=0\\ v'\left(l_1\right)-\lambda w_1&=0\\ \beta u'\left(c_2\right)-\lambda\frac{1}{1+r}&=0\\ \beta v'\left(l_1\right)-\lambda w_1\frac{1}{1+r}&=0\end{aligned}$$

which we can summarize as:

$$\frac{v'\left(l_t\right)}{u'\left(c_t\right)}=w_t \qquad \text{for } t=1,2 \tag{7.4.2}$$

$$u'\left(c_1\right)=\beta\left(1+r\right)u'\left(c_2\right) \tag{7.4.3}$$

Equation (7.4.2) is just like equation (7.2.3), except that now it applies to both periods. In each period the

consumption more than individual workers?

household must be indifferent at the margin between dedicating a unit of time to leisure or to market work. Equation (7.4.3) is the Euler equation again: no matter how the household obtains its income, this equation describes how it distributes its consumption over time.

So far it would seem that looking at the consumption-savings decision and the consumption-leisure decision together doesn't add very much. We just get back the conclusions we got when we looked at the two problems separately. However, looking at the decisions jointly allows us to ask some questions that were not possible before. One that will be important later is the following. Suppose wages increase *temporarily*: how does the household change its labor supply? How does the answer change if the increase is permanent?

Let's look at this mathematically first and then think about what it means. Take equation (7.4.2) and solve for labor supply in period 1, which we denote by $L_1 = 1 - l_1$.

$$\begin{aligned} \frac{v'(l_1)}{u'(c_1)} &= w_1 \\ \Rightarrow v'(l_1) &= w_1 u'(c_1) \\ \Rightarrow l_1 &= (v')^{-1}(w_1 u'(c_1)) \\ \Rightarrow L_1 &= 1 - (v')^{-1}(w_1 u'(c_1)) \end{aligned} \tag{7.4.4}$$

where $(v')^{-1}$ denotes the inverse of v'. Since the marginal utility of leisure is decreasing, $(v')^{-1}$ is a decreasing function. Now compare two experiments: a temporary increase in wages (w_1 rises but w_2 does not) or a permanent increase (w_1 and w_2 both rise). If we look at the right hand side of (7.4.4), w_1 is, by assumption, the same in both experiments. However, c_1 will be different. As we saw in Chapter 6, consumption reacts more strongly to a permanent increase in income than to a temporary increase in income. That means that if we compare the two experiments, the one where the wage increase is *temporary* will have: *lower* c_1, (because the raise is temporary); therefore *higher* $u'(c_1)$ (because $u'(c)$ is a decreasing function); therefore *lower* $(v')^{-1}(w_1 u'(c_1))$ (because $v'(l)$ is a decreasing function); therefore *lower* l_1 and *higher* L_1. Conclusion: the household's labor supply rises more in response to a temporary increase in wages than it does to a permanent one.

What is going on in economic terms? If a wage increase is temporary, the household doesn't really feel much richer than before, so the income effect that pushes it to increase leisure is relatively weak. This makes the substitution effect dominate: time is temporarily very expensive so the household decides to sell more of it and keep less of it.

This effect is at the heart of Uber's surge-pricing strategy. When Uber detects that there are many users who want rides and not a lot of riders available, it increases fares. From the point of view of Uber drivers, this is like a temporary increase in wages. Because the increase is temporary, Uber drivers don't really feel much richer than before, and they react by getting in their car and offering rides to take advantage of the surge pricing.[7] Interestingly, in a study of New York City taxi drivers, Camerer et al. (1997) argued that they did not behave in this way at all and, if anything, reduced their working hours on days where they were temporarily getting higher income per hour. The exact interpretation of these findings is somewhat disputed.

[7]Also, Uber users may react to surge pricing by not taking as many rides. This is just the price adjusting to equate supply and demand! The only difference is that, instead of anonymous market forces, here Uber is actively managing the price adjustment.

7.5 Equilibrium in the Labor Market

So far we have looked at the household's labor supply decision in isolation. Let's put this together with labor demand to find an equilibrium. We'll first look at the frictionless competitive case and then we'll think about what happens if matching workers to jobs involves a process of search. We'll limit ourselves to the one-period case.

A Competitive Labor Market

As we saw in Chapter 4, in a competitive market firms will demand labor up to the point where the marginal product of labor equals the wage. If the production function is $F(K, L)$, then (taking the capital stock K as given), the equation $w = F_L(K, L)$ gives us a labor demand curve. For each possible wage level w, solving this equation for L tells us how much labor the representative firm is willing to hire. This demand curve is downward sloping. This comes from the fact that the production function is assumed to be concave in L, which is equivalent to saying that there is a diminishing marginal product of labor. For example, if $F(K, L) = K^{\alpha}L^{1-\alpha}$ then we have:

$$\begin{aligned} w &= (1-\alpha)K^{\alpha}L^{-\alpha} \\ \Rightarrow L &= (1-\alpha)^{\frac{1}{\alpha}}Kw^{-\frac{1}{\alpha}} \end{aligned}$$

On the workers' side, we can find a labor supply curve by doing the same steps that lead to equation (7.2.4). Replacing c from the budget constraint into the first order condition (7.2.3) we get:

$$\frac{v'(l)}{u'(w(1-\tau)(1-l)+T)} = w$$

This equation implicitly defines a relationship between the wage w and the amount of labor the worker supplies $1 - l$, i.e. a labor supply curve. As we have seen, the slope of this relationship is ambiguous in general, depending on the relative strengths of income and substitution effects, so the labor supply curve could be upward or downward sloping (a downward sloping labor supply curve is sometimes called "backward-bending").

By plotting the labor demand curve and the labor supply curve on the same plot, we can determine the equilibrium wage and the equilibrium amount of labor, as shown in Figure 7.5.1. The figure shows two examples, one where the substitution effect dominates so the labor supply curve is upward-sloping and one where the opposite happens. In either case, we can find the equilibrium wage and the equilibrium amount of labor by finding the intersection of the two curves.

We can also ask what would happen in the labor market in response to various changes. Figure 7.5.2 illustrates two possibilities. In the first panel we have the effects of a shift in the marginal product of labor curve $F_L(K, L)$. At any given wage level, firms want to hire more workers. This leads to an increase in wages and (in this example, where the substitution effect dominates), to an increase in L. The second panel shows an increase in government transfers T. As we have seen, this has an income effect, so for any given wage level, the worker wants to work less, leading to a leftward shift in the labor supply curve. This leads to higher wages

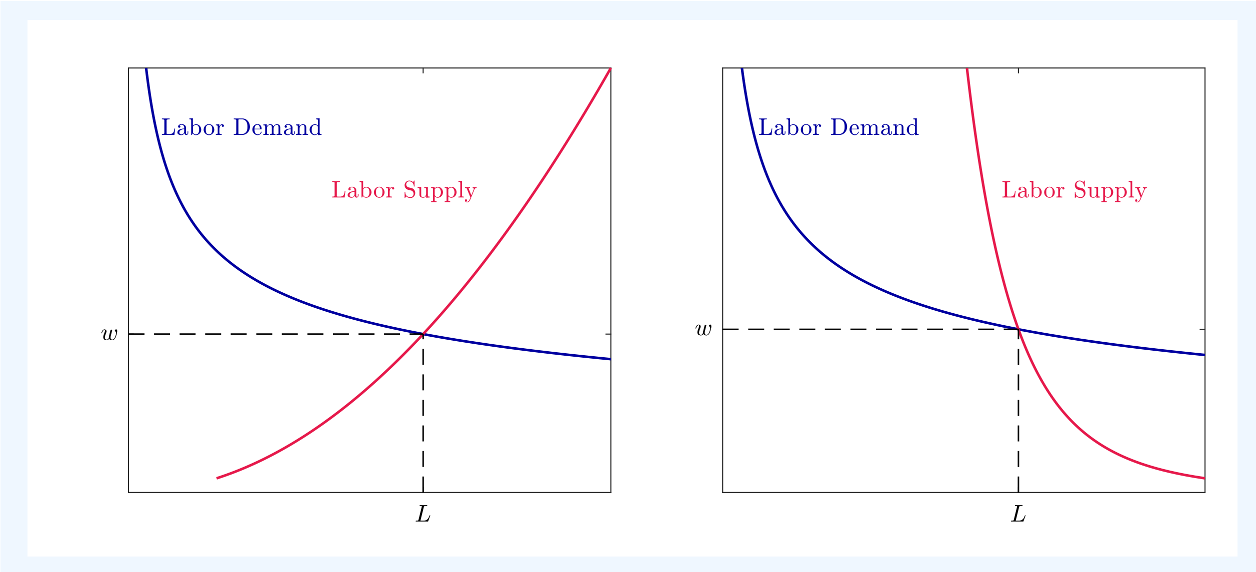


Fig. 7.5.1: *Equilibrium in the labor market.*

and a fall in L.

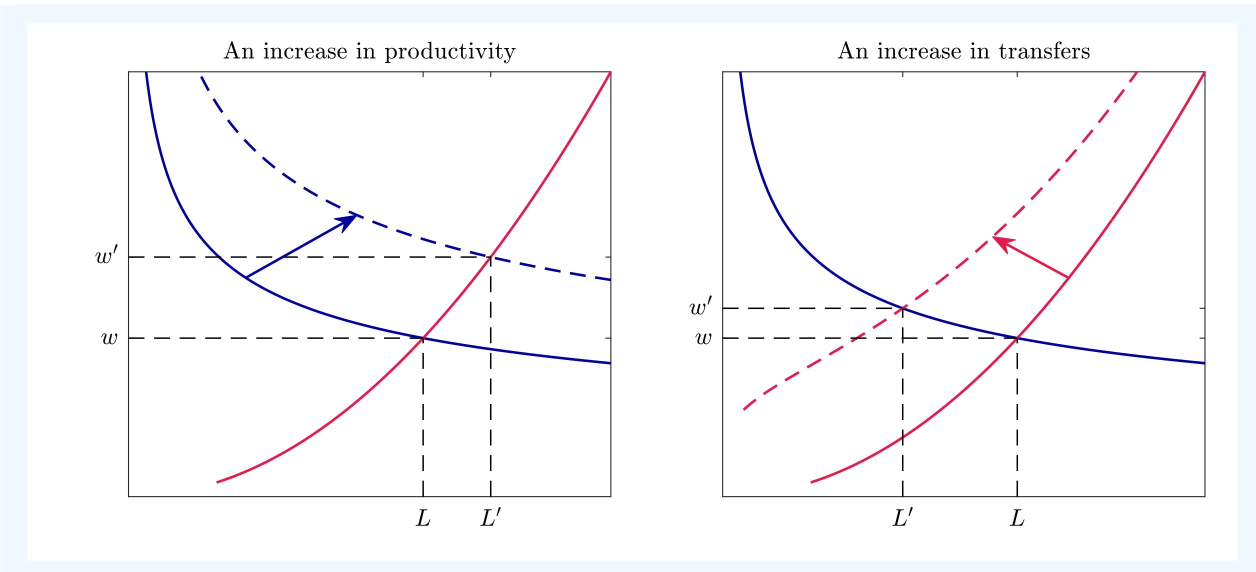


Fig. 7.5.2: *Effects of changes on the labor market.*

Search

So far we have assumed that the labor market is competitive and everyone can work as much or as little as they want at the market wage. This is not a useful way to think about unemployment because, by assumption, there is no unemployment. For that matter, it's also not a good way to think about vacancies either because, by assumption, firms can hire as many workers as they can and they will never have unfilled vacancies. One theory of why unemployment and unfilled vacancies coexist, developed by Diamond (1982), Mortensen (1982) and Pissarides (1985), is built on the assumption that the process of search is frictional: the workers that are looking for jobs and the firms who need them might fail to find each other. Let's look at how this theory works.

To keep things simple, we are going to assume away everything that has to do with the decision of how much to work. We normalize the size of the population to 1 and assume everyone in the population is in the labor force. In order to actually produce output, a worker has to be matched with a firm, because firms have access to the productive technology. If a worker is matched with a firm, we call that a job; it assigns the worker a task in the production function, which results in $y \equiv F_L(K, L)$ additional units of output.

The process by which workers find jobs works as follows. At the beginning of the period, $1 - U$ workers already have jobs, while U are still looking, so U represents the initial unemployment rate, which we take as given. Firms that want workers advertise their vacancies. Advertising a vacancy costs χ units of output. This represents literal advertising costs as well as the cost of things like selecting and interviewing applicants. The number of workers who find jobs depends on the total number of unemployed workers and the total number of vacancies that firms create (denoted by V), according to some function $m(V, U)$. This is known as a "matching function". It attempts to capture, in simplified form, all the process of workers searching for jobs and firms searching for workers. We'll assume that m is increasing and concave in both U and V. This means that more workers will find jobs when there are more vacancies, but the marginal impact of an additional vacancy on the number of jobs created is decreasing; similarly, more firms will be able to fill vacancies if there are more unemployed workers looking for jobs, but the marginal impact of an additional unemployed worker is decreasing. Define the vacancy filling rate as:

$$q(V, U) = \frac{m(V, U)}{V}$$

The concavity of m implies that $q(V, U)$ is decreasing in V: the probability of filling any one vacancy falls with the number of vacancies. Also, $q(V, U)$ is increasing in U.

If a worker and a firm manage to match, then they need to decide what wage the worker will be paid. We are going to assume that they bargain over the wage in a specific way, known as Nash bargaining. They first look at what's going to happen if they cannot reach an agreement. In this case, the firm will have an unfilled vacancy that will produce no output and no profits, while the worker will be unemployed and get some exogenous amount b, which could represent unemployment benefits or the value of leisure. After doing that, they say to each other: "if we do reach an agreement, we'll get a total of y instead of b, so that's $y - b$ better. Let's find a deal so that we split the $y - b$ of surplus that is generated by this job". They then set the wage w so that the worker gets a fraction μ of the surplus and the firm gets $1 - \mu$. The parameter μ measures the bargaining power of the worker. This means that the wage will be: $w = b + \mu(y - b)$ and the firm's profits will be $y - w = (1-\mu)(y-b)$.

Now turn to the firm's decision to post a vacancy. A vacancy costs χ and then, with probability $q(V,U)$, it results in a hire, which gives the firm a profit of $(1-\mu)(y-b)$. Firms will find it profitable to create vacancies until:

$$\chi = q(V,U)(1-\mu)(y-b) \tag{7.5.1}$$

Solving equation (7.5.1) for V tells us how many vacancies will be created, as illustrated by Figure 7.5.3. If V is higher than the number that makes (7.5.1) hold, then the probability of filling a vacancy is too low, firms will on average not recoup the vacancy costs and would prefer not to hire. If V is lower, then firms will on average make profits beyond what is needed to recoup vacancy costs and would open more vacancies.

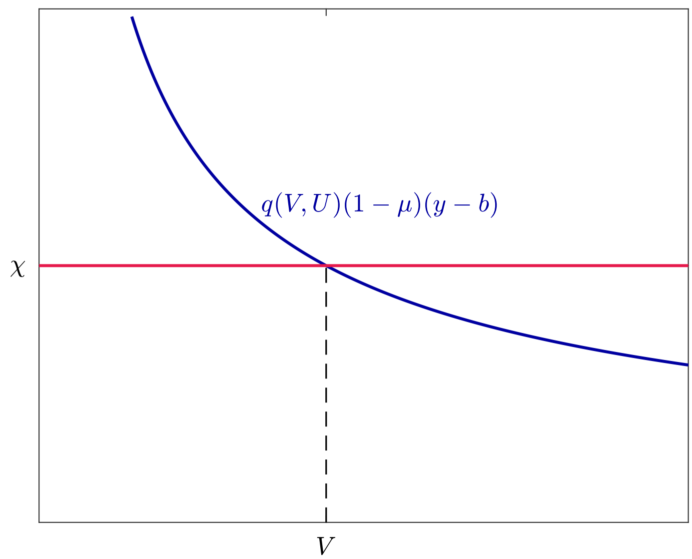


Fig. 7.5.3: *Equilibrium in the Diamond-Mortensen-Pissarides search model.*

Knowing V lets us figure out what the unemployment rate will be once hires have taken place. Denote this by U'. Since a total $m\,(V,U)$ workers find jobs, we have:

$$U' = U - m(V,U) \tag{7.5.2}$$

Let's see what happens to the labor market when something changes. The left panel of Figure 7.5.4 shows how the labor market reacts to an increase in the marginal product of labor y. Other things being equal, this makes posting vacancies more profitable, so firms react by posting more vacancies until condition (7.5.1) is restored. The right panel shows what happens if unemployment benefits b are increased. The effects here are more subtle. Higher unemployment benefits increase workers' outside option if they were fail to reach an agreement while bargaining. This improves their bargaining position, and therefore raises equilibrium wages, lowering firm profits. Firms respond to this by posting fewer vacancies, restoring condition (7.5.1).

Seeing how V reacts to various changes gives us a way to think about the Beveridge Curve. Equation (7.5.2)

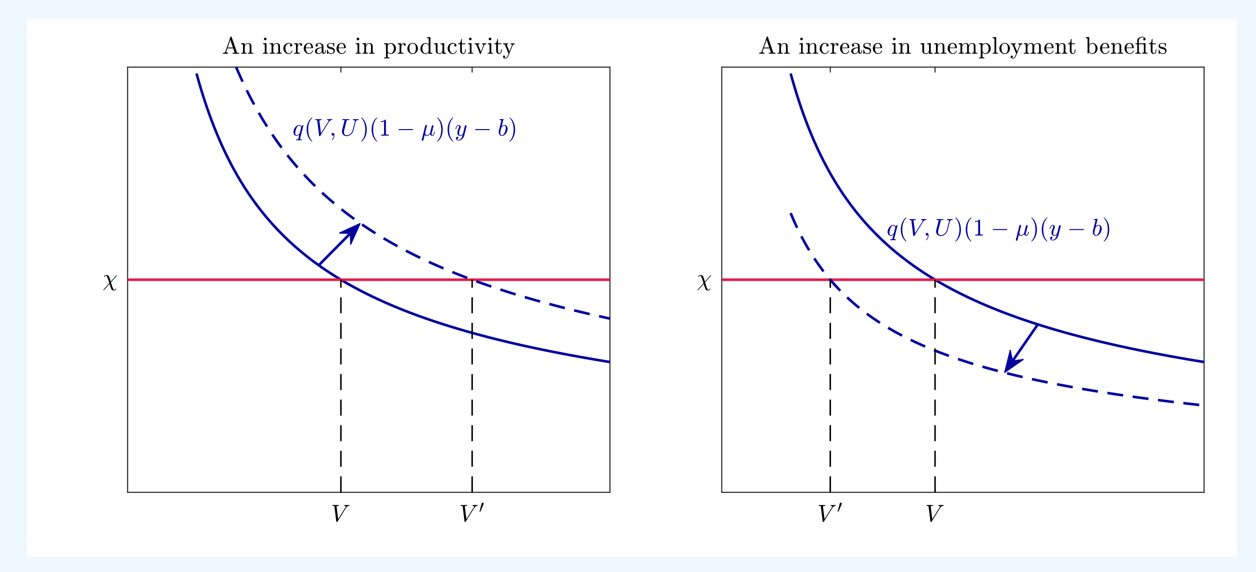


Fig. 7.5.4: *Effects of changes in the Diamond-Mortensen-Pissarides search model.*

implies a negative relationship between vacancies and after-hiring unemployment (taking initial-unemployment as given), which is what the Beveridge Curve says. Other things being equal, anything that leads firms to post more vacancies will result in more workers finding jobs and therefore lower unemployment. There is some debate as to whether this is a satisfactory explanation. Shimer (2005) argued that the movements in the vacancies-to-unemployment ratio produced by this model are too small to fit the data well.

Exercises

7.1 Labor Supply

Suppose household preferences are given by:

$$u(c,l) = c + \theta \log(l)$$

where c is consumption, l is leisure and θ is a parameter. Households have a total of one unit of time and can supply labor in a competitive labor market at a wage w.

(a) Find an expression for the fraction of their time that households spend in market work.

(b) If this was the right model and one looked at households in different countries, how would hours of work correlate with wage levels? How does this compare to the empirical evidence?

7.2 Military Service

An economy is populated by a representative household, whose preferences are described by:

$$u(c) + v(l)$$

where c is consumption and l is leisure. The household has 1 unit of time so its budget is:

$$c = w(1 - l) - \tau$$

where w is the wage rate and τ is a tax that the government collects. Notice that τ is a lump-sum tax, not an income tax: the household must pay the same amount regardless of how much income it earns.

(a) Set up the Lagrangian for the household's optimization program and find first-order conditions.

(b) Use the budget constraint to replace c in the first-order condition to obtain a single equation that relates the household's choice of l to w and τ.

(c) Suppose that the government uses all the tax revenue to hire soldiers for the army, paying them the prevailing wage rate. How many units of soldiers' time can the government afford? Denote this number by m.

(d) Now suppose that instead of taxing citizens to hire soldiers, the government imposes compulsory military service: the representative household has to dedicate m units of time to serving in the army, unpaid. The rest of their time they can use as they please, and they pay no taxes. Set up the household's optimization program under this policy.

(e) Show that the representative household's level of consumption, leisure, army labor and non-army labor are the same under both policies. Explain.

7.3 Consumption Taxes and Labor Supply

Suppose a household solves the following problem:

$$\max_{c,l} u(c) + v(l)$$
$$s.t. \quad (1 + \tau_c) c \leq w(1 - \tau)(1 - l)$$

where c is consumption, l is leisure and τ is an income tax. τ_c is a consumption tax, so that if the household wishes to consume c it must pay $(1 + \tau_c) c$.

(a) Draw the household's budget constraint. What is the slope?

(b) Find the first-order conditions for the household's consumption-leisure decision. How do the effects of τ and τ_c relate to each other? Explain.

(c) If the household chooses c^* and l^*, how much tax revenue does the government collect?

(d) Suppose the government had originally set $\tau_c = 0$ and $\tau > 0$ and now wants to enact a tax reform that uses consumption taxes instead of income taxes. What would be the level of τ_c that leaves the decisions of the household unchanged?

(e) How much revenue does the government collect under the new system? How does it compare to the old system? Explain.

7.4 Puerto Rico

Read the following article: https://www.economist.com/united-states/2006/05/25/trouble-on-welfare-island. Is what was going on in Puerto Rico consistent with the theories discussed in this chapter? Discuss.

7.5 Prescott's Calculation

Suppose preferences for consumption and leisure are:

$$u(c,l) = \log(c) + \alpha \log(l)$$

and households solve:

$$\max_{c,l} u(c,l)$$
$$s.t. \quad c = w(1-\tau)(1-l) + T$$

(a) Find first-order conditions for the consumption-leisure decision.

(b) Use the budget constraint to solve for leisure l. You should get an explicit expression for l as a function of w, τ, T and α.

(c) Suppose $T = 0$. How does l respond to the tax rate τ? What does this mean?

Now suppose that in both Europe and the US we have:

$$\alpha = 1.54$$
$$w = 1$$

but in the US we have:

$$\tau = 0.34$$
$$T = 0.102$$

while in Europe we have:

$$\tau = 0.53$$
$$T = 0.124$$

(d) Compute the amount of leisure chosen in the US and Europe. If we interpret 1 as your entire adult lifetime, what fraction of their adult lives do people in Europe and the US work? Comment on the respective role of taxes and transfers in this analysis using your answers to parts (b) and (c).

(e) The values for τ and T above are not arbitrary. If you did the calculations correctly, you should find that both governments have balanced budgets (up to rounding error), i.e. they redistribute all the tax revenue back as transfers. Check that this is the case.

(f) Assuming the production function is $Y = L = 1 - l$, how much lower is GDP per capita in Europe compared to the US?

(g) Compute the relative welfare of Europe by solving for λ in the following equation:

$$u\left(c_{Europe}, l_{Europe}\right) = u\left(\lambda c_{US}, l_{US}\right)$$

Interpret the number λ that you find

(h) How do the answers to questions (f) and (g) compare? Why?

(i) Suppose a European policymaker sees Prescott's calculation and concludes that Europe could increase its welfare by a factor of $\frac{1}{\lambda}$ by reducing its tax rate and level of transfers to US levels. Do you think they are right? Why? Don't answer this question mechanically: think about what this calculation does and what it leaves out.

In any calculation of this sort, an important parameter is the elasticity of labor supply. One definition of elasticity that is often looked at by labor economists is known as the "Frisch" elasticity. It is based on the answer to the following question: "suppose we increased wages but adjusted the household's income so that consumption remained constant: how would labor supply change?"[8] Let's calculate the Frisch elasticity in Prescott's model.

(j) Use your answer to part (a) to find an expression for labor supply $(1 - l)$ in terms of $w(1-\tau)$, c and α. Notice that now we are holding consumption constant, so the idea is that we *don't* replace c from the budget constraint like we did in part (b).

(k) Use your answer to part (j) to find an expression for $\frac{\partial(1-l)}{\partial w(1-\tau)} \frac{w(1-\tau)}{1-l}$, i.e. the elasticity of labor supply with respect to after-tax wages, holding consumption constant.

(l) Plug in the values of α, τ, w, c and l that you found for the US case into the expression for elasticity. What number do you get? Empirical estimates of this elasticity are usually in the range of 0.4 to 1. How does that compare to the elasticity implied by Prescott's model? Why does that matter for our conclusions about tax policy?

7.6 The Protestant Work Ethic

Suppose that the representative household has preferences given by:

$$u(c, l) = \log(c) + \theta \log(l)$$

where c is consumption, l is leisure and θ is a parameter. The production function is given by:

$$F(K, L) = K^{\alpha} L^{1-\alpha}$$

[8] Notice that it's not exactly the same as isolating the substitution effect as in Figure 7.2.2, because that was holding utility constant instead of consumption; but it's related.

where K is the capital stock, which we take as given and equal to 1, L is labor and α is a parameter. Firms hire workers in a competitive labor market where the wage is w.

(a) What will be the equilibrium wage w?

(b) How does w depend on θ? Suppose the household suddenly develops a so-called Protestant work ethic and stops enjoying leisure so much, what will happen to real wages? Explain.

7.7 The Beveridge Curve

Assume the labor market is well described by the search model presented in Section 7.5.

(a) Suppose there is an improvement in recruiting technology, so that the matching function becomes $Am(V, U)$, with $A > 1$. What happens to the number of vacancies? What happens to the after-hiring unemployment rate? Would this produce something that looks like a Beveridge Curve?

(b) Suppose there is an increase in the initial (before-hiring) unemployment rate. What happens to the number of vacancies? What happens to the after-hiring unemployment rate? Would this produce something that looks like a Beveridge Curve?

CHAPTER 8

Investment

When we looked at the Solow model we assumed that investment was an automatic consequence of savings. In this chapter we are going to think about investment decisions directly. What incentives govern the decisions to invest? Let's start with an example.

Example 8.1.
The world lasts two periods. A firm is considering building a factory in period 1 in order to produce output in period 2. Building the factory costs 1000 (the units here are consumption goods). If the firm builds the factory, this will result in additional revenues of 2500 and additional costs of 1400 in period 2. After period 2, the world is over. Is building the factory a good idea?

Let's start with a naive calculation. Net of costs, in period 2 the firm will get an additional $2500 - 1400 = 1100$ if it builds the factory. Since this is more than the cost of building the factory, this would seem to suggest that building the factory is a good idea.

What's missing from this analysis? Building the factory requires giving up 1000 *in period 1* in order to obtain 1100 *in period 2*. But goods today and goods tomorrow are different goods! Just concluding that you should build the factory is like saying that converting one large apartment into two studios is a good idea because two is more than one. To do it right, you need to know the relative price of apartments and studios. Likewise, in order to decide whether you should build the factory, you need to know the relative price of goods in different periods.

As we saw in Chapter 6, the *real interest rate* is the relative price of goods in different periods. Let's go over why this is the case. Imagine that anyone can borrow or lend as much as they want at the real interest rate r, and all loans are always repaid. This means that anyone can take one good to the market in period 1 and exchange it for $1 + r$ goods in period 2. Conversely, anyone can obtain one good in period 1 in exchange for giving up $1 + r$ goods in period 2. Hence, one good in period 2 is worth $\frac{1}{1+r}$ goods in period 1.

Now let's go back to the investment decision. If the firm invests, it gives up 1000 period-1 goods in exchange for 1100 period-2 goods that are worth $\frac{1100}{1+r}$ period-1 goods. This is a good idea as long as:

$$\frac{1100}{1+r} > 1000$$

$$\Rightarrow r < 0.1$$

If the real interest rate is low enough, then this investment project is worthwhile; otherwise it's not.

Notice that whether this factory is a good idea does not depend on whether the firm has 1000 to begin with. If $r < 0.1$ and the firm doesn't have 1000, it is a good idea to borrow them in order to build the factory: the net revenues from the factory will be more than enough to pay back the loan with interest. Conversely, if $r > 0.1$ and the firm does have 1000 available, it is better off lending and earning interest than building the factory. Of course, this relies on the assumption of limitless borrowing or lending at the same interest rate.

8.1 Present Values

Let's try to generalize from the example above. Most investment projects are expected to produce revenues for more than two periods. How should one evaluate them? The key is to figure out what is the right price at which to value goods that one will receive several periods in the future. How much is a period-t good worth?

Let's imagine that interest rates are constant at r per period. This means you can take one good in period 0, lend it to obtain $(1 + r)$ in period 1, then lend $(1 + r)$ in period 1 to obtain $(1 + r)^2$ in period 2 and so on until you are finally left with $(1 + r)^t$ goods in period t. This means that one good in period t is worth $\frac{1}{(1+r)^t}$ goods in period 0.

Now let's evaluate some arbitrary project. We'll summarize the project in terms of the dividends it will produce at each point in the future. We'll denote the dividend from the project in period t by d_t. The total value of the project is the sum of all of these dividends, each of them valued in terms of the period-0 goods that they are worth. In other words:

$$V = \sum_{t=1}^{\infty} \frac{d_t}{(1 + r)^t} \tag{8.1.1}$$

Formula (8.1.1) is known as a present-value formula: it tells us what is the present value of any possible sequence of dividends.

Armed with formula (8.1.1), deciding whether an investment project is worthwhile is straightforward. Suppose the project costs I, then the *net* present value of the project is defined as:

$$NPV \equiv V - I$$

Projects are worth doing if and only if the net present value is positive. In the example above, the net present value was:

$$\begin{aligned} NPV &= V - I \\ &= \frac{1100}{1 + r} - 1000 \end{aligned}$$

which was positive if $r < 0.1$.

Example 8.2.
Opening a restaurant costs $100,000 dollars in year 0. In year 1 the restaurant will not be very well known, so it is expected to make a loss of $10,000. In year 2, the restaurant will exactly break even. Starting in year 3, the restaurant will be a big success, earning $40,000, $50,000, $60,000 and $70,000 in years 3 to 6 respectively. In year 7, quinoa burgers will suddenly fall out of fashion so the restaurant will close down forever. The interest rate is 10%. Is opening the restaurant a good idea?

Let's compute the net present value:

$$\begin{aligned} NPV &= \frac{-10,000}{1.1} + \frac{0}{1.1^2} + \frac{40,000}{1.1^3} + \frac{50,000}{1.1^4} + \frac{60,000}{1.1^5} + \frac{70,000}{1.1^6} - 100,000 \\ &= 31,881 \end{aligned}$$

In this example, the net present value is positive so opening the restaurant is a good idea.

The Gordon Growth Formula

There is a special case in which formula (8.1.1) becomes very simple. Suppose that the dividends from a project are expected to grow at the constant rate g forever, so that $d_{t+1} = (1+g)\,d_t$. This means that period-t dividends will be $d_t = d_1(1+g)^{t-1}$. In this case, formula (8.1.1) becomes:

$$\begin{aligned} V &= \sum_{t=1}^{\infty} \frac{d_1(1+g)^{t-1}}{(1+r)^t} && \text{(replacing } d_t = d_1(1+g)^{t-1}\text{)} \\ &= \frac{d_1}{1+g} \sum_{t=1}^{\infty} \left(\frac{1+g}{1+r}\right)^t && \text{(rearranging)} \\ &= \frac{d_1}{1+g} \frac{\left(\frac{1+g}{1+r}\right)}{1-\frac{1+g}{1+r}} && \text{(appying the formula for a geometric sum)} \\ &= \frac{d_1}{r-g} && \text{(simplifying)} \end{aligned} \tag{8.1.2}$$

Formula (8.1.2) is known as the Gordon growth formula. It tells us the present value of any project as a function of the initial level of dividends, its growth rate and the interest rate. Notice that V is high when either g is high or r is low. High r pulls V down because it makes future goods less valuable in terms of present goods. High g offsets this and pulls V up because it makes future dividends larger.

Asset Prices, Adjustment Costs and the Q theory of investment

So far we have used formula (8.1.1) to think about the value of *potential* projects. But one can apply the same logic to think about productive projects that are already in place. Formula (8.1.1) is *also* the answer to the following question: if the project is already in place, how much should you be willing to pay to buy it?

This is not merely a hypothetical question. There are many markets where people actively trade assets that are already producing dividends. In the stock market, people trade shares in companies that are already operating; in the commercial real estate market, people trade buildings that are already yielding rental income; in the housing market, people trade houses that are already producing housing services; in the bond market, people trade bonds that are already paying coupons. Formula (8.1.1) tells us what should be the price at which people trade in these markets.

Looking at asset prices in actual markets can be extremely useful in guiding investment decisions. Suppose that a construction company knows that it can build an office building for X dollars. In order to decide whether this is a good idea, they need to decide whether the present value of the rental income they will obtain is more than this construction cost. This can be difficult to predict. However, if there is an active market where comparable office buildings change hands, they can look at what prices people are paying to buy them. If comparable office building sell for more than X, then it must be that investors think the present value of rents is more than X, which would make the project a good idea.

One version of this idea is known as the Q theory of investment. It starts by defining an object called Q, or sometimes Tobin's Q (after James Tobin):

$$Q \equiv \frac{\text{Market Value}}{\text{Book Value}} \tag{8.1.3}$$

The book value of a company measures its accumulated investment, net of depreciation. In theory, this should approximately measure how much you need to invest to build a company just like it. Suppose this was exactly true and you could build an exact replica of a company by investing an amount equal to the firm's book value. Since it's an exact replica, its present value should be equal to the original company's market value. Whenever $Q > 1$, then building a replica is a good idea. Under this extreme assumption of perfect replicability, we should see infinite investment whenever $Q > 1$ and no investment at all whenever $Q < 1$.

A less extreme version of this argument involves adjustment costs. Suppose expanding a firm involves paying an adjustment cost in addition to the cost of the additional capital. This adjustment cost can be the cost of physically installing machines, training workers to use them, or even the cost of making the decision to invest. Adjustment costs can be represented by a function $\Psi(I, K)$ that says how much extra a firm needs to pay if its current capital stock is K and it wants to invest I to expand up to $K' = K + I$. This function could take many possible shapes; we'll assume that:

$$\Psi\left(I, K\right) = \frac{\psi}{2}\frac{I^2}{K} \tag{8.1.4}$$

The parameter ψ scales adjustment costs, so higher ψ means that adjusting the capital stock is more expensive. Furthermore, (8.1.4) implies that the marginal adjustment cost is:

$$\frac{\partial \Psi(I, K)}{\partial I} = \psi \frac{I}{K}$$

which is higher when investment is higher as a proportion of existing capital. In other words, proportionately small adjustments to the firm's scale are cheap to do, but large adjustments are expensive. This is an assumption, and it's not obvious whether it's a good one or not. Indeed, some economists have argued that

$\Psi(\cdot)$ could involve a fixed adjustment cost of undertaking any non-zero level of investment. Still, we are going to assume (8.1.4) holds and see what the implications are.

Consider the problem of a firm that needs to decide how much to invest today. Its capital stock is K and its stock market value is QK. Furthermore, it knows that if increases its capital stock to K', the stock market value will rise to QK'.[1] The firm wants to maximize its stock market value net of the cost of investment. The problem of the firm is:

$$\max_I Q\left(K+I\right) - \left(I + \frac{\psi}{2}\frac{I^2}{K}\right)$$

The first order condition is:

$$\begin{aligned} &-1 - \frac{\psi I}{K} + Q = 0 \\ \Rightarrow &\frac{I}{K} = \frac{1}{\psi}\left(Q-1\right) \end{aligned}$$

Without adjustment costs, we had that investment would be infinite when $Q > 1$ and zero when $Q < 1$. With adjustment costs, we get a less extreme version of the same idea: the investment-to-capital-ratio $\frac{I}{K}$ depends positively on Q. If adjustment costs are low (so $\frac{1}{\psi}$ is high), then small changes in Q will lead to a strong response of $\frac{I}{K}$, and vice versa.

Figure 8.1.1 shows the relationship between Q and investment for the aggregate US economy. The relationship is positive, and has become especially strong in the last couple of decades.

8.2 Risk

In everything we've done so far, we have assumed that future dividends are perfectly known. Of course, uncertainty is a central aspect of any investment decision. How should this be taken into account?

Let's think again about a two-period problem, although the same principles apply to problems with any horizon. Imagine a project that will pay a dividend in period 2 but the size of the dividend depends on the "state of the world". The probability of a state s is denoted by $\Pr(s)$ and the dividend that the project will pay in state s is denoted $d(s)$.

Example 8.3.
The project is a farm. The dividend it will pay in period 2 depends on whether it rains. The interest rate is 10%.

[1] We sometimes distinguish between "average Q", which is defined by equation (8.1.3), and "marginal Q, which is defined by:

$$\text{Marginal } Q \equiv \frac{\partial \text{Market Value}}{\partial \text{Investment}}$$

A rational firm should make its investment decisions on the basis of marginal Q: how much does the value of the firm change with an additional unit of investment. By assuming that the stock market value will be QK' for any K' we are assuming that marginal Q is equal to average Q.

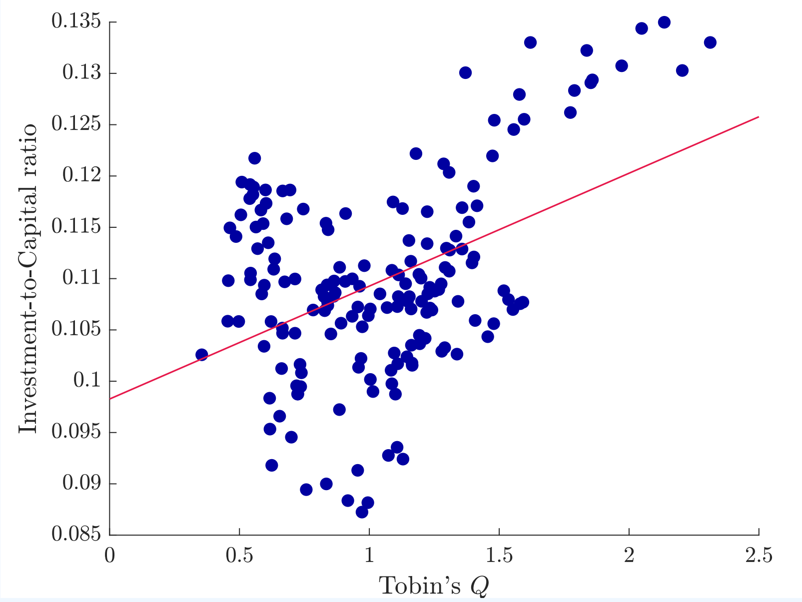


Fig. 8.1.1: *Tobin's Q and investment. Source Andrei et al. (2019).*

State of the world	Probability $\Pr(s)$	Dividend $d(s)$
$s = \textit{Rain}$	0.75	100
$s = \textit{Drought}$	0.25	8

We are going to look at this problem from the perspective of a household who has to answer the question: is it worth investing p in period 1 to buy a unit of this project? What is the price p that makes the household indifferent between investing in the project or not? One naive way to do this would be to take an average of the possible dividends and discount it at the interest rate. This would give the following answer:

$$p = \frac{0.75 \times 100 + 0.25 \times 8}{1 + 0.1} = \frac{77}{1.1} = 70 \tag{8.2.1}$$

On average, the farm will pay a dividend of 77. Discounting it at an interest rate of 10% gives a present value of 70. However, this way of thinking about the problem is not quite right because it ignores the fact that people are risk averse and the farm is risky. What's the right way to do it?

It turns out that the answer depends on how this particular project fits with the rest of the household's decisions. Suppose that, if it does not invest in this project, the household will be consuming c_1 in period 1 and $c_2(s)$ in period 2 in state of the world s. Now let's ask the household how many units of the project it wants to buy, assuming that it has to pay p for each unit. Mathematically, the household's problem is:

$$\max_x u(c_1 - xp) + \beta \sum_s \Pr(s)\, u(c_2(s) + xd(s)) \tag{8.2.2}$$

Buying x units of the project reduces period-1 consumption by xp, bringing it down to $c_1 - xp$. On the other hand, owning x units of the project increases period-2 consumption by $xd(s)$ in state of the world s, bringing it up to $c_2(s)+xd(s)$. Equation (8.2.2) describes the objective function of a household that wants to maximize expected utility.

The first order condition for this problem is:

$$-pu'(c_1-xp)+\beta\sum_s \Pr(s)\,u'(c_2(s)+xd(s))\,d(s)=0 \tag{8.2.3}$$

Suppose we wanted to find the price at which the household is exactly indifferent between buying a little bit of the asset and not buying at all. We would look for the price such that $x=0$ satisfies equation (8.2.3). Setting $x=0$ and solving for p we get:

$$p=\frac{\beta\sum_s \Pr(s)\,u'(c_2(s))\,d(s)}{u'(c_1)} \tag{8.2.4}$$

Now let's assume that the choices of c_1 and $c_2(s)$ are consistent with intertemporal maximization, so that the Euler equation holds:[2]

$$u'(c_1)=\beta(1+r)\sum_s \Pr(s)\,u'(c_2(s))$$

Replacing $u'(c_1)$ into (8.2.4) and simplifying gives:

$$p=\frac{1}{1+r}\frac{\sum_s \Pr(s)\,u'(c_2(s))\,d(s)}{\sum_s \Pr(s)\,u'(c_2(s))} \tag{8.2.5}$$

Equation (8.2.5) contains everything you might ever want to know about finance. It says that when there is risk, the present value of an asset comes not just from discounting average dividends but by discounting a *weighted* average of dividends, where the weights are proportional to marginal utility. If $u'(c_2(s))$ was the same for all s, then formula (8.2.5) reduces to;

$$\begin{aligned}
p &= \frac{1}{1+r}\frac{u'(c_2)\sum_s \Pr(s)\,d(s)}{u'(c_2)\sum_s \Pr(s)} && (u'(c_2)\ \text{factors out}) \\
&= \frac{\sum_s \Pr(s)\,d(s)}{1+r} && \text{(probabilities add up to one)} \\
&= \frac{\mathbb{E}(d)}{1+r} && (8.2.6)
\end{aligned}$$

which is what our naive calculation in (8.2.1) was doing. The reason this is not quite right in general is that $u'(c_2(s))$ might differ for different s.

Let's break down formula (8.2.5) a bit more. Recall from statistics the definition of a covariance between two random variables X and Y:

$$Cov(X,Y)\equiv\mathbb{E}(XY)-\mathbb{E}(X)\,\mathbb{E}(Y)$$

[2] This is the generalized version of the Euler equation that has *expected* marginal utility on the right hand side, as in equation (6.2.17)

Rearranging we obtain:

$$\frac{\mathbb{E}(XY)}{\mathbb{E}(X)} = \mathbb{E}(Y) + \frac{Cov(X,Y)}{E(X)} \tag{8.2.7}$$

Letting $X = u'(c)$ and $Y = d$, we can use formula (8.2.7) to rewrite the right hand side of (8.2.5) and obtain:

$$p = \frac{\mathbb{E}(d) + Cov(d, u'(c))}{1+r} \tag{8.2.8}$$

Formula (8.2.8) shows us the consequences of using a marginal-utility-weighted average rather than a simple average to value the project's dividends. Compared to (8.2.6) there is an additional term: $Cov(d, u'(c))$. How much a household is willing to pay for a marginal unit of a project depends on average dividends *and* on how those dividends co-vary with marginal utility. What is this telling us? The marginal utility of consumption measures how much the household values extra consumption in a particular state of the world. If dividends co-vary positively with marginal utility, then they provide the household extra consumption exactly when the household values it the most. This makes the asset attractive, so the household is willing to pay more than $\frac{\mathbb{E}(d)}{1+r}$ for it. Conversely, if dividends co-vary negatively with marginal utility, then the asset gives the household extra consumption exactly when the household values it the least. The household will pay less than $\frac{\mathbb{E}(d)}{1+r}$ to hold such an asset.

Let's see some examples.

Example 8.4.
The asset is a bet on a (fair) coin flip. If the coin turn out heads, it pays one dollar; if it turns up tails, it pays zero. The interest rate is zero.

Here $\frac{\mathbb{E}(d)}{1+r} = 0.5$. Furthermore, $Cov(d, u'(c)) \approx 0$. Why? Because the event "the coin turn up heads" is independent of the events that determine whether the household has high or low consumption (such as losing a job, getting a promotion, etc.). Therefore the household is willing to pay $p = 0.5$ for this asset.

Example 8.5. The asset is a bet on a (fair) coin flip. If the coin turn out heads, it pays one million dollars; if it turns up tails, it pays zero. The interest rate is zero.

Here $\frac{\mathbb{E}(d)}{1+r} = 500{,}000$. However, $Cov(d, u'(c)) < 0$. Why is this different from the previous example? Because now, if the household buys the asset, the outcome of the coin flip is a very big deal: the household will consume much more if the coin comes up head than if it comes up tails. Therefore marginal utility of consumption will be lower exactly when the asset pays a high dividend. A rational, risk-averse household will value this asset at $p < 500{,}000$.

Example 8.6. The asset is one dollar in Facebook shares and Sheryl is a Facebook employee.

Here it's likely that $Cov(d, u'(c)) < 0$. Why? The dividend from the asset is not by itself a big determinant of Sheryl's consumption because she only owns one dollar of it. A bigger determinant is how she's doing at her job: whether she gets a raise, gets fired, etc. However, she is more likely to get a raise and less likely to get fired if Facebook is doing well. Hence the asset tends to have higher dividends in the states of the world where Sheryl's consumption is high and values them less. Sheryl will be willing to pay less than $\frac{\mathbb{E}(d)}{1+r}$ for this asset.

Example 8.7.
The asset is a one-dollar health insurance contract. A household member gets sick with probability 0.1. If this happens, the insurance policy gives the household one dollar that can be used towards medical expenses. The household has no other medical insurance. The interest rate is zero.

Here $\frac{\mathbb{E}(d)}{1+r} = 0.1$ but $Cov\left(d, u'\left(c\right)\right) > 0$. Why? Absent other insurance, getting sick is expensive, so the household has to cut back on consumption to pay medical bills if a household member gets sick. This means that the asset pays exactly in those states of the world where marginal utility is high. Therefore a risk averse household would be willing to pay $p > 0.1$ for this asset.

8.3 The Marginal Product of Capital and Aggregate Investment

So far we have been talking about the decisions to invest in individual projects. In macroeconomics we usually care about what determines the *overall* level of investment. To think about that, we are going to abstract from the features of each individual investment project and go back to assuming that every project is identical. What determines how many investment projects take place?

We are going to imagine that the representative investment project consists simply of converting one unit of output into a unit of capital. If this is done in period t, then in period $t+1$, this unit of capital can be rented out to a firm at the rental rate of capital r^K_{t+1}. As we saw in Chapter 4, with competitive markets we have $r^K_{t+1} = F_K\left(K_{t+1}, L_{t+1}\right)$. In addition to the rental income, the investor gets back the depreciated capital, so in total he gets $F_K\left(K_{t+1}, L_{t+1}\right) + 1 - \delta$ goods at $t+1$. The net present value of the project is:

$$NPV = \frac{1 + F_K\left(K_{t+1}, L_{t+1}\right) - \delta}{1 + r_{t+1}} - 1 \tag{8.3.1}$$

The first term is what the investor gets in period $t+1$, valued in terms of period-t goods. Minus 1 is the cost of the project in period t. The NPV of a representative investment project is a decreasing function of the following-period capital stock K_{t+1}, other things being equal. Why? Because a higher capital stock means a lower marginal product of capital and therefore a lower rental rate of capital.

How does formula (8.3.1) help us figure out the total level of investment? The key thing to notice is that *the NPV of the representative investment project must be exactly zero.* Why? If it was positive, there would be positive-NPV projects left undone; conversely, if it was negative, it means negative-NPV projects are being done. Neither of these possibilities is consistent with projects being carried out whenever they have positive NPV. Setting the NPV to zero in (8.3.1) we get:

$$\frac{1 + F_K\left(K_{t+1}, L_{t+1}\right) - \delta}{1 + r_{t+1}} - 1 = 0$$

$$\Rightarrow F_K\left(K_{t+1}, L_{t+1}\right) - \delta = r_{t+1} \tag{8.3.2}$$

We have already seen equation (8.3.2) before. It's identical to equation (4.4.12) in Chapter 4. There we were asking the question the other way around: given a level of investment, what must the interest rate be? Here we are asking: given an interest rate, what will be the level of the capital stock? The level of the capital stock must be such that the rental rate of capital makes the NPV of the representative investment project equal to zero.

Formula (8.3.2) is stated in terms of the *level* of the capital stock. In order to know the level of *investment*, we use that:

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

and replace K_{t+1} to get:

$$F_K\left((1-\delta)\, K_t + I_t, L_{t+1}\right) - \delta = r_{t+1} \tag{8.3.3}$$

Figure 8.3.1 shows equation (8.3.3) graphically. The left-hand side of the equation is a downward-sloping curve, which inherits the shape of the marginal product of capital curve. Sometimes this is known as an "investment demand" schedule, meaning that it tells us how much investment would be carried out at each possible level of interest rates. The right hand side is represented with a horizontal line since we are taking interest rates as given.

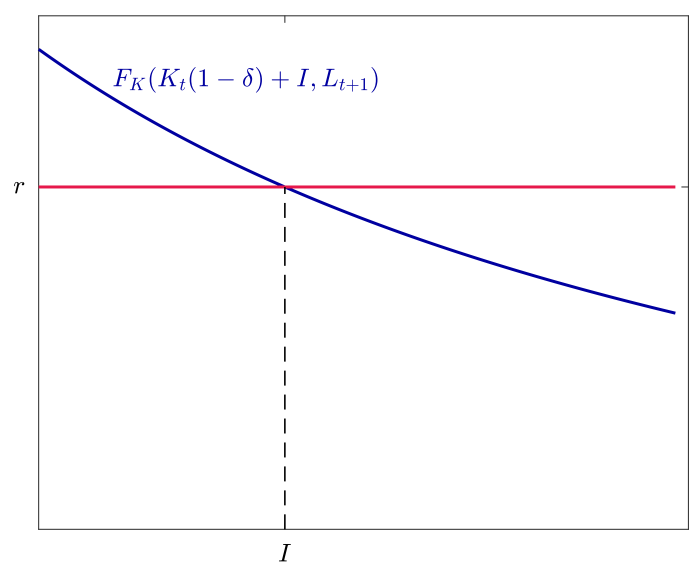


Fig. 8.3.1: *The determination of aggregate investment.*

We can use equation (8.3.3), either graphically or algebraically, to ask how the level of investment responds to different changes. The left panel of Figure 8.3.2 shows how investment responds to an increase in expected productivity, represented by an upward shift in the F_K curve. At the original level of investment, higher productivity would make the representative investment project positive-NPV. This encourages additional investment until the decreasing marginal product of capital ensures that NPV equals zero again. Hence, investment rises. The right panel shows how investment responds to a rise in the interest rate. At the original level of investment, a higher interest rate would make the representative investment project negative-NPV. This leads to a fall in investment until the higher marginal product of capital ensures that NPV equals zero again. Hence, investment falls.

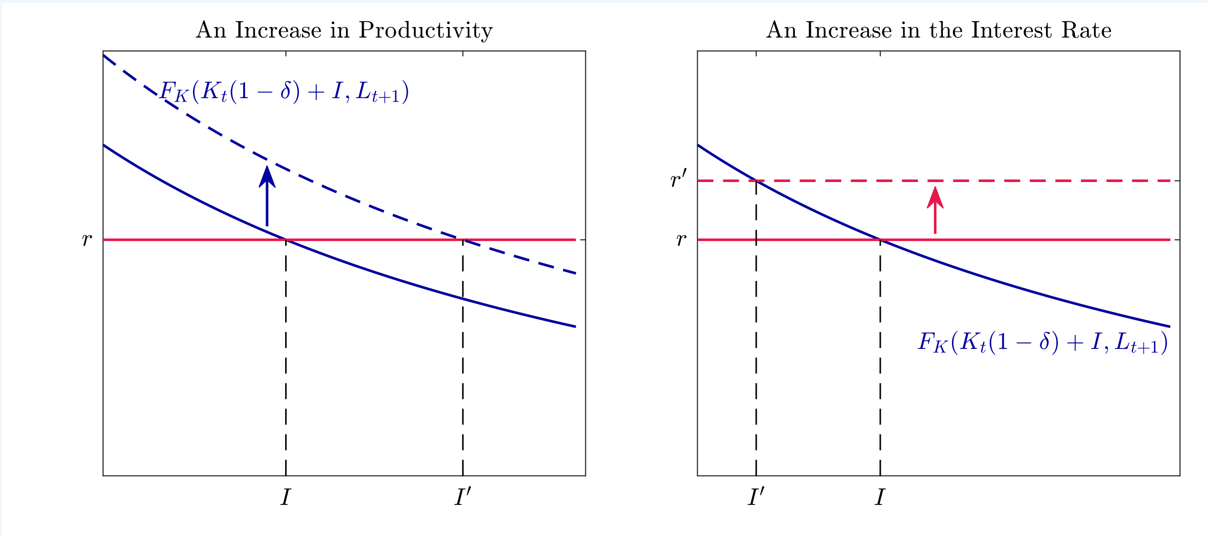


Fig. 8.3.2: *Investment response to changes in productivity and interest rates.*

Exercises

8.1 **Valuation**

Charlie's cheese factory has a very precise business plan for 2019-2028, shown below (you can download it as an Excel spreadsheet from the book website):

Year	Profits	Investment	Dividend
2019	80	50	30
2020	77	49	28
2021	86	49	37
2022	91	47	44
2023	98	47	51
2024	98	48	50
2025	109	200	-91
2026	120	57	63
2027	119	57	62
2028	125	61	64

(in 2025 the main storage facility will need to be replaced, hence the higher investment). From 2029 onwards, dividends will increase at a rate of 2% a year forever. The interest rate is 6%.

(a) Use the Gordon growth formula (8.1.2) to calculate the value that the factory will have in 2028 after paying dividends (i.e. the value not including the value of the dividends it will pay in 2028).

(b) Compute the present value of the entire infinite stream of dividends that starts in 2019.

8.2 **Bond Pricing**

Suppose the interest rate is 4% and is expected to stay at 4% forever. On January 1, 2018, the government issued a T-year bond, which pays coupons of 4 cents every year on January 1 (starting in 2019) and then pays 1 dollar (in addition to the 4-cent coupon) on January 1 of the year $2018 + T$.

(a) Use formula (8.1.1) to compute what the market price of the bond should be. How does it depend on T? Explain.

(b) Suppose that after the bond has been issued, market conditions change and interest rates fall to 3%, and are expected to remain at 3% forever. What is the market price of the bond now? How does it depend on T? Explain.

8.3 **Looking at the Stock Market**

Specific Mills is a company that has just started operations. Its business consists of buying wheat from farmers, grinding it to produce flour and selling the flour. It is very small relative to both the wheat market and the flour market. Its assets consist of a single plant that cost $10 million to build. It issued 1 million shares, which currently trade in the stock market at a price of $10 per share.

(a) What is Q for Specific Mills?

(b) Specific Mills' share price suddenly rises to $15 per share. What is Q now?

Management is considering expanding the plant. It has calculated that in order to expand the plant to λ times its current capacity, it is going to have to carry out additional investment that will cost

$$(\lambda - 1) + \psi(\lambda - 1)^2$$

times its original investment, with $\psi = 1$. Assume that management trusts the stock market investors' opinion about the profitability of the grain-processing industry.

(c) By how much should it expand its capacity?

(d) How much will it spend in total to do so?

(e) Now suppose that instead of Specific Mills, which is small relative to the industry it operates in, the same situation arose for a company that was large relative to its market. How does its reaction to its stock market price compare to that of Specific Mills? Why?

8.4 **The Magnitude of Adjustment Costs**

The slope of the relationship illustrated in Figure 8.1.1 is approximately 0.01, so an increase in Q of 1 is associated with an increase in the investment-to-capital ratio $\frac{I}{K}$ of 0.01.

(a) If adjustment costs are given by formula (8.1.4), what value of ψ would be consistent with this observation?

(b) If a firm's investment is equal to 10% of its capital stock, how much does it have to spend in adjustment costs as a fraction of its total investment? How about if a firm's investment is 20% of its capital?

8.5 **A Risky Asset**

The world lasts two periods. Danube.com is a large internet-based retailer. At $t = 2$, it will pay a dividend of \$150 per share if the economy is doing well and \$50 per share if the economy is doing poorly. The representative household will consume \$40,000 if the economy is doing well and \$30,000 if the economy is doing poorly. Its preferences are given by $\mathbb{E}\left(\frac{c^{1-\sigma}}{1-\sigma}\right)$. The interest rate is 10%. Use formula (8.2.5) to determine at what price the representative household would be indifferent with respect to buying shares in Danube.com, for the following values of σ: 0.5,1,2 and 10. Explain.

8.6 **A Risky Investment**

Ingrid plans to live for two periods and only cares about consumption in the second period. She is considering getting her MBA in the first period. If she does it, it's going to cost her \$200,000 between tuition and foregone wages. Once she graduates, there is a 50% chance that in the second period she will get the job she wants, which will pay her \$5,000,000 and an 50% chance she'll get a regular job that will pay her \$80,000. If she doesn't get her MBA, she'll put \$200,000 in the bank, where it will earn 10% interest, and then get a regular job that pays her \$80,000. In all cases, since she will only live for two periods, she will consume everything she has in period 2. Suppose Ingrid's preferences over period-2 consumption are given by $\mathbb{E}\left(\frac{c^{1-\sigma}}{1-\sigma}\right)$.

(a) For what values of σ is it a good idea to get an MBA? Explain.

(b) Now suppose $\sigma = 1.2$. For what values of the interest rate is it a good idea to get an MBA? Explain.

8.7 **Credit Constraints**

Two world lasts two periods. A firm's production function is $F(K, L) = A_t K^\alpha L^{1-\alpha}$, where A_t can take different values in each period. In each period, the firm can hire labor in a competitive labor market at the same wage w, which the firm takes as given. However, there is no market for renting capital: the firm can only use capital that it owns. If the firm owns K units of capital and decides to hire L workers, then it earns $F(K, L) - wL$. The firm starts off having K_1 units of capital in the first period. There is no depreciation. At the end of period 1, the firm can buy capital for period 2 by either using its earnings or by borrowing. Loans must be paid back in period 2. The interest rate is r. The firm can also use its earnings to pay dividends to its shareholders in period 1. The maximum amount that lenders are willing to lend to the firm is b.

(a) Set up the problem of the firm that needs to decide how many workers to hire in each period. Note that this problem can be solved period-by-period taking K_t as given. Find an expression for the period-t profits of a firm that takes as given its capital stock K_t and chooses how much labor to hire, i.e. for $\pi(K_t) \equiv \max_L F(K_t, L) - wL$.

(b) Set up the problem of the firm that must decide how much to pay in dividends, how much to borrow and how much to invest.

(c) Assume $b = \infty$. Show that the optimal amount of investment depends on A_2 but not on A_1. Explain.

(d) Assume $b = 0$. Show that if A_1 is sufficiently high, then the optimal amount of investment is the same as in part (c). Find the minimum level of A_1 such that this is the case, and denote it A_1^*. Show that for $A_1 < A_1^*$, investment depends on A_1. Explain.

(e) Suppose $A_1 < A_1^*$. How does the firm react to an increase in A_2? How does the firm react to an increase in b? Explain.

8.8 **An Earthquake**

Suppose an earthquake destroys a large part of the capital stock at time t. Assume interest rates and future labor supply are not affected by the earthquake, and there are no adjustment costs.

(a) What will happen to investment?

(b) How does K_{t+1} compare with and without the earthquake?

8.9 **Aggregate Investment with Risk**

The world lasts two periods. The aggregate production technology for period 2 is given by:

$$F(K, L) = AK^{\alpha}L^{1-\alpha}$$

K is the aggregate capital stock in period 2 and L is the labor force, which is exogenous and normalized to $L = 1$. Period 2 is the end of the world, so capital depreciates fully ($\delta = 1$) and the representative household will consume $F(K, L) = AK^{\alpha}L^{1-\alpha}$. The utility function is $u(c) = \log(c)$. A is a random variable, which can take two possible values: $A_H = 1 + \varepsilon$ or $A_L = 1 - \varepsilon$, with equal probability. The interest rate between periods 1 and 2 is r.

Now consider the following investment project: investing an additional unit in period 1 to obtain an additional unit of capital in period 2.

(a) What is the dividend produced by the marginal unit of capital? Express it as a function of the aggregate capital stock K and realized productivity A.

(b) Suppose $\varepsilon = 0$. For what value of K is the net present value of additional investment exactly zero?

(c) Now suppose $\varepsilon > 0$. For what value of K is the net present value of additional investment exactly zero? (Use equation (8.2.5) to guide your answer)

(d) How does K depend on ε? Explain.

CHAPTER 9

General Equilibrium

In Chapters 6-8 we have studied the decisions of households and firms in isolation. In this chapter we look at how they all fit together.

In microeconomics we say that there's an equilibrium in a competitive market for some good if supply equals demand: everyone buys or sells as much as they want and the outcome is that sales equal purchases. General equilibrium is the same idea but applied to many goods at once.

The model economy we'll be looking at will have:

- a representative household making consumption-savings-leisure decisions
- a representative firm choosing to hire labor and rent capital
- a representative investment firm carrying out investment

We'll first look at this in a simplified two-period model and then we'll extend the analysis to an infinite-period case.

9.1 General Equilibrium in a Two-Period Economy

We are going to assume that the economy lasts only for two periods. In the first period, there is already some level of capital K_1 in place, and it's owned by the representative household. Each period, the production function is $F(K, L)$ and there are competitive markets for labor and renting capital. The initial capital stock depreciates at rate δ between period 1 and period 2, and it's possible to invest in order to build more capital. After period 2 it's the end of the world, so capital depreciates completely $(\delta_2 = 1)$ and there is no more investment.

Household, Firm and Investment Firm Problems

The representative household must choose how much to consume and also how much of its time to dedicate to leisure and consumption. It solves a problem like the one we looked at in Section 7.4:

$$\max_{c_1,l_1,c_2,l_2} u(c_1) + v(l_1) + \beta\left[u(c_2) + v(l_2)\right]$$
$$s.t. \tag{9.1.1}$$
$$c_1 + \frac{1}{1+r}c_2 \leq w_1(1-l_1) + \frac{1}{1+r}w_2(1-l_2) + K_1\left(1 + r_1^K - \delta\right) + \Pi$$

The only thing that is different between this problem and the one we know from Section 7.4 is that the household starts with some initial wealth: it owns the initial capital stock and earns a rental (net of depreciation) for it in the first period. Also, the household owns the firms, so if they were to make profits, the present value of those profits, which we denote by Π, would be part of the household's budget. Π is given by $\Pi = \Pi_1^F + \frac{\Pi_2^F}{1+r} + \Pi^I$, where Π_t^F are the profits of productive firms in period t and Π^I are the profits of investment firms. In equilibrium, it will be the case that $\Pi = 0$, so we can just ignore this part.

There is a representative firm. In each period, the firm solves the profit maximization problem we saw in Chapter 4:

$$\Pi_t^F = \max_{K,L} F(K,L) - w_t L - r_t^K K$$

Investment consists of building period-2 capital. We are going to imagine that there is a representative investment firm that carries out all the investment. This firm buys the $(1-\delta)K_1$ units of used period-1 capital, adds I units of investment and obtains $K_2 = (1-\delta)K_1 + I$ units of period-2 capital (here we are assuming there are no adjustment costs like the ones we had in Section 8.1). The investment firms then rents out the K_2 units of capital at the rental rate r_2^K (given that the second period is the end of the world, they *fully* depreciate in period 2). Converted back to present value, this means the investment firm earns $\frac{r_2^K}{1+r}K_2$ from renting out the capital. The problem of the investment firm is:

$$\Pi^I = \max_I \underbrace{\frac{r_2^K}{1+r}\left[(1-\delta)K_1 + I\right]}_{\text{Present Value of } K_2 \text{ rental}} - \underbrace{\left[(1-\delta)K_1 + I\right]}_{\text{Cost of } K_2} \tag{9.1.2}$$

Equilibrium Definition

Definition 9.1.
Given an initial K_1, a competitive equilibrium consists of:

1. An allocation $\{c_1, l_1, c_2, l_2, I, K_2, L_1, L_2\}$.
2. Prices $\left\{w_1, w_2, r_1^K, r_2^K, r\right\}$.

such that:

1. $\{c_1, l_1, c_2, l_2\}$ solves the household's problem, taking prices as given.
2. L_t, K_t solve the firm's problem for $t = 1, 2$, taking prices as given.

3. I solves the investment firm's problem, taking prices as given.

4. Markets for goods and labor in each period clear:

   (a) Goods:

$$\underbrace{F\left(K_1, L_1\right)}_{\text{GDP}} = \underbrace{c_1}_{\text{Consumption}} + \underbrace{I}_{\text{Investment}} \tag{9.1.3}$$

$$\underbrace{F\left(K_2, L_2\right)}_{\text{GDP}} = \underbrace{c_2}_{\text{Consumption}} \tag{9.1.4}$$

   (b) Capital:

$$K_2 = K_1(1-\delta) + I \tag{9.1.5}$$

   (c) Labor:

$$L_t + l_t = 1 \tag{9.1.6}$$

A competitive equilibrium satisfies two basic properties:

- Everyone is, individually, making the best choices they can. This is represented by conditions 1-3, which say that each individual household and each individual firm makes its choices in their own best interest.

- Things add up, i.e. everyone's choices are consistent with everyone else's choices. This is represented in condition 4. Condition 4(a) says that all the output in the economy is used for either consumption or investment in the first period and, since it's the end of the world, only for consumption in the second period. (Recall that this is a closed economy with no government, so there is no other use for output). Condition 4(b) says that the capital that firms want to use in period 2 is equal to the amount of period-1 capital that remains plus the amount that investment firms chose to build. Condition 4(c) says that the labor that firms choose to hire (L_t) plus the amount of time the households choose to dedicate to leisure (l_t) add up to the entire amount of time available, which is normalized to 1.

Describing an Equilibrium

In this section we'll find a system of equations whose solution represents the economy's competitive equilibrium. For now we are going to leave it as a mathematical expression and the economics it contains might be a little hard to discern. We'll use these equations to think more about economics in later chapters.

We know from Section 7.4 that the first order conditions are (7.4.2) and (7.4.3), which we just restate here:

$$\frac{v'\left(l_t\right)}{u'\left(c_t\right)} = w_t \tag{9.1.7}$$

$$u'\left(c_1\right) = \beta\left(1+r\right) u'\left(c_2\right) \tag{9.1.8}$$

Equation (9.1.7) is the same as (7.4.2) and it describes how the household trades off leisure and consumption. Equation (9.1.8) is the same as (6.3.9) and (7.4.3) and it describes how the household trades off present and future consumption.

As we saw in Chapter 4, the solution to the firm's problem can be summarized by the first-order conditions:

$$F_K(K_t, L_t) - r_t^K = 0 \tag{9.1.9}$$

$$F_L(K_t, L_t) - w_t = 0 \tag{9.1.10}$$

(which are the same as (4.4.1) and (4.4.2)).

The investment firm's problem (9.1.2) is linear in I. This means that unless $1 + r = r_2^K$, investment firms would be able to make infinite profits by choosing either $I = \infty$ or $I = -\infty$ depending on which way the inequality goes. This would be inconsistent with capital-market clearing. Therefore it must be that in equilibrium:

$$1 + r = r_2^K \tag{9.1.11}$$

which is just equation (4.4.12) when depreciation is set to 1. Recalling the definition of net present value from Chapter 8, this says that in equilibrium the NPV of investment must be zero.

Replacing (9.1.10) into (9.1.7) we obtain:

$$\underbrace{\frac{v'(l_t)}{u'(c_t)}}_{\text{Marginal Rate of Substitution}} = \underbrace{F_L(K_t, L_t)}_{\text{Marginal Rate of Transformation}} \tag{9.1.12}$$

Equation (9.1.12) summarizes how this economy will allocate the use of time. On the left hand side, the expression $\frac{v'(l_t)}{u'(c_t)}$ describes how the representative household is willing to trade off leisure against consumption. On the right hand side, $F_L(K_t, L_t)$ describes how the available technology is able (at the margin) to convert time into output.

Replacing (9.1.11) and (9.1.9) into (9.1.8) and we obtain:

$$\underbrace{\frac{u'(c_1)}{\beta u'(c_2)}}_{\text{Marginal Rate of Substitution}} = \underbrace{F_K(K_2, L_2)}_{\text{Marginal Rate of Transformation}} \tag{9.1.13}$$

Equation (9.1.13) summarizes how this economy will allocate output between the present and the future. On the left hand side, the expression $\frac{u'(c_1)}{\beta u'(c_2)}$ describes how the representative household is willing to trade off present consumption against future consumption. On the right hand side, $F_K(K_2, L_2)$ describes how the available technology is able, by building a marginal unit of capital, to convert current output into future output.

Equations (9.1.12) and (9.1.13) summarize what the general equilibrium will look like. We'll come back to them many times.

9.2 The First Welfare Theorem

A competitive economy will allocate resources in one specific way. What if we could choose a different allocation? Would that be a good idea?

In order to answer this question it is convenient to invoke the metaphor of a "social planner". The idea is to imagine that, instead of making their decisions individually, everyone delegates decisions to a benevolent social planner. Is it the case that the social planner would want to change the allocation of resources to something other than the competitive equilibrium? We'll show that, under the assumptions we've made so far, the answer is "no". Even a social planner that was perfectly benevolent and had no practical, political or cognitive difficulties in choosing among all the possible allocations of resources would be satisfied with the outcome of competitive markets.

The fictitious social planner solves the following problem:

$$\begin{gathered}\max_{c_1,l_1,c_2,l_2,L_1,L_2,K_2} u\left(c_1\right)+v\left(l_1\right)+\beta\left[u\left(c_2\right)+v\left(l_2\right)\right]\\ s.t.\\ K_2 \leq (1-\delta) K_1+F\left(K_1, L_1\right)-c_1\\ c_2 \leq F\left(K_2, L_2\right)\\ L_t \leq 1-l_t \quad \text{for } t=1,2\\ K_1 \text{ given}\end{gathered} \tag{9.2.1}$$

What does this optimization problem represent, in economic terms? First, the planner indeed is benevolent. Its objective function is the same as the objective of the representative household: the planner wants to make the representative household happy. Second, the planner is quite powerful: it can tell everyone exactly how much to work, consume and invest without worrying about whether they will obey the instructions. In that regard, our fictitious social planner is much more powerful than a government could possibly be. Lastly, the planner is not all-powerful: it is constrained by the technological possibilities of the economy. The constraints, which are identical to conditions (9.1.3)-(9.1.6), say that in order to accumulate capital it is necessary to give up consumption and in order to produce output it is necessary to use labor, which requires giving up leisure. These constraints are not exactly budget constraints since the planner is not buying or selling from anyone; rather, they are technological constraints.

Let's replace the constraints in the objective function and then solve the planner's problem. Replace $c_1=(1-\delta) K_1+F\left(K_1, L_1\right)-K_2$, $c_2=F\left(K_2, L_2\right)$ and $l_t=1-L_t$ into the objective to get:[1]

$$\max_{L_1,L_2,K_2} u\left((1-\delta) K_1+F\left(K_1, L_1\right)-K_2\right)+v\left(1-L_1\right)+\beta\left[u\left(F\left(K_2, L_2\right)\right)+v\left(1-L_2\right)\right]$$

The first order conditions with respect to L_1, L_2 and K_2 respectively are:

$$\begin{aligned} u'\left((1-\delta) K_1+F\left(K_1, L_1\right)-K_2\right) F_L\left(K_1, L_1\right)-v'\left(1-L_1\right)=0\\ \beta\left[u'\left(F\left(K_2, l_2\right)\right) F_L\left(K_2, L_2\right)-v'\left(1-L_1\right)\right]=0\\ -u'\left((1-\delta) K_1+F\left(K_1, L_1\right)-K_2\right)+\beta u'\left(F\left(K_2, l_2\right)\right) F_K\left(K_2, L_2\right)=0\end{aligned}$$

which simplify back to (9.1.12) and (9.1.13), which are the equations that describe the competitive equilibrium.

[1] Alternatively, we can set up a Lagrangian with all the constraints.

Therefore the equations that define the solution to the social planner's problem are the same as those which define the competitive equilibrium! This is not a coincidence. The following result proves that the allocations that result from a competitive equilibrium must achieve the optimum in the planner's problem.

Proposition 9.1 (First Welfare Theorem)**.** If an allocation $\{c_1, l_1, c_2, l_2, I, K_2, L_1, L_2\}$ is part of a competitive equilibrium then it solves the social planner's problem.

Proof. Proceed by contradiction. Suppose that an allocation $\{c_1, l_1, c_2, l_2, I, K_2, L_1, L_2\}$ with prices $\{w_1, w_2, r_1^K, r_2^K, r\}$ is a competitive equilibrium but there is another allocation $\left\{\hat{c}_1, \hat{l}_1, \hat{c}_2, \ \hat{l}_2, \hat{I}, \hat{K}_2, \hat{L}_1, \hat{L}_2\right\}$ that satisfies the constraints on the social planner's problem but achieves strictly higher utility for the household. This implies that:

$$\hat{c}_1 + \frac{1}{1+r}\hat{c}_2 > w_1\left(1-\hat{l}_1\right) + \frac{1}{1+r}w_2\left(1-\hat{l}_2\right) + K_1\left(1+r_1^K-\delta\right) + \Pi \tag{9.2.2}$$

Equation (9.2.2) says, if prices are the equilibrium prices, then the household cannot afford the consumption-leisure combination $\left\{\hat{c}_1, \hat{l}_1, \hat{c}_2, \hat{l}_2\right\}$. Otherwise, since it's presumed to be strictly better, the household would have chosen it. Furthermore, the fact that in equilibrium firms and investment firms are maximizing profits implies that by choosing the equilibrium quantities they make at least as much profit as they would by choosing the alternative quantities:

$$\begin{aligned}
\Pi_1^F &= F\left(K_1, L_1\right) - w_1 L_1 - r_1^K K_1 \geq F\left(K_1, \hat{L}_1\right) - w_1\hat{L}_1 - r_1^K K_1 \\
\Pi_2^F &= F\left(K_2, L_2\right) - w_2 L_2 - r_2^K K_2 \geq F\left(\hat{K}_2, \hat{L}_1\right) - w_2\hat{L}_1 - r_2^K \hat{K}_2 \\
\Pi^I &= \left(\frac{r_2^K}{1+r} - 1\right)\left[(1-\delta)K_1 + \hat{I}\right] \geq \left(\frac{r_2^K}{1+r} - 1\right)\left[(1-\delta)K_1 + \hat{I}\right]
\end{aligned}$$

Therefore, using the definition:

$$\Pi \equiv \Pi_1^F + \frac{\Pi_2^F}{1+r} + \Pi^I$$

to add the profits of production firms and investment firms, we have that:

$$\Pi \geq F\left(K_1, \hat{L}_1\right) - w_1\hat{L}_1 - r_1^K K_1 + \frac{F\left(\hat{K}_2, \hat{L}_1\right) - w_2\hat{L}_1 - r_2^K\hat{K}_2}{1+r} + \left(\frac{r_2^K}{1+r} - 1\right)\left[(1-\delta)K_1 + \hat{I}\right] \tag{9.2.3}$$

Replacing (9.2.3), $\hat{L}_t \leq 1 - \hat{l}_t$ and $\hat{I} = K_2 - (1-\delta)K_1$ into (9.2.2) and simplifying, we obtain:

$$\hat{c}_1 + K_2 + \frac{1}{1+r}\hat{c}_2 > F\left(K_1, \hat{L}_1\right) + (1-\delta)K_1 \frac{F\left(\hat{K}_2, \hat{L}_1\right)}{1+r}$$

This implies that one of the following statements must be true. Either:

$$\hat{c}_1 + K_2 > F\left(K_1, \hat{L}_1\right) + (1-\delta)K_1$$

or

$$\hat{c}_2 > F\left(\hat{K}_2, \hat{L}_1\right)$$

either of which contradicts the assumption that $\left\{\hat{c}_1, \hat{l}_1, \hat{c}_2, \hat{l}_2, \hat{I}, \hat{K}_2, \hat{L}_1, \hat{L}_2\right\}$ satisfies the constraints on the planner's problem. □

What is the economic logic behind the First Welfare Theorem? Why are the social planner's choices the same as those the market produces? The social planner chooses an allocation to maximize utility subject to technological possibilities. Conversely, in a competitive economy, each household maximizes utility subject to prices (i.e. wages, the rental rate of capital and the interest rate), as equations (9.1.7) and (9.1.8) indicate; however, those prices in turn reflect technological possibilities, as equations (9.1.9) and (9.1.10) indicate. Therefore the household is also, indirectly, maximizing utility subject to technological possibilities.

What we have shown is really just a special case of the First Welfare Theorem. The result is much more general. In particular, it is still true if:

- There are many different goods at each date (e.g. apples, oranges, etc.) instead of just a general consumption good.
- There is (exogenous) technological progress.
- There is uncertainty, as long as there are "complete markets", i.e. markets to trade insurance against every possible state of the world, at competitive prices.
- There are many different households with different preferences, different abilities and different wealth instead of a representative household. However, for this case the theorem needs to be stated slightly differently. If an economy has different households, there is no unique way to define the social planner's problem because the planner would have to decide how much the utility of each different household matters. Hence, the more general version of the FWT says that if an allocation is part of a competitive equilibrium then it is Pareto optimal: there is no technologically feasible way to make someone better off without making someone else worse off.

The FWT is an extremely useful guide for thinking about public policy. In an economy where the conditions for the theorem hold, then *any* policy that makes a household better off, relative to the competitive equilibrium, necessarily makes someone else worse off. Does this mean that no policy is ever justified? Some people interpret the theorem to imply just that, but this conclusion requires an extra bit of political philosophy. Depending on one's views on the nature of justice, it is quite possible to advocate for a policy that benefits some groups of people at the expense of others (for instance, policies that benefit the poor at the expense of the rich or the old at the expense of the young). What the FWT does clarify is that, if the economy is competitive, the *only* possible economic justification for a policy is that one views the resulting redistribution as desirable.

The conditions for the FWT are quite strict and no one believes that they hold exactly in practice. Two of the main things that would make the FWT *not* hold are:[2]

[2] Other features that would make the FWT fail include asymmetric information, incomplete markets and borrowing constraints.

- Monopoly power. Monopolists reduce quantity, relative to a competitive producer, in order to raise prices. Therefore it is not true that they equates the value of the marginal product to the factor prices, which is the key step that leads to equations (9.1.12) and (9.1.13). Hence, the FWT does not hold. In Chapter 14 we study a model that features monopoly power.

- Externalities. In an economy with externalities, the private value and the social value of a good do not coincide. If I decide to hire a (good) orchestra to play in my garden and all my neighbors enjoy it, the social value this produces exceeds the private value that I obtain. When I decide whether or not to hire the orchestra, I will ignore the benefit to my neighbors while a benevolent social planner would take it into account, so I will hire an orchestra less often than what the social planner would want. Similarly, I will over-pollute, under-invest in preventing communicable diseases, etc.

Again, the FWT helps to organize the analysis of public policy. If a policy is not about redistribution, then it can only be justified economically on the basis of which failure of the FWT it's designed to address.[3] For instance, zoning regulations can (perhaps) be justified as a way to deal with externalities: if I could build a tall building next to your house, I would make your garden less sunny. Antitrust policies can (perhaps) be justified as a way to deal with monopoly power. It is often useful to begin thinking about a policy problem by asking what is the failure of the FWT that the policy is designed to address.

9.3 General Equilibrium in an Infinite-Period Economy

Let's now look at what happens in an economy that lasts infinite periods. We'll proceed in the same way as before: describe the household's problem, the firm's problem, the investment decision and finally define and describe an equilibrium.

The representative household solves the following problem:

$$\max_{c_t, l_t, a_{t+1}} \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t\right) + v\left(l_t\right) \right] \tag{9.3.1}$$

$$s.t.$$

$$a_{t+1} = \left(1 + r_t\right) a_t + w_t \left(1 - l_t\right) + \Pi_t^F + \Pi_t^I - c_t \tag{9.3.2}$$

$$a_0 = K_0 \left(1 + r_0^K - \delta\right) \text{ given}$$

$$\text{a no-Ponzi-game condition like (6.3.7)}$$

In each period, the household obtains utility from both consumption c_t and leisure l_t. Equation (9.3.2) is the budget constraint; more precisely, there is an infinite number of budget constraints, each on them linking one period to the following one. Π_t^F and Π_t^I represent the profits of the productive and investment firms respectively, which will turn out to be zero.

[3] There's a gray area with policies that are justified on the basis that people make the wrong choices. Should a benevolent social planner want to give people what they themselves would choose or what the social planner "knows" is best for them? This issue famously comes up in discussions of drug policy but also, for instance, in financial regulation. The answer involves a philosophical discussion that economists usually don't specialize in. Thaler and Sunstein (2008) discuss many policy issues related to this question.

The firm's problem is unchanged: it just maximizes profits period-by-period. The investment firm's problem is similar to (9.1.2), except that capital does not fully depreciate in one period, so a firm that invests at time t knows that it will get back depreciated capital at time $t+1$ in addition to a rental. Therefore, the investment firm solves:

$$\Pi_t^I = \max_I \frac{r_{t+1}^K + 1 - \delta}{1 + r_{t+1}} \left[K_t(1-\delta) + I\right] - \left[K_t(1-\delta) + I\right] \tag{9.3.3}$$

Definition 9.2.
A competitive equilibrium consists of:

1. An allocation $\{c_t, l_t, I_t, K_{t+1}, L_t\}_{t=0}^{\infty}$.
2. Prices $\left\{w_t, r_t^K, r_t\right\}_{t=0}^{\infty}$.

such that:

1. $\{c_t, l_t\}_{t=0}^{\infty}$ solves the household's problem, taking prices as given.
2. L_t, K_t solve the firm's problem for every t, taking prices as given.
3. I_t solves the investment firm's problem for every t, taking prices as given.
4. Markets for goods and labor in each period clear:
   (a) Goods:
   $$\underbrace{F\left(K_t, L_t\right)}_{\text{GDP}} = \underbrace{c_t}_{\text{Consumption}} + \underbrace{I_t}_{\text{Investment}} \tag{9.3.4}$$
   (b) Capital:
   $$K_{t+1} = K_t(1-\delta) + I_t \tag{9.3.5}$$
   (c) Labor:
   $$L_t + l_t = 1 \tag{9.3.6}$$

As you can see, there isn't much conceptual difference between how we define an equilibrium in a two-period model and in an infinite-period model. In fact, the description of how the equilibrium behaves is also very similar. The first order conditions for the household's problem are:

$$\frac{v'\left(l_t\right)}{u'\left(c_t\right)} = w_t \tag{9.3.7}$$

$$u'\left(c_t\right) = \beta\left(1 + r_{t+1}\right) u'\left(c_{t+1}\right) \tag{9.3.8}$$

For the firms problem:

$$F_K\left(K_t, L_t\right) - r_t^K = 0 \tag{9.3.9}$$

$$F_L(K_t, L_t) - w_t = 0 \tag{9.3.10}$$

And setting the NPV of investment to zero to prevent investment firms from choosing infinite investment:

$$r_{t+1} = r^K_{t+1} - \delta \tag{9.3.11}$$

which is just equation (4.4.12).

Replacing (9.3.10) into (9.3.7) we obtain:

$$\underbrace{\frac{v'(l_t)}{u'(c_t)}}_{\text{Marginal Rate of Substitution}} = \underbrace{F_L(K_t, L_t)}_{\text{Marginal Rate of Transformation}} \tag{9.3.12}$$

and replacing (9.3.11) and (9.3.9) into (9.3.8) and we obtain:

$$\underbrace{\frac{u'(c_t)}{\beta u'(c_{t+1})}}_{\text{Marginal Rate of Substitution}} = \underbrace{(1 + F_K(K_{t+1}, L_{t+1}) - \delta)}_{\text{Marginal Rate of Transformation}} \tag{9.3.13}$$

All of these equations have the same interpretation as in the two-period economy. The infinite-horizon model will just let us ask some additional questions, such as "what will the economy look like in the long run?" or "how does the economy react today to news about things that will happen in the distant future?"

Dynamics

When we studied the Solow model in Chapter 4, we concluded that an economy with a constant savings rate and a fixed labor supply would converge towards a steady state, where the capital stock and output were constant and investment was just enough to make up for depreciation. Now that we have a theory of what determines the savings rate, we can ask the same questions again: what will the economy look like in the long run? How will it behave in the meantime?

To keep things relatively simple and focus just on the consumption / investment problem, we are going to go back to the assumption that the labor supply is fixed, so instead of equation (9.1.12) we'll just have $L_t = 1$. Furthermore, we'll assume that the representative household has the utility function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, which we first introduced in Chapter 2. With this utility function and setting $L_t = 1$, equation (9.3.13) reduces to:

$$\frac{c_{t+1}}{c_t} = [\beta(1 + F_K(K_{t+1}, 1) - \delta)]^{\frac{1}{\sigma}} \tag{9.3.14}$$

Equation (9.3.14) gives us a relationship between the *rate of growth* of consumption between t and $t+1$ and the *level* of the capital stock at time $t+1$. If the capital stock is *low,* consumption should be *growing* over time. What is the economic logic that it represents? Suppose, for instance, that the capital stock at $t+1$ is *low.* This implies that the marginal product of capital is *high*, due to diminishing marginal product; therefore the rental rate of capital will be *high*, by equation (9.1.9); therefore the interest rate will be *high*, by equation (9.1.11); therefore current consumption is expensive relative to future consumption, as we saw in Chapter 6;

therefore the household chooses to consume more in the *future* than in the *present*, by the Euler equation (9.1.8), which is exactly what a high rate of growth of consumption means.

In addition, we have the market clearing condition (9.1.3), which we can rewrite as:

$$K_{t+1} = (1-\delta)\, K_t + F\,(K_t, 1) - c_t \tag{9.3.15}$$

This gives us a relationship between the future capital stock, the current capital stock, and consumption. The economic logic here is simple: more consumption implies less investment and therefore a lower capital stock in the following period.

We can represent equations (9.3.14) and (9.3.15) by means of a phase diagram. This is a graph that shows us, for each possible level of c and k, in what direction c and k are supposed to be moving if they are to satisfy (9.3.14) and (9.3.15). This is shown in Figure 9.3.1.

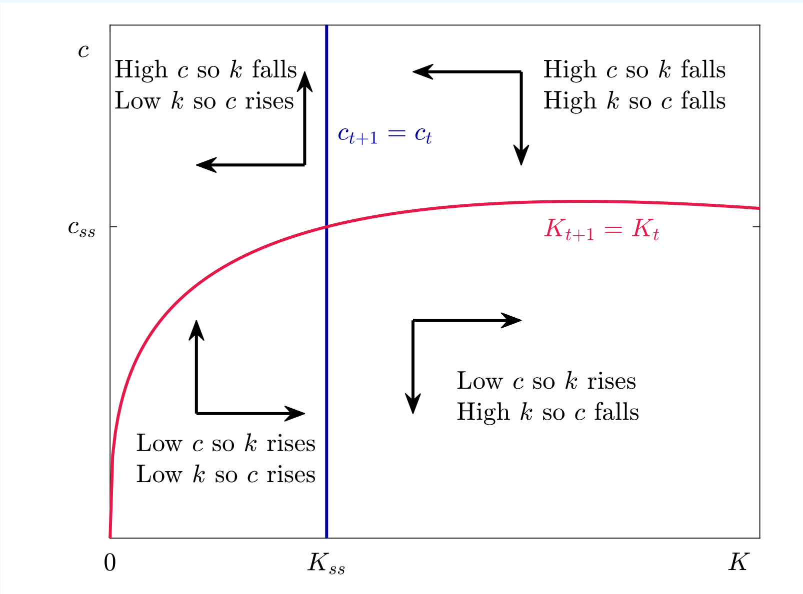


***Fig. 9.3.1:** Dynamics of the Neoclassical Growth Model*

The vertical line represents equation:

$$\frac{c_{t+1}}{c_t} = 1$$

$$\Rightarrow \left[\beta\,(1 + F_K\,(K_{t+1}, 1) - \delta)\right]^{\frac{1}{\sigma}} = 1 \tag{9.3.16}$$

$$\Rightarrow F_K\,(K_{t+1}, 1) - \delta = \frac{1}{\beta} - 1 \tag{9.3.17}$$

so it tells us what level of capital is consistent with constant consumption. At any point to the left, the lower capital stock means a higher interest rate so consumption must be growing; at any point to the right, the higher capital stock means a lower interest rate so consumption must be falling. Only if K solves equation (9.3.17) is the interest rate exactly $\frac{1}{\beta} - 1$, which persuades the household to keep consumption constant over time.

The curved line represents the equation:

$$\begin{aligned} K_{t+1} &= K_t \\ \Rightarrow (1-\delta) K_t + F(K_t, 1) - c_t &= K_t \\ \Rightarrow c_t &= F(K_t, 1) - \delta K_t \end{aligned}$$

so it tells us, for each level of K, how much that household needs to consume in order to invest enough to exactly make up for depreciation, thus keeping the capital stock constant. For all the points above the curve, higher consumption implies that depreciation exceeds investment so the capital stock shrinks; for all the points below the curve, lower consumption implies that investment exceeds depreciation and the capital stock grows.

Mathematically, (9.3.14) and (9.3.15) are two difference equations in terms of K_t and c_t. If we knew the initial conditions K_0 and c_0, we could trace out the entire path of both variables over time. The initial condition for K_0 is easy. We assumed it's exogenous so we just take as given its initial value. How about c_0? How much will the household consume in the initial period? Figure 9.3.2 shows, for two possible levels of the initial capital stock (K_0 and K_0'), the paths of K_t and c_t that result from different possible values of c_0.

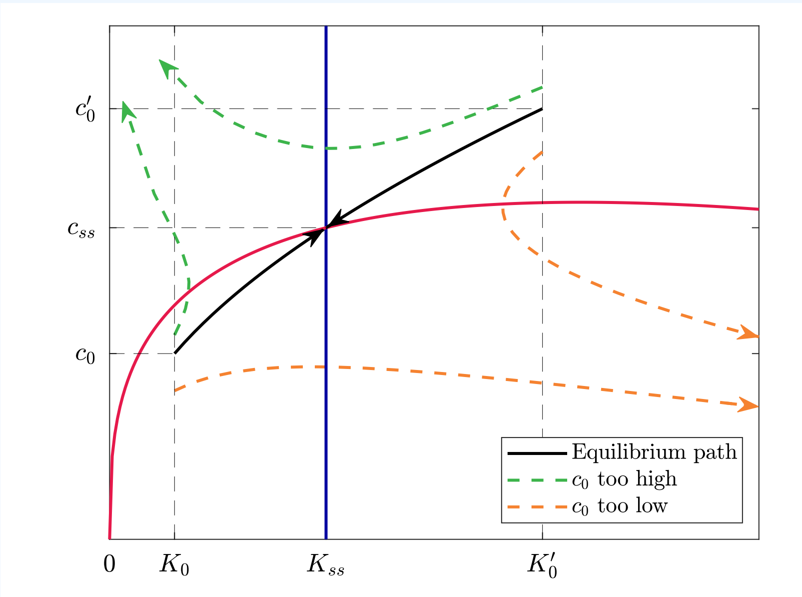


Fig. 9.3.2: *Paths for the economy dictated by equations (9.3.14) and (9.3.15), starting from different levels of c_0 and K_0.*

In each case, the higher path has c_0 too high. Starting from this level of c_0, the Euler equation (9.3.14) dictates path of consumption that becomes ever-increasing, but there is not enough output so the economy starts to deplete the capital stock, eventually depleting it completely. In the lower paths, c_0 is too low. The economy accumulates more and more capital over time and after some time the Euler equation starts to dictate falling consumption. Eventually, all the output is being invested and consumption falls to zero. The only levels of c_0 that are consistent with optimality are the ones that gives rise to the middle paths. Here both c_t and K_t

converge to c_{ss}, K_{ss}. Hence this economy, just like the Solow economy with an exogenous savings rate, has a steady state.

In the theory of difference equations, the paths that lead to the steady state are known as "saddle paths" and the steady state is known as a "saddle point". Starting from any K_0, there is a unique level of c_0 such that the dynamics implied by equations (9.3.14) and (9.3.15) lead towards the steady state. Furthermore, once c_t is on the saddle path, it will stay on the saddle path, so the economy will always be on this path.

The Golden Rule

When we looked at the Solow model we defined a concept called the Golden Rule. This described the level of capital K_{gr} (and the savings rate needed to attain it) such that steady state consumption is maximized. We saw that, depending on the savings rate, an economy could end up with either more or less capital than prescribed by the Golden Rule. Now that we have a theory of the savings rate we can ask how the model predicts that K_{ss} and K_{gr} will compare. Will the economy accumulate more or less capital than prescribed by the Golden Rule? Recall from Chapter 4 that equation (4.3.4) says K_{gr} satisfies:[4]

$$F_K\left(K_{gr}, 1\right) = \delta$$

Instead, with endogenous savings, the steady state level of capital satisfies (9.3.17). Rearranging, we have

$$\begin{aligned}
F_K\left(K_{ss}, 1\right) &= \frac{1}{\beta} - 1 + \delta \\
&> \delta && \text{(as long as the household is impatient so } \beta < 1\text{)} \\
&= F_K\left(K_{gr}, 1\right) \\
\Rightarrow K_{ss} &< K_{gr}
\end{aligned}$$

Therefore this economy will *not* attain the Golden Rule level of capital. It will remain below this.

What do we make of this? The Golden Rule seemed like a pretty desirable outcome, and the First Welfare Theorem says that the equilibrium maximizes the household's utility: why doesn't the social planner implement the Golden Rule? The answer comes from the household's impatience. The Golden Rule maximizes consumption *in the long run*. An impatient household cares about the short run as well as the long run. It would rather consume a little bit more in the present even if it means a lower level of consumption later. Note that mathematically, $K_{ss} \to K_{gr}$ if $\beta \to 1$, so as households become very patient the economy indeed comes closer to the Golden Rule. Note also that the argument does not depend on starting with a low level of capital. If the economy were to start at K_{gr}, the household would choose to invest less than required to maintain the capital stock, consuming more than c_{gr} in the short run at the expense of lower consumption later.

Figure (9.3.3) shows how the steady state compares to the Golden Rule. The graph shows, for each value of K, the level of consumption that is consistent with maintaining a constant capital stock equal to K, i.e. $c = F(K, 1) - \delta K$. If the economy were to maintain K_{gr}, then it could sustain a level of consumption $c_{gr} > c_{ss}$. However, attaining and maintaining such a high capital stock requires sacrificing too much present

[4] In Chapter 4 we had population growth. Here we are setting $n = 0$ but the argument works regardless.

consumption for future consumption and the household is better off with the equilibrium that converges to K_{ss}.

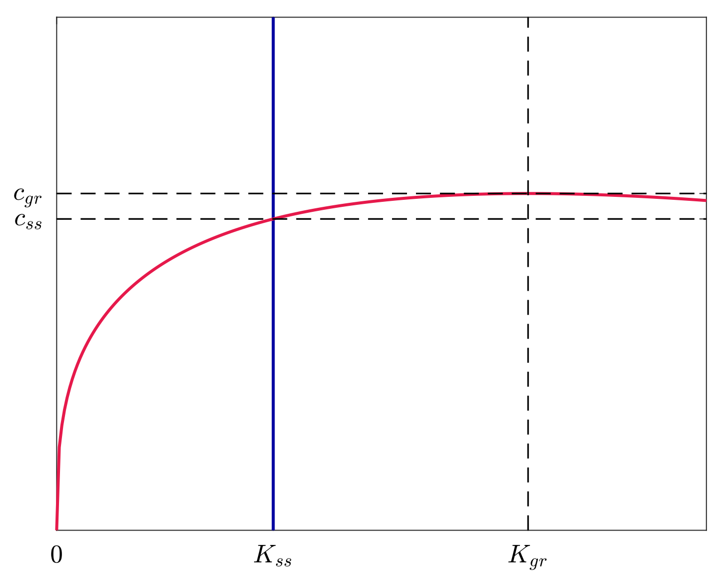


Fig. 9.3.3: *How the steady state compares to the Golden Rule.*

Anticipation Effects

Thinking in terms of general equilibrium can be useful for thinking about how anticipation of things that will happen in the future can affect decisions in the present. Let's consider an example. Suppose the economy is in steady state and suddenly everyone anticipates that a technological breakthrough will lead to a change in the production function from $F(K, 1)$ to $AF(K, 1)$ starting in year T, where $A > 1$. Figure 9.3.4 shows what the effect of this would be.

Once it happens, the technological improvement shifts the $c_t = c_{t+1}$ line to the right: higher A means that it takes higher K to have $AF_K(K, 1) = \frac{1}{\beta} - 1 + \delta$ (which is the condition for consumption stay constant). In addition, the $K_t = K_{t+1}$ curve shifts up: higher productivity means it is possible to afford more and still maintain the capital stock. Once period T arrives, we can apply the analysis we did for the constant-technology case. From T onward, the economy must be on the new saddle path that leads to the new steady state. But before period T, the dynamics of capital and consumption are still governed by the old technology. As illustrated in the figure, the initial level of consumption must be such that, by the time period T arrives, the dynamics under the *old* technology lead to the saddle path of the *new* technology.

In economic terms, what happens is that anticipation of a technological improvement leads to higher consumption through a wealth effect. Since technology has not improved yet, higher consumption implies that the economy is investing less than is necessary to maintain the old capital stock, so the capital stock begins to shrink. A lower capital stock means that the marginal product of capital is higher and therefore the interest rate is higher than $\frac{1}{\beta} - 1$. This means that the household chooses a rising path for consumption. Hence, in

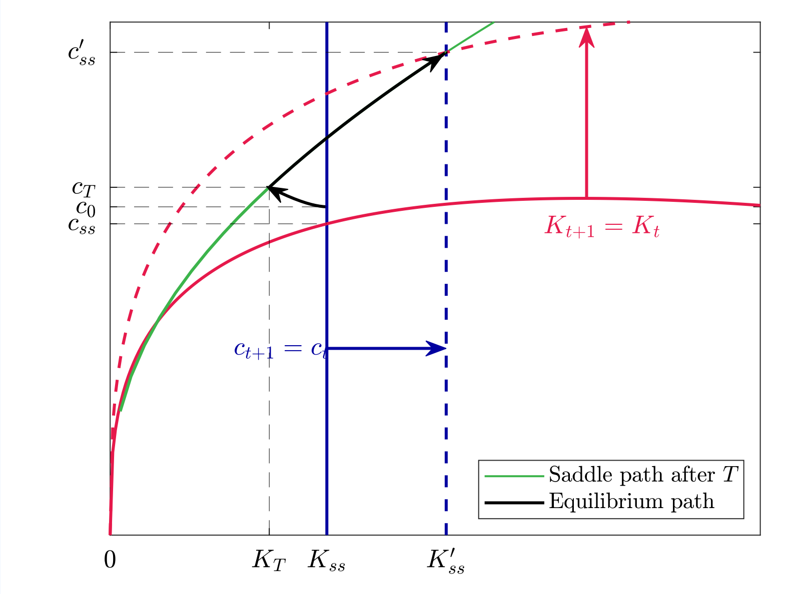


Fig. 9.3.4: *An anticipated technological improvement.*

anticipation of the technological improvement, consumption first jumps from c_{ss} to c_0 and then gradually rises over time. When period T arrives, the capital stock has shrunk to K_T and consumption has reached c_T, so the economy is exactly in the saddle path that will lead it to the new steady state K'_{ss}, c'_{ss}.

Exercises

9.1 **The First Welfare Theorem**

Consider the following policies. In each case, explain whether, in your view, the policy is justified and why.

(a) Setting a maximum price that electrical utilities may charge.

(b) Mandatory vaccinations.

(c) Workplace safety standards.

(d) Rent control.

(e) A 35-hours-per-week limit on working hours.

(f) Banning grocery stores from opening on Sundays.

(g) Banning self-service at gas stations.

9.2 **Storage**

Suppose the production function is the following:

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

and the rate of depreciation is $\delta = 1$.

(a) Why is this production function called " storage"?

(b) What will be the real interest rate and the wage in this economy?

(c) With standard preferences, will consumption increase over time, remain constant over time or decrease over time?

9.3 **Baumol's Cost Disease**

The representative household consumes two different goods: x (clothes) and y (performances by string quartets). Its preferences are given by:[5]

$$u(x, y) = \left(\alpha^{\frac{1}{\epsilon}} x^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)^{\frac{1}{\epsilon}} y^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}$$

Each of the two goods is produced using only labor. If L_x units of labor work in producing clothes, output of clothes is $x = A_x L_x$; if L_y units of labor play in string quartets, output is $y = A_y L_y$. The household supplies one unit of labor inelastically and is indifferent as to how much it supplies to each industry. The price of clothes is denoted p_x and the price of a quartet performance is p_y. The wage rate is w.

(a) Use a software like Matlab or Excel to plot two sets of indifference curves for this utility function. Set $\alpha = 0.5$ in both cases, $\epsilon = 0.5$ in one of them and $\epsilon = 2$ in the other.

(b) Set up the problem of a household that obtains income w from supplying one unit of labor inelastically and has to decide how much to consume each of the two goods. Show that the household will consume:

$$x = \frac{w}{p}\alpha\left(\frac{p_x}{p}\right)^{-\epsilon}$$
$$y = \frac{w}{p}(1-\alpha)\left(\frac{p_y}{p}\right)^{-\epsilon}$$

where $p \equiv \left(\alpha p_x^{1-\epsilon} + (1-\alpha) p_y^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$. [This takes a bit of work. If you want, you can look up the derivation on the book website].

(c) Set of the problem of a firm in each of the two industries that needs to decide how much labor to hire to maximize profits. What values of $\frac{p_x}{w}$ and $\frac{p_y}{w}$ are consistent with each industry hiring a positive but not infinite amount of labor?

(d) What will be the price of string quartet performances relative to clothes $\frac{p_y}{p_x}$?

(e) Suppose that over time A_x rises and A_y stays the same. What does this mean? Do you find this plausible? What will be the effect on the relative price of string quartets? Explain.

(f) Use your answers to parts (b) and (d) to compute the relative quantities of each of the two goods that will be consumed, i.e. solve for the equilibrium level of $\frac{y}{x}$.

[5] This is known as a "constant elasticity of substitution" utility function. The parameter ϵ represents the elasticity of substitution between goods x and y: a higher number means that that are close to perfect substitutes. For reference, when $\epsilon \to 1$, these preferences converge to the Cobb-Douglas case $u(x, y) = x^{\alpha} y^{1-\alpha}$.

(g) Use your answer to part (f) to compute the relative allocation of labor across the two industries, i.e. solve for the equilibrium level of $\frac{L_y}{L_x}$.

(h) Suppose that over time A_x rises and A_y stays the same. What happens to the quantity of string quartet performances over time? How does the answer depend on ϵ? Explain.

9.4 **Pro-Worker Policies**

Consider the following one-period economy. The production function is $F(K, L)$. The representative household owns all the capital stock K, which is exogenously given. The household's preferences are $u(c, l)$ where, as usual, $l = 1 - L$. The labor and capital markets are perfectly competitive.

Denote by L^* the amount of labor that households supply in a competitive equilibrium with no government intervention. Now suppose that the government dictates a new law that prohibits the household from working more than $L^* - \varepsilon$, where ε is a small positive number.

(a) Show that the policy will increase wages and lower the rental rate of capital.

(b) Show that the policy will make the representative household worse off.

(c) Now suppose that there are two representative households in the economy. Household A owns the capital stock and does not work. Its preferences are given by $u(c)$. Household B has preferences $u(c, l)$ and does not own any capital. Suppose the government enacts the same policy as before (i.e. it prohibits household B from working more than $L^* - \varepsilon$, where L^* is the amount it works under a no-intervention equilibrium). Show that this policy makes household A worse off and household B better off.

9.5 **Differences in Preferences**

There are two countries: Industria and Lethargia. Both countries have a population of 1 and identical production functions:

$$Y = AL$$

where L is the total amount of time spent in market labor.

Preferences are given by:

$$u(c, l) = \log(c) + \theta \log(l) \tag{9.3.18}$$

where c is consumption and l is leisure. Within each country, everyone is identical but θ is different for residents of the two countries. Denote the two values by θ^I and θ^L and assume $\theta^L > \theta^I$.

Both countries run free market economies, with perfectly competitive labor markets.

(a) What is the wage in each country?

(b) Set up the problem of a representative household that has to decide how to divide its time (normalized to a total of 1) between market work and leisure.

(c) What fraction of its time will the representative household spend on market work in each country?

(d) Which country has higher GDP?

(e) Suppose a researcher does the following:

i. assumes preferences are the same across the world, with functional form (9.3.18),

ii. estimates θ using data from Lethargia,

iii. gathers data on consumption and leisure in each country,

iv. computes λ: the relative welfare of residents of Industria, taking Lethargia as a benchmark with $\lambda = 1$ the way Jones and Klenow took the US as a benchmark.

Find an expression for λ. Which country would the researcher conclude has a higher standard of living? Explain.

9.6 **Capital Income Taxes**

An economy has the production function

$$Y = K^{\alpha} L^{1-\alpha}$$

It is populated by two types of households: workers and capitalists.

Workers supply a total of L units of labor inelastically and consume all their income in every period, so they don't really make any decisions. Their consumption in any given period is given by:

$$c_t = w_t L + T_t$$

where w_t is the wage and T_t is a transfer they get from the government.

Capitalists have preferences given by:

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

They do not work: they get their income from capital, and it is taxed at a rate τ. However, they can choose how much to consume and how much to save in the standard way. Their budget constraint is:

$$K_{t+1} = -c_t + \left(1 + (1-\tau)\left[r_t^K - \delta\right]\right) K_t$$

where K_t is the capital stock in period t, δ is the rate of depreciation, τ is the tax rate and r_t^K is the rental rate of capital (which each capitalist household takes as given but depends on the total capital stock in the usual way).

What does this budget constraint mean? For each unit of capital the capitalist has, it obtains a rental. The government taxes this rental (net of depreciation) at a rate τ. Assume that the government taxes interest income in the same way as it taxes income from renting capital, so that equation (9.1.11) holds.

(a) Set up the maximization problem of a capitalist household.

(b) From the first order conditions, find an after-tax version of the Euler equation. (You can skip steps if you want)

(c) If in the long run this economy reaches a steady state in which consumption of the capitalists is constant, what are the pre-tax and after-tax interest rates in this steady state?

(d) If the economy is in a steady state, what is the rental rate of capital? How does it depend on τ?

(e) If the economy is in a steady state, what is the capital stock?

(f) How much revenue does the government collect from the capital-income tax in steady state?

(g) What is the level of wages in steady state?

(h) Suppose that the government uses all the revenue from the capital-income tax to finance transfers to the workers. What level of consumption do workers attain in steady state? How does it depend on τ? Explain.

9.7 **Getting Old**

An economy has two types of households: young and old. Let μ be the fraction of households that are young. We hold the total population constant and consider the effects of higher or lower μ (which could be the result of differences in fertility and mortality). Each young household supplies x units of labor inelastically, so the total amount of young labor is $L^Y = \mu x$; old households supply z units of labor, also inelastically, so the total amount of old labor is $L^O = (1-\mu)\, z$. Assume $x > z$.

The production function is:

$$Y = K^\alpha L^{1-\alpha}$$

where K is the capital stock and L is the total amount of labor: $L \equiv L^Y + L^O$.

(a) Holding K constant, how do GDP, wages and and the rental rate of capital depend on μ?

(b) How will the steady-state capital stock capital compare between two economies with different levels of μ? Explain.

(c) Now imagine instead that the economy doesn't use capital and the production function is:

$$Y = \left(L^Y\right)^\gamma \left(L^O\right)^{1-\gamma}$$

How do the wage levels of young and old workers depend on μ?

9.8 **A New Technology**

Consider a one-period economy where the representative worker has preferences:

$$\log(c) + \theta \log(l)$$

where c is consumption and l is leisure. The worker must choose how to divide its time between leisure l and market work L, so:

$$l + L = 1$$

The labor market is perfectly competitive. The equilibrium wage is denoted w.

(a) Set up the problem of the representative worker, who must decide how to allocate his time. Assume the worker has no source of income other that its earnings from labor. Solve for the household's choice of L.

The production technology uses labor and capital to produce goods according to:

$$Y = K^{\alpha} L^{1-\alpha}$$

where K is the capital stock, which is exogenously given. The capital-rental market is also perfectly competitive. The equilibrium rental rate of capital is denoted r^K.

(b) Set up the problem of the representative firm, which must choose capital and labor inputs to maximize profits. Derive first order conditions and find the level of w and r^K such that markets clear.

(c) How do w and r^K depend on K and θ? Explain.

(d) Suppose a team of scientists develops a new technology for producing goods. Using this new technology, it's possible to produce goods according to:

$$Y = AK$$

where A is a parameter. The scientists tried out the technology at a very small scale, so anything they did was too small to affect equilibrium prices. Prove that the new technology turned out to be profitable if and only if $K > \bar{K}$, where $\bar{K}$ is some number. Find an expression for $\bar{K}$. How does $\bar{K}$ depend on θ? Explain.

9.9 An Oil-Producing Economy

Suppose that the production function is given by $F(K, L) = K^{\alpha} L^{1-\alpha} + e$. GDP is the sum of regular output, which is produced using capital and labor, and oil extraction e, which is exogenous. The representative household has preferences:

$$\sum_{t=0}^{\infty} \beta^t \log(c_t)$$

and supplies $L = 1$ units of labor inelastically. Capital depreciates at rate δ.

(a) On the same graph, draw two phase diagrams, one for $e = 0$ and one for $e = \bar{e} > 0$ and label the steady state in each case.

(b) Suppose that the economy starts at $t = 0$ at the $\bar{e}$ steady state but everyone suddenly realizes that oil will run out at time T, so that:

$$e_t = \begin{cases} \bar{e} & \text{if } t < T \\ 0 & \text{if } t \geq T \end{cases}$$

Using the phase diagrams, plot the evolution of K_t and c_t from $t = 0$ until $t = T$ and from $t = T$ onwards. Describe in words why consumption and investment evolve the way they do, and what happens to wages and the interest rate.

9.10 Wizards and Cows

The Calevingians have two basic economic activities: gathering fruit and raising cows.

Wild fruit grows in the Western part of the kingdom. The Calevingians have not mastered the art of growing fruit trees, so they have no control over how much fruit is available. They just visit the fruit trees every week and pick whatever fruit is ripe. Luckily, the kingdom includes fruits with different seasonal patters so there is ripe fruit more or less evenly throughout the year. Not all years are the same, however. Some years are warm and fruit is plentiful, while others are colder and result in less fruit. The Calevingians have complete trust in their chief wizard, who gives them a weather forecast (and thus a fruit forecast) with a horizon of a few years.

Cows graze in pastures in the Eastern part of the kingdom. The grass on which cows feed grows just as well in cold and warm weather. Each family owns a plot of land and keeps their own cows there. Cows are raised primarily for meat; the Calevingians have not learned to milk them. One of the main decisions that Calevingians need to make is how many of their cows to slaughter for meat each year and how many (including newborn calves) to keep fattening from one year to the next. They understand that the more cows they keep in their plot of land, the less grass each of them will have available for grazing.

(a) Write down mathematically the economic decision problem faced by a Calevingian household. Explain the meaning of each equation you write down.

(b) Derive first-order conditions (you may skip steps if you want).

(c) Suppose the chief wizard announces that the next few years will be cold and therefore yield less fruit. How do households react to this announcement?

(d) Despite their rather simple economy, the Calevingians have a fairly sophisticated legal system; the concept of a sale, a loan and an interest rate are well-established and contracts are enforced very effectively. It's possible, for instance, to borrow in order to buy more cows or, conversely, to sell one's cows and lend the proceeds of the sale to someone else. What happens to interest rates when the wizard announces the upcoming cold weather?

9.11 **Patience and Investment**

Suppose the production function is $F(K, L) = K^\alpha L^{1-\alpha}$, and the labor supply is exogenous and equal to $L = 1$. The representative household's preferences are given by:

$$\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}$$

and the depreciation rate is δ.

(a) Use equation (9.3.17) to find an expression for the level of capital that this economy will have in the steady state.

(b) Find an expression for the level of GDP that the economy will have in the steady state.

(c) Find an expression for the steady-state investment-to-GDP ratio. How does it depend on β? Explain.

9.12 **Optimal vs Fixed Savings Rates**

Suppose the economy is as in Exercise 9.11, with $\alpha = 0.4$, $\beta = 0.95$, $\delta = 0.08$ and $\sigma = 2$. The economy starts with an initial capital stock $K_0 = 2$. We are going to compare how this economy behaves under a fixed investment rate and under an optimally-chosen investment rate.

(a) First suppose that, like we assumed in Chapter 4, the investment rate is exogenous and constant, and equal to the steady-state investment-to-GDP ratio that you found in Exercise 9.11. Compute GDP, consumption, investment and the capital stock for the first 150 years of this economy.

(b) Now suppose that consumption and investment are chosen in accordance with (9.3.13). Find the level of initial consumption that is consistent with the system of equations (9.3.15) - (9.3.13) and the economy converging to a steady state. (You can find a Matlab code that will compute this at the book website). Compute GDP, consumption, investment and the capital stock for the first 150 years of this economy.

(c) Plot the consumption paths of the two economies over the first 150 years on the same graph.

(d) Replace the computed paths of consumption from parts (a) and (b) into the utility function to compute how much utility the representative household will obtain in the first 150 years of each economy.

(e) By what fraction would one have to reduce consumption in the optimal-savings economy for the household to get the same utility as in the exogenous-savings economy?

9.13 **The Enclosure Acts**

The Enclosure Acts were a series of laws that converted common property into private property in Great Britain starting around the XVII century. This exercise asks you to think about their macroeconomic effects.

Consider the following one-period economy.

- There are two sectors:
  - Agriculture produces output according to:

    $$Y_A = N^\alpha L_A^{1-\alpha}$$

    where Y_A is agricultural output, N is natural resources (land) and L_A is labor dedicated to agriculture.
  - Industry produces output according to:

    $$Y_I = K^\alpha L_I^{1-\alpha}$$

    where Y_I is industrial output, K is capital and L_I is labor dedicated to industry.
  - For simplicity, assume that both sectors produce the same good (or that both goods are perfect substitutes), so GDP is $Y = Y_A + Y_I$.

- The industrial sector operates as a competitive market, with a wage rate w and a rental rate of capital r^K. The capital stock is exogenously given and is owned by capitalists, who do not work.
- Agriculture is organized as follows:
  - no one owns the land and anyone can use it,
  - once all the agricultural output is produced, it gets shared equally among agricultural workers.
- The amount of land available and the capital stock are exogenously given.
- Each worker supplies one unit of labor. The total number of workers is L.
- Each worker freely decides whether to work in industry or in agriculture:
  - if the worker chooses industry, he gets a wage w,
  - if the worker chooses agriculture, he gets a share of total agricultural. output

(a) Find an expression for the ratio $\frac{L_A}{L_I}$ that would lead to maximum GDP.

(b) If there are L_A workers in agriculture, what is the income of a worker who chooses agriculture?

(c) If there are L_I workers in industry, what is the income of a worker who chooses industry?

(d) What is the ratio $\frac{L_A}{L_I}$ that would make workers indifferent between choosing agriculture and industry? Argue that this is the allocation of labor that the economy will have. How does it compare to the allocation that would maximize GDP? Explain.

(e) Suppose there is an Enclosure Act that does the following:

- It splits up the land equally among agricultural workers, giving each of them ownership of a specific piece of land.
- It establishes a competitive market for renting land and hiring agricultural workers.

After the reform:

i. Which of the sectors will increase the number of workers it employs? Explain.

ii. What will happen to GDP? Explain.

iii. What will happen to industrial wages? Explain.

iv. What will happen to the rental rate on industrial capital? Explain.

(f) How would the answers to part (e) change if instead of being split up equally the land was given to only some of the agricultural workers? Why?

PART IV

Money and Inflation

This part of the book looks at the role of money in the economy.

In Chapter 10 we study what money is, how the quantity of money is determined and why people choose to hold it.

In Chapter 11 we study how the equilibrium in the money market is determined, what determines the price level and how inflation comes about.

CHAPTER 10

Money

10.1 What is Money?

This question is trickier than is seems.[1] The standard answer is that "money" can be anything that can serves as:

1. A store of value. One can save it (for instance, by keeping it in one's pocket) and use it later.
2. A unit of account. We can express the prices of things in terms of how much money it takes to buy them.
3. A medium of exchange. Money changes hands when people pay for things.

Many different things have been used as money at various times and places: pieces of paper with the faces of historical figures, gold and silver, cigarettes.

Why do we use money at all? The main reason is that it solves what's known as the "double coincidence of wants" problem. Using money, I don't need to find someone who has exactly what I need and wants exactly what I have in order to trade. I can accept money in payment for the goods I sell knowing that others will accept money in payment for the goods I want.

For something to be convenient to use as money, it typically needs to have several properties:

1. It has to be hard to counterfeit. If you want to pay for something with money, you don't want the seller to be wondering whether you are giving them real money or fake money.
2. It has to be easy to carry, since transactions happen in many different places.
3. It has to be durable, otherwise it's not a very good store of value. It makes more sense to use coffee beans than strawberries as money because strawberries are likely to spoil before they can be used in the next transaction.
4. It has to be easily divisible. For transactions to go smoothly it's important to be able to pay the exact price without too much rounding and to be able to make change. The wonderful book by Sargent and

[1] Asmundson and Oner (2012) have a good introduction to this question.

Velde (2014): *The big problem of small change* tells the history of how Europe dealt with the problem of making change.

5. It has to be commonly accepted. Money is only useful is everyone agrees that it is indeed money and accepts it as payment. Sometimes this acceptance is purely a social convention, sometimes it is reinforced by laws.

The properties listed above are satisfied to different degrees by different assets. As a result, there is no unique measure of what is the type of money that is used in any economy or how much of it there is. By convention, several definitions of money are typically studied, each of which draws a somewhat arbitrary line between "money" and "not money". Here are the main ones:

Monetary base	M0	M1	M2
Physical Currency	Physical Currency	Physical Currency	Physical Currency
Central Bank Reserves		Demand deposits	Demand deposits
			Savings deposits
			Small time deposits
			Money market mutual funds

Let's start from M0. This measure counts as money only physical bills and coins. It quite clear that physical currency meets the conditions for something to be money pretty well. Not perfectly, though: it's possible to counterfeit and there are some places where it's not accepted as a means of payment. M1 is a broader measure of money because it also includes demand deposits (basically checking accounts). For most purposes, a checking account satisfies the definition of money: for most transactions, either a check or a debit card is acceptable.[2] M2 is a still broader measure because it includes other types of deposits (savings accounts and smaller time deposits held by individuals) as well as shares in money market mutual funds held by individuals.[3] The assets in this broader measure of money are slightly less easy to use in transactions. That's why these assets have an intermediate degree of "moneyness" and are included in the broader measures of money but not the narrower ones.

The "monetary base" has this name because it's the only part of money that is under the direct control of the government. Therefore it forms the "base" on which every other measure is built. To see this, it is useful to go into the mechanics of how the quantity of money is determined.

10.2 The Supply of Money

The monetary aggregates defined above are in part chosen by the government, which decides how much physical currency to issue, and in part determined by what happens in the banking sector, since deposits are bank liabilities. How does this all fit together?

[2] But not for every transaction. Some stands in the Palo Alto farmers' market will only take physical cash. So (I've been told) will most drug dealers.

[3] A mutual fund is an investment vehicle by which investors each own a proportion of a pool of assets. A mutual fund is called a "money market" fund if it invests in very safe assets so that the total value of the pool of assets doesn't move much, making it very similar to money as a store of value.

To address this, a good place to start is by looking at what a bank balance sheet looks like. Here is a typical bank balance sheet:

Balance Sheet	
Assets	**Liabilities**
Reserves Bonds	Deposits
Loans	Other Liabilities
Other Assets	**Net Worth**

On the left are all the bank's assets: loans, government bonds, etc. One of the main sources of income for banks is the interest it earns on these assets (the other is fees of various kinds). One asset in particular will be relevant to us: Central Bank reserves. The Central Bank acts as a bank for banks and reserves is just the name given to the deposits that banks hold at the Central Bank (i.e. this is an asset for banks and a liability for the Central Bank). Typically, these reserves earn either no interest or a very low rate.[4] Why do banks hold them?

There are two reasons, whose relative importance has been different at different times in history. One reason is that reserves are a way to meet unexpected withdrawals of deposits. Most bank assets are relatively long term and hard to sell so if depositors want their money right away it's useful for the bank to have an asset that can be converted into cash very quickly, and reserves provide this: in most countries, the Central Bank stands ready to exchange reserves for cash whenever banks want it. Nowadays this is usually not the main reason banks keep reserves since deposit insurance has made bank runs quite rare and there are explicit arrangements for banks to get emergency loans to meet deposit withdrawals. Instead, the main reason banks keep reserves is that they are required to do so by regulation. Typically, banks are required to hold a certain minimum level of reserves, set as a percentage of the bank's deposits. The exact percentage usually depends on the type of deposit, though the details vary a lot from one country to another.

On the right hand side are the bank's liabilities: mostly deposits but sometimes also non-deposit borrowing such as long term bonds that the bank has issued. The bank's net worth is the difference between its assets and liabilities.

Let's compute the supply of money in a simplified example. There are three relevant entities:

- the Central Bank,
- a single commercial bank, which is meant to represent the sum of all banks in the economy,
- a single representative household.

These are their balance sheets:

[4] This is changing. In recent years many central banks are paying interest on reserves. We look into this further below.

Central Bank		Private Bank		Household	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
Bonds: b	Reserves: ρd	Reserves: ρd	Deposits: d	Currency: c	0
	Currency: c	Bonds: B		Deposits: d	
	Net Worth	Loans: L	**Net Worth**		**Net Worth**
	$b-\rho d-c$		$\rho d+B+L-d$		$c+d$

The household owns currency and deposits and, in this example, has no liabilities or any other assets. The bank's assets are made up of loans, government bonds and Central Bank reserves. The reserve requirement in this example is set at a fraction ρ of deposits and the bank is satisfying it exactly, so it has ρd in reserves. The Central Bank owns a certain amount of government bonds b and its liabilities are the private bank's reserves and the outstanding currency.[5]

In this economy we have that, following the definitions above:

- the monetary base is $c+\rho d$,
- M0 is c,
- M1 is $c+d$ (assuming the deposit is a demand deposit).

10.3 Changing the Supply of Money

In this section we'll study the traditional way in which central banks would change the supply of money. The way central banks operate has been changing over the last decade or so. Exercise 10.3 asks you to think about how more modern operating procedures compare with this traditional approach.

Suppose that the Central Bank wants to change the supply of M1 money. We'll come to the reasons why the Central Bank might want to do that later on, but for now let's just accept that the Central Bank wants to do this. The Central Bank doesn't directly control the amount of deposits, which represent the biggest component of M1, but it affects it indirectly through "open market operations". Let's see how this works.

An open market operation is a trade by the Central Bank where the Central Bank either buys bonds and pays for them with reserves or sells bonds and accepts reserves as payment. It's called open market operation because the Central Bank is trading just like anyone else in the open market. Let's work through how an open market operation takes place in the example above. Suppose that the Central Bank buys government bonds worth Δ from the private bank and pays for them by crediting Δ reserves to the private bank. These are new reserves: they come into existence because the Central Bank creates them. Balance sheets are now:

[5] Thinking of currency as a liability of the Central Bank is a bit counterintuitive at first: the Central Bank doesn't really have an obligation to pay anything to holders of currency. This wasn't always the case: it used to be that currency represented a promise by the Central Bank to deliver gold to the holder. Then currency was in every sense a liability, and the accounting reflected that.

Central Bank		Private Bank		Household	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
	Reserves: $\rho d + \Delta$	Reserves: $\rho d + \Delta$	Deposits: d	Currency: c	0
Bonds: $b + \Delta$	Currency: c	Bonds: $B - \Delta$			
	Net Worth	Loans: L	**Net Worth**		**Net Worth**
	$b - \rho d - c$		$\rho d + B + L - d$		$c + d$

But this is not the end of the story. Now the private bank has reserves of $\rho d + \Delta$ but it's only required to have ρd. It now has "excess reserves". Since reserves don't earn any interest, the bank will try to lend out these excess reserves.[6] Let's imagine that the bank makes a loan of Δ to the household. What exactly happens when the bank makes a loan?

- The bank gives the borrower a check for Δ in exchange for a promise that the borrower will pay it back with interest later.
- The borrower deposits the check. Maybe the borrower deposits it in the same bank; maybe the borrower deposits it in a different bank; maybe the borrower hands over the check to someone else (for instance, someone who sells him a car) and then that person deposits the check in their bank. Since we are adding up over all banks and over all households, all these variants are equivalent.

The balance sheets are now:

Central Bank		Private Bank		Household	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
	Reserves: $\rho d + \Delta$	Reserves: $\rho d + \Delta$	Deposits: $d + \Delta$	Currency: c	Loans: Δ
Bonds: $b + \Delta$	Currency: c	Bonds: $B - \Delta$		Deposits: $d + \Delta$	
		Loans: $L + \Delta$			
	Net Worth		**Net Worth**		**Net Worth**
	$b - \rho d - c$		$\rho d + B + L - d$		$c + d$

Notice that even though the bank "lent out the excess reserves" the reserves don't actually disappear. They are still there. It's more accurate to say: the private banks take advantage of the relaxation of the reserve-to-deposit ratio to expand loans and deposits.

Now suppose that the borrower (or anyone that the borrower made a payment to) wants to take out a fraction of the Δ new deposits in cash. Call that fraction χ, so that the borrower wants to have $\chi\Delta$ cash and $(1 - \chi)\,\Delta$ deposits. The borrower goes to the ATM and makes a withdrawal. What exactly happens when the borrower does this?

- The bank asks the Central Bank for cash (note that in our example the bank held zero cash to begin with).

[6] This assumes that there are lending opportunities out there where the bank will in fact earn a positive interest rate. Later we'll think about what happens if interest rates are zero, or when the Central Bank pays interest on reserves.

- The Central Bank prints physical currency and gives it to the bank. In return, it reduces the amount of reserves owed to the bank.

- The bank hands over the cash to the borrower. In return, it reduces the balance on the borrower's deposit.

The balance sheets are now:

Central Bank		**Private Bank**		**Household**	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
	Reserves: $\rho d+(1-\chi)\,\Delta$	Reserves: $\rho d+(1-\chi)\,\Delta$	Deposits: $d+(1-\chi)\,\Delta$	Currency: $c+\chi\Delta$	Loans: Δ
Bonds: $b+\Delta$	Currency: $c+\chi\Delta$	Bonds: $B-\Delta$		Deposits: $d+(1-\chi)\,\Delta$	
		Loans: $L+\Delta$			
	Net Worth		**Net Worth**		**Net Worth**
	$b-\rho d-c$		$\rho d+B+L-d$		$c+d$

Notice that nobody's net worth changes in this whole series of transactions. The participants are just exchanging different types of assets and liabilities with each other.

The process is not over. The bank has $\rho d+(1-\chi)\,\Delta$ reserves but the reserve requirement is only $\rho\,(d+(1-\chi)\,\Delta)$ so there are excess reserves of $(1-\rho)\,(1-\chi)\,\Delta$. Therefore the process repeats itself, just scaled down by $(1-\rho)\,(1-\chi)$, and then by $((1-\rho)\,(1-\chi))^2$, and then by $((1-\rho)\,(1-\chi))^3$, etc. This defines a geometric series, so we can compute the total effect as follows:

$$\text{Change in Deposits} = \sum_{n=0}^{\infty}((1-\rho)\,(1-\chi))^n \underbrace{(1-\chi)\,\Delta}_{\text{First round effect}} = \frac{(1-\chi)\,\Delta}{\rho+\chi-\rho\chi}$$

$$\text{Change in Currency Holdings} = \sum_{n=0}^{\infty}((1-\rho)\,(1-\chi))^n \underbrace{\chi\Delta}_{\text{First round effect}} = \frac{\chi\Delta}{\rho+\chi-\rho\chi}$$

$$\text{Change in Reserves} = \Delta - \text{Change in Currency Holdings} = \frac{\rho(1-\chi)}{\rho+\chi-\rho\chi}\Delta$$

$$\text{Change in the Monetary Base} = \text{Change in Reserves} + \text{Change in Currency Holdings} = \Delta$$

$$\text{Change in M1} = \text{Change in Deposits} + \text{Change in Currency Holdings} = \frac{\Delta}{\rho+\chi-\chi\rho}$$

At the end, we have balance sheets that look like this:

Central Bank		**Private Bank**		**Household**	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
	Reserves:	Reserves:	Deposits:	Currency:	Loans:
	$\rho d + \frac{\rho(1-\chi)}{\rho+\chi-\rho\chi}\Delta$	$\rho d + \frac{\rho(1-\chi)}{\rho+\chi-\rho\chi}\Delta$	$d + \frac{1-\chi}{\rho+\chi-\chi\rho}\Delta$	$c + \frac{\chi}{\rho+\chi-\chi\rho}\Delta$	$\frac{1}{\rho+\chi-\chi\rho}\Delta$
Bonds: $b+\Delta$	Currency:	Bonds: $B-\Delta$		Deposits:	
	$c + \frac{\chi}{\rho+\chi-\chi\rho}\Delta$			$d + \frac{1-\chi}{\rho+\chi-\chi\rho}\Delta$	
		Loans:			
		$L + \frac{1}{\rho+\chi-\chi\rho}\Delta$			
	Net Worth		**Net Worth**		**Net Worth**
	$b-\rho d-c$		$\rho d+B+L-d$		$c+d$

so now the private banks are meeting the reserve requirement exactly:

$$\frac{\text{Reserves}}{\text{Deposits}} = \frac{\rho d + \frac{\rho(1-\chi)}{\rho+\chi-\rho\chi}\Delta}{d + \frac{1-\chi}{\rho+\chi-\chi\rho}\Delta} = \rho$$

The quantity:

$$\omega \equiv \frac{\text{Change in M1}}{\text{Change in the Monetary Base}} = \frac{1}{\rho+\chi-\chi\rho} \tag{10.3.1}$$

is known as the M1 money multiplier. It's called a multiplier because whenever the Central Bank changes the monetary base (which it controls directly), the magnitude of the change in M1 is the change in the monetary base times the multiplier.

The money multiplier depends on two things:

- ρ : the ratio of reserves to deposits ρ. For the most part, the Central Bank can change this number by changing the legal reserve requirement.[7]
- χ : the fraction of their M1 money that people want to hold in physical cash. This is not under the control of the Central Bank: it can vary over time depending of the evolution of payment systems or the public's confidence in the banking system.

In most of what we'll do later, we'll just say that the Central Bank controls the money supply. What we mean by this is that the Central Bank directly controls the monetary base. By understanding how the money multiplier works, it can control M1 fairly accurately, albeit indirectly.

What if the Interest Rate is Zero?

The analysis above was built on the assumption that banks try to maintain reserves as low as possible. This makes sense as long as reserves pay zero interest and other assets pay positive interest. However, if the interest rate fell all the way down to zero (or if the Central Bank started paying market interest rates on reserves), this logic would break down. Banks would be perfectly willing to hold reserves above the legal requirement since the alternatives are not better. This would mean that changes in the monetary base need not lead to changes

[7] As mentioned before, this assumes that banks are trying to keep the minimum possible level of reserves so that the legal requirement is binding.

in the M1 money supply, since all the new reserves would just sit in bank balance sheets without triggering an expansion in loans and deposits.

This scenario has been realized in recent years, as shown in Figure 10.3.1. Nominal interest rates fell to almost zero in late 2008. At around the same time, the Federal Reserve decided to start paying interest on excess reserves (i.e. reserves above the legal reserve requirement). Reserves were suddenly a more attractive asset for banks to hold, and banks started holding large amounts excess reserves. As a result, the M1 money multiplier fell from about 2 to less than 1. The monetary base was increased almost fivefold but M1 increased much less. We'll think more about what happens when the interest rate is near zero in Chapter 15.

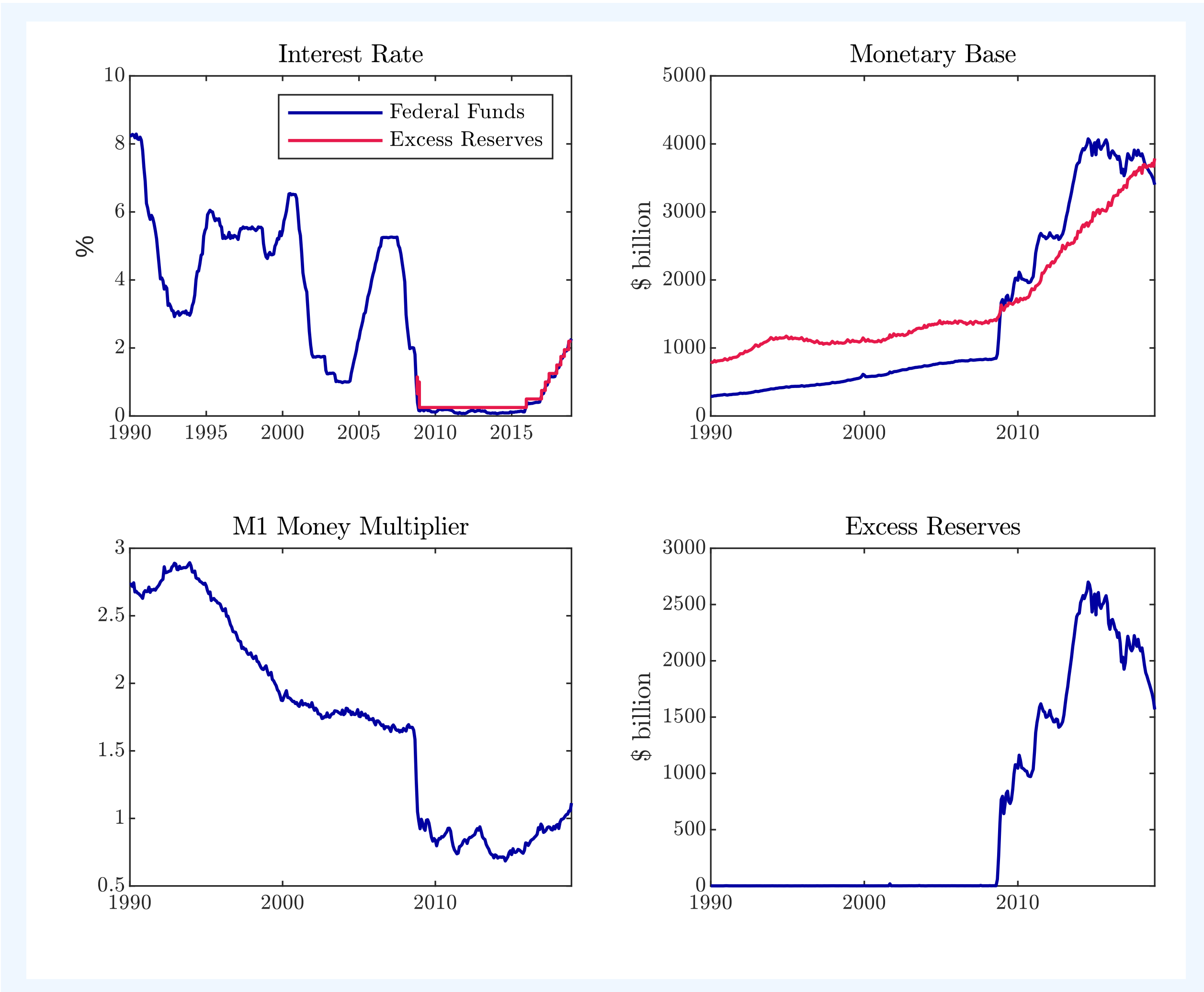


Fig. 10.3.1: *Monetary Aggregates in the US when nominal interest rates reached zero. Source: Board of Governors of the Federal Reserve System.*

10.4 The Demand for Money

Let's focus on M1 as our definition of money. Why do people hold money? Why is it that instead of only holding assets that earn interest (like physical capital, government bonds, etc.) people choose to hold physical currency and checking deposits, which earn no interest?[8] We'll propose a simple model based on the idea that money is necessary to carry out transactions. This model is known as the Baumol-Tobin model since it was first analyzed by Baumol (1952) and Tobin (1956).

Let's imagine that there are two types of assets:

- M1 money, which does not pay any interest,
- assets that do pay interest, all of which pay the same nominal interest rate i.

In the course of a period (for instance, the period can be a year), a household will spend Y in real terms (Y stands for real GDP). The price of a good is p, so in nominal terms the household will spend pY. This spending is not all at once: it's spread evenly over the period. For instance, if the period is a year, the household spends $\frac{pY}{365}$ each day.

Whenever the household wants to pay for something, it must use money. However, this does not mean that the household needs to have pY dollars all at once. Whenever it wants, the household can "go to the bank"' and swap some of its interest-bearing assets for money. One way of "going to the bank" is to go to the ATM and get physical currency (which is money) from one's savings account (which is not money under the M1 definition). But "going to the bank" need not mean literally going to a physical bank branch. Another way of "going to the bank" is to go to their online brokerage account and sell some bonds (which are not money), depositing the proceeds in a checking account (which is money under the M1 definition). We are going to assume that there is a fixed cost F (in real terms) of going to the bank. F can literally represent ATM fees but also the time and mental cost of dealing with the issue.

The problem of the household is to decide how many times per period it goes to the bank. The advantage of going to the bank many times is that it allows the household to have very low levels of money, so that most of the household's wealth is earning interest most of the time. The disadvantage is that it requires paying the fixed cost F many times. Let N denote the number of times per period that the household goes to the bank. Figure 10.4.1 shows how the amount of money held by the household evolves over time for two values of N.

Each time the household goes to the bank, it brings up its money balance to $\frac{pY}{N}$. Then the balance starts to decrease gradually as the household spends the money. Eventually, when the balance reaches zero, the household goes to the bank again to get more money. It's clear from the picture that the household will, on average, hold less money the more often it goes to the bank. Indeed, the average money balance is simply:

$$M = \frac{pY}{2N} \tag{10.4.1}$$

How does the household choose N? Mathematically, it solves the following problem:

$$\min_N pFN + i\frac{pY}{2N} \tag{10.4.2}$$

[8] Some checking deposits do earn interest, but it's typically much lower than what one could earn by holding some other asset.

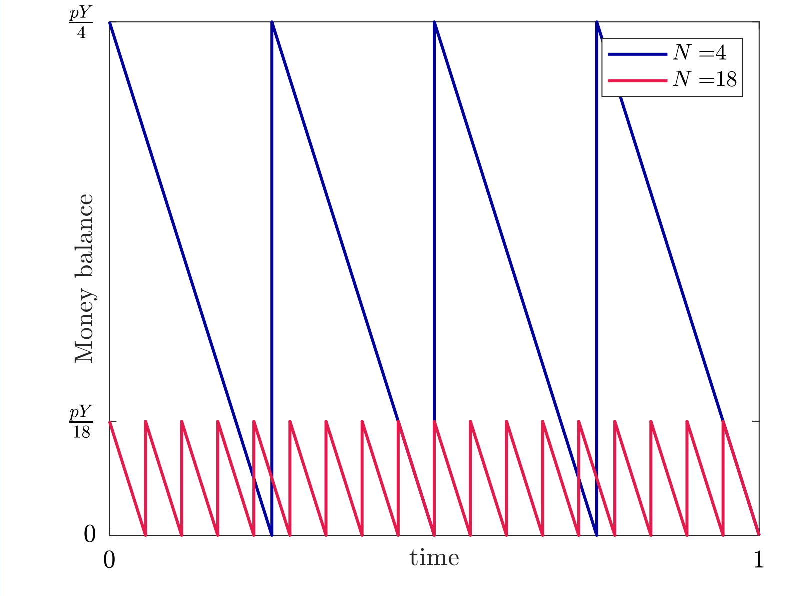


Fig. 10.4.1: *Money balances over time in the Baumol-Tobin model.*

What does this mean? The household is trying to minimize the overall cost of having money for transactions. This cost has two parts. First, if it goes to the bank N times, it pays the cost F each time. Expressed in nominal terms, this gives us pFN. Second, if it goes to the bank N times it will on average hold $\frac{pY}{2N}$ dollars in money. Since this money does not earn interest, there is an opportunity cost of holding it: the foregone interest that the household could have earned if it had held less money. If the interest rate is i, then $i\frac{pY}{2N}$ is the foregone-interest cost of the household's money holdings.

The first-order condition for problem (10.4.2) is:

$$pF - \frac{ipY}{2}N^{-2} = 0$$

so we can solve for N to get:

$$N = \sqrt{\frac{iY}{2F}} \tag{10.4.3}$$

Equation (10.4.3) tells us that the household will go more times to the bank if i is high and if F is low. What's the economic logic of this? If i is high, then the opportunity cost of holding money is high and the household will be willing to go to the bank many times in an effort to hold low amounts of money. On the other hand, if F is low, going to the bank is cheap and the household will, other things being equal, be willing to go to the bank more times.

Replacing (10.4.3) into (10.4.1) and rearranging, we get an expression for the average money balances:

$$M = p\sqrt{\frac{YF}{2i}}$$

or, dividing by the price level, for average "real" money balances:

$$\frac{M}{p} = \sqrt{\frac{YF}{2i}} \tag{10.4.4}$$

"Real money balances" are the answer to the question: "how many goods would the household be able to buy with the amount of money it holds?" Equation (10.4.4) is telling us that real money balances will be higher when:

- Y is high. If the household wants to spend more, this will involve more payments and therefore the household will choose to carry higher real money balances.
- i is low. i is the opportunity cost of holding money. If this is low, the household will choose to hold higher money balances to save on trips to the bank.
- F is high. If going to the bank is costly, the household will choose to hold higher money balances to save on trips to the bank.

Figure 10.4.2 shows the shape of the money-demand function that results from the Baumol-Tobin model. The quantity of money that people want to hold decreases with the interest rate. A higher cost of going to the bank shifts the entire demand schedule to the right.

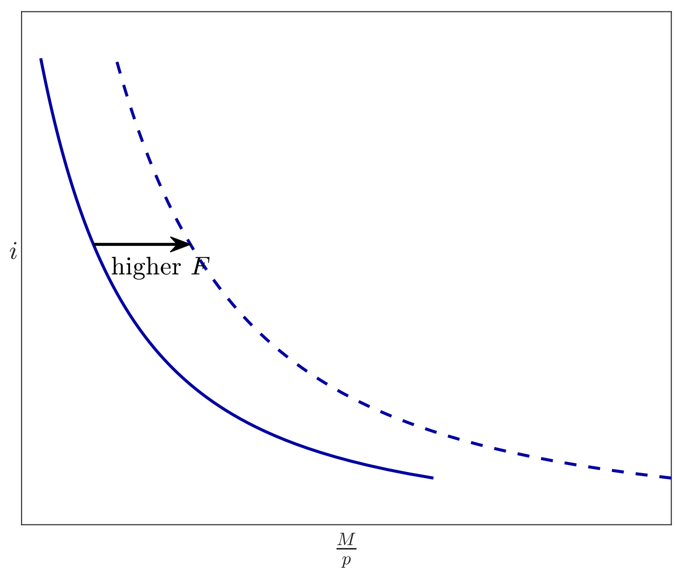


Fig. 10.4.2: *The money demand function in the Baumol-Tobin model.*

The Baumol-Tobin model makes very specific assumptions about how exactly households manage their money: every trip to the bank costs the same, spending is spread out exactly over time and perfectly predictable, etc. We will sometimes want to think about the basic economic forces that the Baumol-Tobin model captures while not expecting the exact formula (10.4.4) to hold. For this purpose, we will sometimes want to think of a generalized money-demand function $m^D(Y,i)$, increasing in Y and decreasing in i. $m^D(Y,i) = \sqrt{\frac{YF}{2i}}$ is just a special case of this more general formula.

There is some debate as to whether the money demand function is sufficiently stable over time to be a useful thing to look at. At times, the quantities of M0, M1 and M2 that people held have moved around quite a bit without changes in interest rates or GDP that would account for them. However, Lucas and Nicolini (2015) and Kurlat (2019) argue that if one constructs an appropriately-weighted composite measure that takes into account how people substitute between different subcomponents of money like physical currency, checking accounts, savings accounts, etc., the resulting money demand has a relatively stable relationship with the interest rate, as predicted by the theory. This is illustrated in Figure 10.4.3.

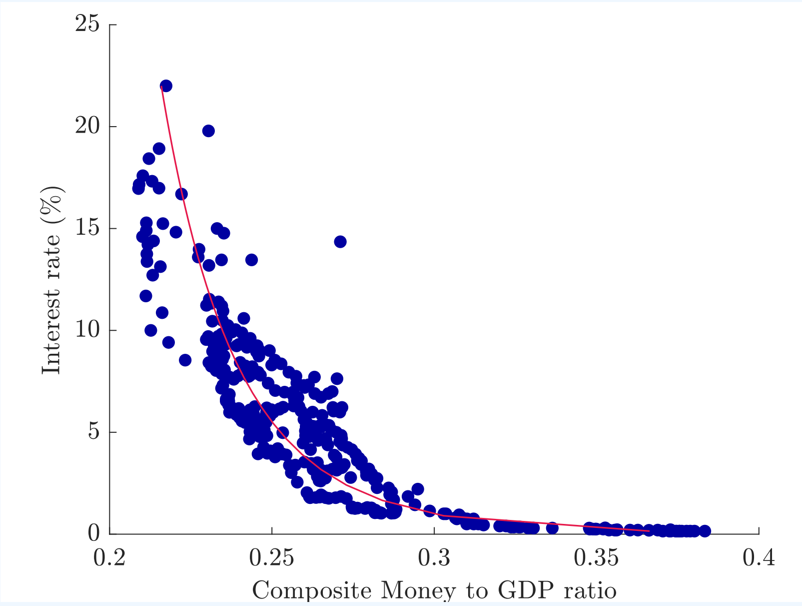


Fig. 10.4.3: *Money demand in the United States. Each dot represents one month between 1980 and 2013. Source: Kurlat (2019)*

Exercises

10.1 **Central Bank Instruments**

Suppose the Central Bank wants to reduce the M1 money supply but does not want to change the monetary base. In what direction should it change reserve requirements?

10.2 **Pickpockets**

Suppose there is an increase in the number of pickpockets. How would that change the fraction of their money that people want to have in cash as opposed to checking deposits? If the Central Bank keeps the monetary base constant, what will happen to the M1 money supply?

10.3 **Interest on Reserves**

Consider an economy where households almost use no cash (so $\chi \to 0$). Furthermore, assume that the Central Bank makes lowers the reserve requirement to almost zero (so $\rho \to 0$).

(a) What is the value of the money multiplier as χ and ρ become small?

(b) Suppose that the Central Bank decides to pay interest on reserves at the market rate i. Does equation (10.3.1) still hold? Why or why not?

(c) Under this policy regime, what determines the quantity of money?

10.4 **ATMs**

Suppose one day, suddenly and unexpectedly, ATMs are invented, which make getting cash more convenient than before. What would be the effect of this invention on the demand for money?

10.5 **Going to the Bank**

The average person in the US has about \$10,000 in M1 money and spends around \$60,000 per year. Suppose this person uses its checking account to pay for all its purchases, behaves according to the Baumol-Tobin model and only has two investment options: its checking account (which pays zero interest) and an investment account (which pays interest). Define "going to the bank" as transferring a balance from the investment account to the checking account.

(a) How often does the average person go to the bank?

(b) Suppose the interest rate is 2%. What must be the perceived cost of each visit to the bank for this type of behavior to be optimal?

(c) How often would the average person go to the bank if the interest rate rose to 3%?

(d) Now suppose the checking account does pay interest, but it only pays 0.5% instead of the full 2% one can get from the investment account. What is the opportunity cost of holding balances in the checking account? What do we infer now about the perceived cost of going to the bank?

CHAPTER 11

The Price Level and Inflation

11.1 Measurement

Inflation is defined as a generalized increase in the level of prices. If the prices of all goods increased by the same percentage, then measuring inflation would be straightforward. It becomes harder when different prices are changing at different rates, or even going in different directions. How do we define the “overall” level of change?

We already encountered this issue when we discussed real and nominal GDP in Chapter 1. There the question was how to measure the “overall” change in output when prices of different goods were changing by different percentages. Here we are interested in prices for their own sake.

The basic idea is going to be to define what is known as a “basket” of goods (i.e., a list of specific quantities of various goods) and measure how the total price of the basket changes. We call this total price a price index. The different methods of measuring the total change in prices have to do with different ways of choosing and updating the basket of goods.

The GDP Deflator

The GDP deflator is a price index that is a side product of the calculation of real GDP. It is defined as

$$\text{GDP deflator} = \frac{\text{Nominal GDP}}{\text{Real GDP}} \times 100.$$

Look back at Example 1.13 from Chapter 1. Suppose that Expandia computes real GDP at 2017 prices, so that real GDP in 2018 is \$2,550, while nominal GDP is \$1,860. Then the GDP deflator would be

$$\text{GDP deflator}_{2018} = \frac{\$1,860}{\$2,550} \times 100 \approx 73.$$

By definition, the GDP deflator is 100 in the base year, so in this case we would say that overall prices went down. Notice that not all prices went down: some went up and some went down. By using the GDP deflator as a price index we are implicitly choosing to weigh each good in proportion to its share of GDP.

The Consumer Price Index

The most commonly used price index weighs the prices of different goods by how much they are consumed rather than how much they are produced. The basket for the CPI is constructed by conducting a survey asking households how much they consume of each good. The CPI is then defined by:

$$CPI = \frac{\sum_j q_j p_{jt}}{\sum_j q_j p_{j0}} \times 100$$

where:

- q_j is the quantity of good j in the basket,
- p_{jt} is the price of good j in period t,
- p_{j0} is the price that good j used to have in the base year.

Example 11.1.

The residents of Luxuria consume only three goods: Ferraris, caviar, and champagne.

Good (j)	Quantity (q_j)	Price in 2017 (p_{j0})	Price in 2018 (p_{jt})
Ferrari (units)	2	100	115
Caviar (kg)	20	4	3
Champagne (liters)	10	2	4
Total basket	1	300	330
CPI		**100**	**110**

Inflation

Having measured a price index, we calculate inflation (denoted by the letter π) by applying the following formula:

$$\pi_t = \frac{P_t}{P_{t-1}} - 1$$

where P_t is a price index in period t. In Example 1.13, inflation (in terms of the GDP deflator) was -27%. When inflation is negative we call it deflation: a general fall in prices. In Example 11.1, inflation (in terms of the CPI) was 10%.

Notice that for a given country there will coexist several measures of inflation, each derived from a different price index. Usually it doesn't make much difference which price index one looks at because the production-based basket that is used in constructing the GDP deflator and the consumption-based basket that is used in constructing the CPI are not that different, at least in the US. It could make a bigger difference in countries that produce and consume very different goods. For instance, in a country that produces oil and exports most of it, a rise in the price of oil would result in a big rise in the GDP deflator but not as much in the CPI.

Figure 11.1.1 shows the evolution of CPI inflation in the US. Inflation was very variable until the 1950s, with times of over 20% inflation and over 10% deflation. Between the 1960s and the early 1980s inflation tended to increase. Since the mid-1980s inflation has been quite low and stable.

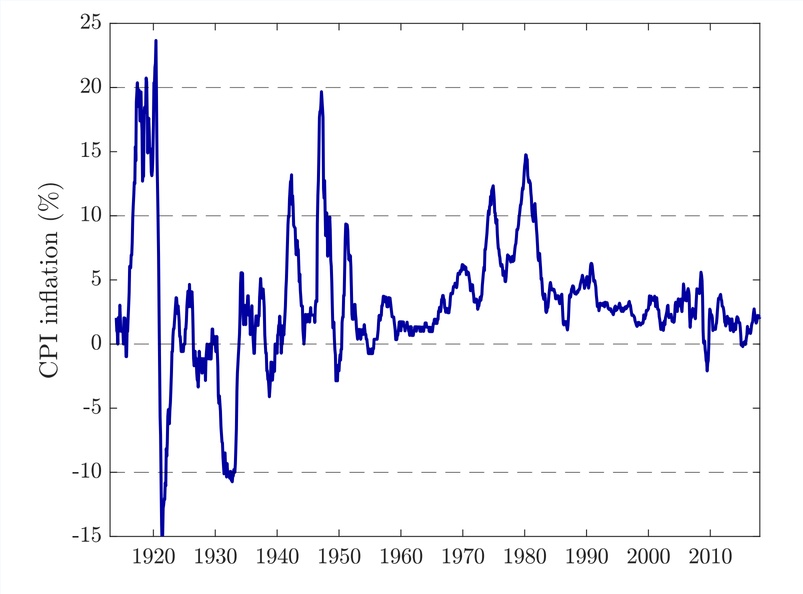


Fig. 11.1.1: *CPI inflation in the US. Source: BLS.*

Nominal and Real Interest Rates

A basic lending transaction works as follows:

- a lender gives a borrower one dollar in period t,
- the borrower pays back $1 + i_{t+1}$ dollars in period $t + 1$.

i_{t+1} is the interest rate between periods t and $t + 1$.[1]

We will often talk about "the" interest rate, although in reality there is no single interest rate for all loans. Typically, the interest rate on government debt is the lowest rate in the country (at least in the US, where the government is perceived as reliable) and rates paid by private borrowers are higher, which compensates for administrative costs, the probability of default, etc.

We are often interested in expressing interest rates in terms of goods rather than in terms of dollars.

Example 11.2.
The interest rate in Usuria on a one year loan that is issued in January 2018 and will be paid back in

[1] An interest rate always involves more than one period: when the loan starts and when it ends. We will adopt the convention to label interest rates according to the period when the loan has to be paid back. Hence i_{t+1} refers to the interest rate on loans that are issued in period t and are due in period $t + 1$.

January 2019 is 11%. Everyone expects that inflation between those dates will be 2%. Suppose someone lends 100 dollars in January 2018. What are they giving up? What do they get in return?

	January 2018	January 2019
Price index	100	102
Loan issued / repayment received (in dollars)	100	111
Loan issued / repayment received (in goods)	$\frac{100}{100} = 1$	$\frac{111}{102} \approx 1.088$

In the example, the 100 dollars of the original loan would be enough to buy exactly 1 consumption basket at the time the loan is granted. By the time the loan is repaid, the 111 dollars that are paid back are not enough to buy 1.11 consumption baskets because prices have risen in the meantime: it is only enough to buy 1.088 consumption baskets. In other words, for each good that the lender gave up at the beginning, he is getting back 1.088 goods one year later. The 0.088 extra goods that the lender obtains are what we call a real interest rate. We call it "real" because it is expressed in terms of goods as opposed to a "nominal" rate that is expressed in dollars. Whenever we have referred to interest rates in Chapters 4-9 we have meant real interest rates because we were thinking about exchanges of real goods over time. Instead, when we studied money demand in Chapter 10 it was the nominal interest rate that mattered because that's what determines the opportunity cost of holding money.

In general, if i_{t+1} is the nominal interest rate, the real interest rate is defined by the following expression:

$$
\begin{aligned}
1 + r_{t+1} &= \frac{\text{Goods you can afford with loan repayment}}{\text{Goods you could afford with the loan when issued}} \\
&= \frac{\left(\frac{1+i_{t+1}}{P_{t+1}}\right)}{\frac{1}{P_t}} = \frac{1 + i_{t+1}}{\frac{P_{t+1}}{P_t}} \\
&= \frac{1 + i_{t+1}}{1 + \pi_{t+1}} \\
\Rightarrow r_{t+1} &\approx i_{t+1} - \pi_{t+1} \qquad (11.1.1)
\end{aligned}
$$

(The last approximation is accurate when π_{t+1} is small.) Equation (11.1.1) is known as the Fisher equation, named after Irving Fisher.

It is not always easy to know what the real interest rate is. There is always some uncertainty as to what inflation is going to be. If one lends or borrows in dollars, as is usual, then until the end of the loan one is not certain how many goods the future dollars are going to be worth. Sometimes we make the distinction between ex-ante real interest rates (meaning the real rate that was expected at the beginning, based on expected inflation) and ex-post real interest rates (based on what inflation turned out to be).

11.2 Equilibrium in the Money Market

In Chapter 10 we looked at the money supply and the money demand separately. An equilibrium in the money market requires that supply equals demand: all the money that is created jointly by the Central Bank and the private banks must be held by someone, voluntarily. We can write the money-market equilibrium condition as:

$$M^S = m^D(Y, i) \cdot p \tag{11.2.1}$$

The left hand side of (11.2.1) is the money supply. We are going to imagine that the Central Bank simply chooses the money supply, by choosing the monetary base and understanding the money multiplier. The right hand side of (11.2.1) is the money demand. This is the result of households' decisions of how much money to hold.

How does a money market equilibrium come about? Suppose that the central bank increases M^S, what changes to induce households to increase their money holdings? The right hand side of (11.2.1) gives us a list of the things that could possible change to restore equilibrium:

- p. The price level could rise. If the price level is higher, then the same amount of real transactions requires more money, so households would want to hold extra money.
- i. Nominal interest rates could fall. If interest rates are lower, the opportunity cost of money is lower and households would be willing to hold more of it.
- Y. GDP could go up. If GDP is higher, there are more real transactions to carry out, which requires more money.

There are different views on which of these three variables tends to respond and why. This turns out to be an extremely important issue. In this chapter we'll look at the so-called "classical" view, which postulates that the real side of the economy is separate from anything having to do with money. Real variables like real GDP and real interest rates are determined purely by real factors (technology, preferences, etc.) that do not change when the money supply changes. One way of stating this view is to say that money is *neutral*. In everything we have done so far we have implicitly adopted this classical view: we studied the forces that determine real variables without any reference to the money supply. In Chapter 14 we'll think about reasons why money might *not* be neutral. For now, let's see how prices and inflation behave if the classical view is correct.

An Economy in Steady State with a Constant Money Supply

Imagine first that the economy is in a steady state where Y and r are constant and the Central Bank holds the money supply M^S constant as well. We'll conjecture that in this economy the price level will be constant as well, and then verify that this is consistent with an equilibrium in the money market.[2] If indeed the price level is constant, then the nominal interest is equal to the real interest rate. Therefore, solving for p in (11.2.1), we get:

$$p = \frac{M^S}{m^D(Y, r)} \tag{11.2.2}$$

which indeed is constant, because we have assumed that M^S, Y and r are constant. Equation (11.2.2) tells us that an economy where the money supply is higher will, other things being equal, have higher prices.

[2] This method of figuring out the equilibrium of a model is sometimes called "guess and verify". Technically, we will show that there is an equilibrium where prices are constant but not that it's the only equilibrium.

People want a certain level of real money balances given by $m^D(Y, r)$, so the price level will be such that $\frac{M^S}{p}$ corresponds to these desired real money balances.

An Economy in Steady State with a Growing Money Supply

Maintain the assumption that Y and r are constant but now assume that the money supply grows at a constant rate μ, i.e. $M_{t+1}^S = (1+\mu) M_t^S$. We'll conjecture that in this economy the price level will also grow at rate μ, and then check that this is consistent with equilibrium in the money market. If:

$$p_{t+1} = (1+\mu) p_t$$

then inflation π_{t+1} is:

$$\pi_{t+1} \equiv \frac{p_{t+1}}{p_t} - 1 = \mu$$

and therefore the nominal interest rate is:

$$i_{t+1} = r + \pi_{t+1} = r + \mu$$

If the money market is in equilibrium in period t, then:

$$\begin{aligned}
M_t^S &= m^D(Y, r+\mu) p_t \\
\Rightarrow M_t^S (1+\mu) &= m^D(Y, r+\mu) p_t (1+\mu) \\
\Rightarrow M_{t+1}^S &= m^D(Y, r+\mu) p_{t+1}
\end{aligned}$$

which implies that the money market is also in equilibrium in period $t+1$. This confirms our conjecture.

Economically, what's going on is the following. Since GDP and nominal interest rates are constant, people want to hold constant real money balances. Since the money supply is growing, prices must be growing too in order to keep money balances constant.

Figure 11.2.1 looks at data on inflation and the growth rate of the money supply over a long period in many countries. Comparing across countries, the data shows that inflation is almost exactly proportional to the growth rate of the money supply.

A Growing Economy

Now suppose that the economy is in a steady-state-with-growth, with Y growing at a constant rate g and a constant real interest rate r. The money supply grows at a constant rate μ. Let's try to find the inflation rate in this economy. Start from (11.2.1) and take the derivative with respect to time:

$$\frac{dM^S}{dt} = \left[\frac{\partial m^D(Y,i)}{\partial Y}\frac{dY}{dt} + \frac{\partial m^D(Y,i)}{\partial i}\frac{di}{dt}\right] \cdot p + m^D(Y,i)\frac{dp}{dt}$$

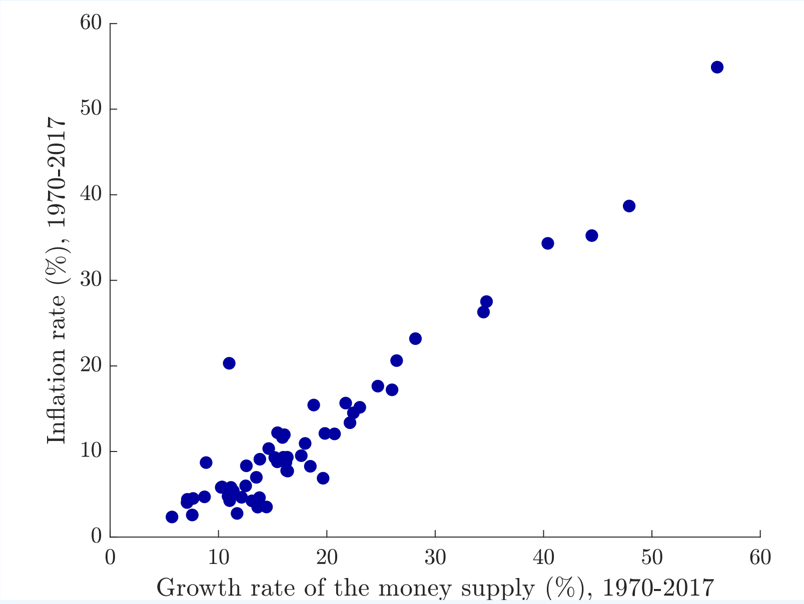


Fig. 11.2.1: *Inflation and the growth rate of the money supply in the long run. Each dot represents one country. Source: World Bank.*

Now divide by (11.2.1) on each side:

$$\frac{\frac{dM^S}{dt}}{M^S} = \left[\left(\frac{\partial m^D\left(Y,i\right)}{\partial Y}\frac{Y}{m^D\left(Y,i\right)}\right)\frac{\frac{dY}{dt}}{Y} + \frac{\frac{\partial m^D(Y,i)}{\partial i}}{m^D\left(Y,i\right)}\frac{di}{dt}\right] + \frac{\frac{dp}{dt}}{p}$$

$$\mu = \left[\left(\frac{\partial m^D\left(Y,i\right)}{\partial Y}\frac{Y}{m^D\left(Y,i\right)}\right)g + \frac{\frac{\partial m^D(Y,i)}{\partial i}}{m^D\left(Y,i\right)}\frac{di}{dt}\right] + \pi$$

Let $\eta \equiv \frac{\partial m^D(Y,i)}{\partial Y}\frac{Y}{m^D(Y,i)}$. η represents the elasticity of money demand with respect to GDP. It is the answer to the question: if GDP rises x%, by what percent does the demand for real money balances increase? Assume the function m^D is such that this elasticity is constant, so:

$$\mu = \left[\eta g + \frac{\frac{\partial m^D(Y,i)}{\partial i}}{m^D\left(Y,i\right)}\frac{di}{dt}\right] + \pi$$

If inflation is constant, then $i = r + \pi$ will be constant so $\frac{di}{dt} = 0$. Then the equation reduces to:

$$\mu = \eta g + \pi$$

and therefore:

$$\pi = \mu - \eta g \tag{11.2.3}$$

so, indeed, inflation is constant. Equation (11.2.3) tells us that, other things being equal, a growing economy will have lower inflation. Why is this? A growing economy means a growing number of transactions and therefore a growing demand for real money balances. This means that the economy can absorb growing quantities of money without resulting in inflation. Why does η show up in the formula? η measures *how much* the demand for money increases when the economy grows. The higher this number, the faster the money supply can grow without leading to inflation. Note that formula (11.2.3) encompasses the steady-state-without-growth examples as special cases.

A One-Time Increase in the Money Supply

Suppose that, starting from a steady state with a constant money supply, at time t there is a sudden, unexpected increase in the money supply, from M^S to $M^{S\prime}$. After this, the money supply is expected to remain constant at $M^{S\prime}$ forever. What's going to happen to the price level?

Before time t we had that $p = \frac{M^S}{m^D(Y,r)}$. After that, the money supply will again be constant, except that the level will be higher. Therefore we are going to be back in a constant-money-supply steady state, where $p' = \frac{M^{S\prime}}{m^D(Y,r)}$. The effect on prices is therefore:

$$\frac{p'}{p} = \frac{M^{S\prime}}{M^S}$$

In other words, prices jump immediately to their new level, and the size of the jump is proportional to the size of the increase in M^S.

The exact causal chain that leads from an increase in the money supply to a rise in the price level is a matter of some debate. Why does everyone immediately raise their prices? In terms of the logic of the model, what happens is that when the Central Bank increases the money supply, everyone suddenly has more money than they would like, so they try to reduce their money balances. But it is impossible for everyone to do this at the same time because someone has to hold the money. Everyone immediately realizes what's going on so money immediately loses value, which is exactly what an increase in the price level means. Note that one condition for this reasoning to be correct is that prices must be flexible, reacting immediately to changes in the supply of money. Starting in Chapter 14, we'll think about the possibility that prices might be "sticky" and react slowly to changes in the money supply. This will be a source of monetary nonneutrality, i.e. of interaction between money and the real economy.

A Change in the Rate of Growth of the Money Supply.

Now let's do a slightly more subtle exercise. Suppose we start at a steady state with the money supply growing at rate μ and, therefore, an inflation rate of μ. At time t there is a sudden, unexpected increase in the *rate of growth* of the money supply, form μ to μ'. After this, the rate of growth of the money supply is expected to remain at μ' forever. What's going to happen?

A naive guess would be to say that inflation will simply increase from μ to μ'. This guess is not wrong, but it's incomplete. If the inflation rate changes from μ to μ', then the nominal interest rate rises from $i = r + \mu$ to $i = r + \mu'$. Using (11.2.1), this implies that real money balances must fall. Higher nominal interest rates increase the opportunity cost of holding money, so people want to hold less of it. But the *level* of M^S does not change at time t: it simply starts growing at a different rate. What makes the money market clear? The price level must rise!

Economically, this is what's going on. People are all simultaneously trying to reduce their money balances because the opportunity cost of holding them has gone up. Since the total (nominal) supply of money has not changed, money loses value, which is the same thing as saying the prices rise.

Figure 11.2.2 shows how the price level evolves over time in the different examples above.

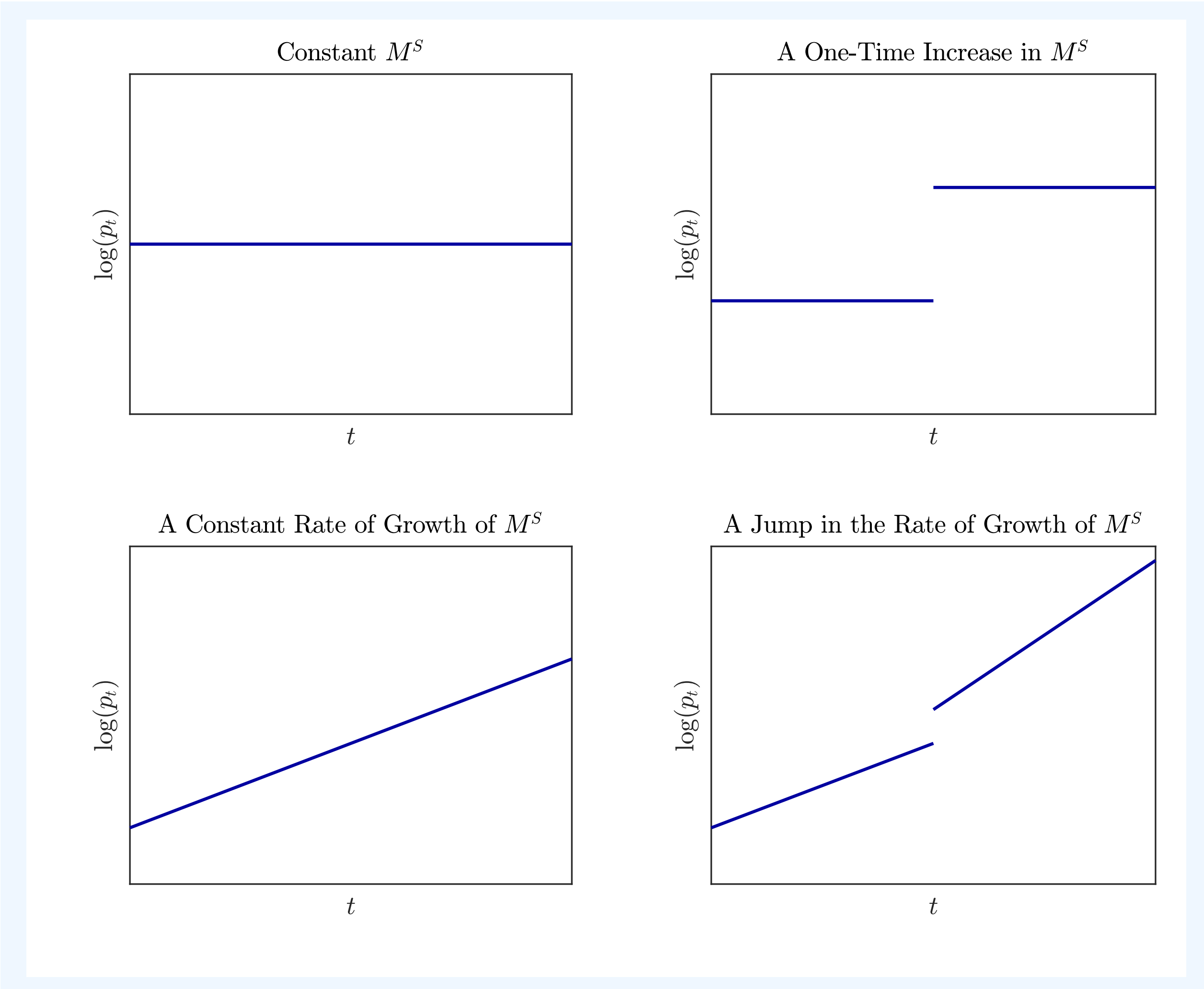


Fig. 11.2.2: *The evolution of prices in several examples.*

The Velocity of Money

The "velocity" of money refers to the number of times a unit of money is used per period. Let's see an example.

Example 11.3.
The money supply is \$2. At the beginning of the period, Ann and Bob each hold one dollar. Over the course of the year, the following things happen:

- Ann produces an apple and sells it to Bob for \$1. Bob produces a banana and sells it to Ann for \$1
- Ann produces asparagus and sells it to Bob for \$1. Bob produces a blueberry and sells it to Ann for \$1
- Ann produces an apricot and sells it to Bob for \$1. Bob produces a blackberry and sells it to Ann for \$1

In the example, nominal GDP is \$6, the money supply is \$2 and and each dollar changes hands 3 times, so the velocity of money is 3. In general, we have that:

$$\underbrace{M}_{\text{Money Supply}} \quad \underbrace{V}_{\text{Velocity}} \equiv \underbrace{\underbrace{p}_{\text{Price Level}} \quad \underbrace{Y}_{\text{Real GDP}}}_{\text{Nominal GDP}} \tag{11.2.4}$$

Equation (11.2.4), sometimes known as the quantity equation, is a *definition*. It's true because this is the way we define the velocity of money.

How does equation (11.2.4) relate to the money-market equilibrium condition (11.2.1)? We can use (11.2.1) to replace $\frac{M}{p}$ in (11.2.4) and rearrange to obtain:

$$V = \frac{Y}{m^D(Y,i)} \tag{11.2.5}$$

Equation (11.2.5) says that any theory of money demand, summarized by a function $m^D(Y,i)$, is also a theory of velocity. Once we have a $m^D(Y,i)$ function, we can simply plug it into (11.2.5) to obtain velocity as a function of Y and i.

Our theory of money demand implies that velocity is an increasing function of the nominal interest rate. We can see this in equation (11.2.5) by noting that m^D is decreasing in i, which implies V is increasing in i. Economically, what this is saying is that if interest rates are higher, people will hold less money, so in order to carry out the same amount of real transactions, each dollar will have to change hands more times.

Figure 11.2.3 shows how the velocity of money has evolved over time. Notice that the velocity of M1 is higher than the velocity of M2. Recall from the definitions of M1 and M2 that M2 includes more things than M1. Using (11.2.4), this implies that the velocity of M2 must be lower.

One assumption that people sometimes make is that V is constant. Figure 11.2.3 shows that this is not completely justified, since velocity has moved around quite a bit over time. Furthermore, a standard model of the money demand says that velocity should not be expected to remain constant: when interest rates rise, money demand falls and therefore velocity rises.

Nevertheless, sometimes it is useful to assume that velocity can be held constant in an "other things being

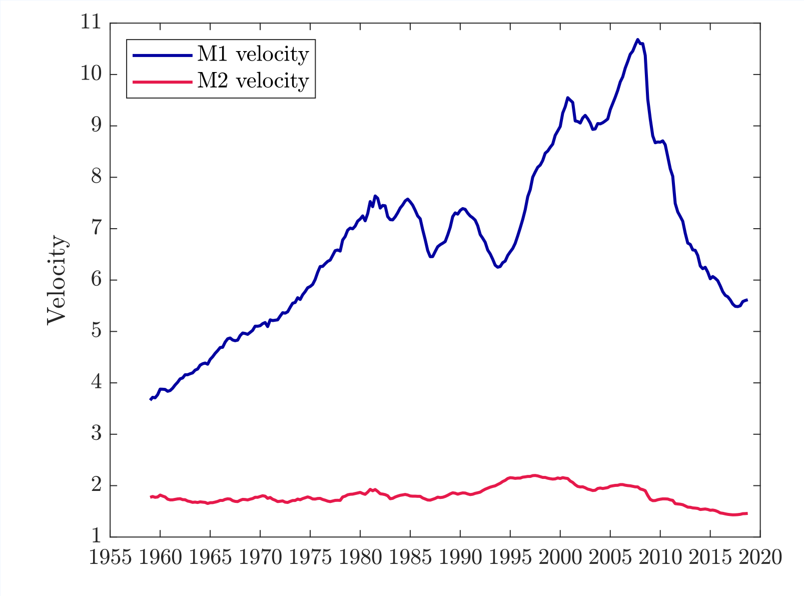


Fig. 11.2.3: *The velocity of money in the US, using M1 and M2 as definitions of money. Source: Federal Reserve Bank of St. Louis.*

equal" sense when one considers some other change in the economic environment. If one assumes that V is constant, then equation (11.2.4) changes from being a *definition* to being a *theory*. In fact, it is sometimes known as the "quantity theory", because it says that that the price level will be exactly proportional to the quantity of money.

The evidence from Figure 11.2.1 is sometimes interpreted as supportive of the quantity theory, since the quantity theory implies that there should be an exact linear relationship between changes in the money supply and changes in prices. Notice that the relationship between money growth and inflation becomes much closer for countries with high inflation. Even if velocity is not exactly constant, compared to the scale of changes in the money supply in those countries it doesn't move that much, so the quantity theory is not such a bad approximation.

11.3 Seignorage

Nowadays, most countries tend to keep inflation quite low, though usually not at zero. We'll look at some of the arguments in favor of positive inflation later on. Historically, one of the reasons why inflation has sometimes been high is that governments used inflation to obtain seignorage.

The term seignorage, which derives from the French word for lord ("seigneur"), originally referred to the profit made in the production of coins. Back when coins were usually made of precious metals, the value of a minted coin was typically above the value of the metal used to produce it. Why? Because minted coins were better money than raw metal since they were standardized to be useful in transactions. The profit earned by the mint by turning metal into money was known as seignorage. Nowadays the term is used more broadly to refer to the resources obtained thanks to the ability to create money.

When we looked at the process of money creation in Chapter 10, we didn't pay too much attention to the Central Bank's balance sheet, but if you go back to it, you'll notice that, in the process of increasing the monetary base by Δ, the Central Bank increased both its assets and its liabilities by Δ. This doesn't seem like a big deal, but there is one important difference. The assets that the Central Bank obtains (government bonds) earn interest while the liabilities that it issues (currency and reserves) typically do not. Standard accounting still treats them as liabilities but in a certain sense they are not. Furthermore, the Central Bank is just a branch of the government.[3] Therefore increasing the monetary base is a way for the government to get a loan that it will never have to repay and doesn't pay any interest. Indeed, one way to write down the government's budget constraint is:

$$B_{t+1} + p_t \tau_t + \left[M_{t+1}^B - M_t^B\right] = p_t G_t + (1 + i_t) B_t \tag{11.3.1}$$

where:

- B_{t+1} is nominal public debt,
- G_t is real government spending,
- τ_t is real government tax revenue,
- p_t is the price level,
- i_t is the nominal interest rate,
- M_t^B is the monetary base.

Let's go through the terms in (11.3.1) to see what it means. The right hand side represents all the payments the government must make, in nominal terms. $p_t G_t$ is how much the government must pay for the current period's spending. $(1 + i_t) B_t$ is how much it must pay on the debts it had at the beginning of the period, including the interest that accrued in the current period. The left hand side represents all the resources the government can use to make its payments. $p_t \tau_t$ is how much it raises in taxes, in nominal terms. $M_{t+1}^B - M_t^B$ is the increase in the monetary base. This is all the payments the government can make just by virtue of having created extra money. B_{t+1} is the amount of payments the government can make by virtue of issuing debt that will have to be repaid in the future.

We can also rearrange (11.3.1) to express it as:

$$M_{t+1}^B + B_{t+1} = p_t \left[G_t - \tau_t\right] + (1 + i_t) B_t + M_t^B$$

This formulation makes it easier to see that the monetary base is just like debt (in the sense that it enters the government budget in the same way), except that it doesn't pay interest.

Historically, governments have used expansion of the monetary base as a way to satisfy the government budget in various circumstances. Sometimes it's a result of difficulties in collecting regular taxes, due to tax

[3] This is literally true in some countries and sort-of-true in others. In the US, the Federal Reserve has a mixed governance structure with some influence from the private sector. However, it rebates its profits to the Treasury, so in that sense it's part of the Federal Government.

evasion or political indecision about what other taxes to use. Sometimes it's a result of a rapid increase in government spending that leaves no time to increase regular taxes, as in wartime. Sometimes it's the result of the inability to borrow, perhaps because lenders don't trust the government to pay back its debts. In other instances it may have been to a misperception that increasing the monetary base is a way for the government to obtain resources without really taking them away from anyone. Often it could be several of these reasons at the same time.

One subtle question is who exactly the government is taxing when it increases the monetary base. It's clear that the government can, at least to some extent, pay for goods and services with monetary expansion. But nobody seems to be paying for this. How does it all add up? The answer is that anyone who holds money is implicitly paying a tax to the government when the monetary base expands. We know that in a money market equilibrium, an expansion of the monetary base leads (through the money multiplier) to an increase in the money supply and then to an increase in the price level. Anyone who holds the monetary base *while it's losing value* is implicitly giving up some of their wealth to the government, just as they would if they were paying a regular tax. That's why seignorage revenue is also sometimes referred to as an inflation tax.[4]

Governments usually limit how much seignorage revenue they try to raise because they want to avoid creating high inflation. But even if they didn't care about inflation there is a limit to how much seignorage revenue they can obtain. To obtain high revenue, they need to expand the monetary base very fast. But fast expansion of the monetary base leads to high inflation, which leads to high nominal interest rates, which means that the real money demand falls. Since implicitly seignorage is a tax on money holdings, this means that the tax base shrinks. Exercise 11.6 asks you to work out the limit on seignorage revenue and compare it to some historical experiences of very high inflation.

11.4 The Cost of Inflation

Inflation is generally seen as undesirable. There are several reasons for this.

One reason can be understood directly from the Baumol-Tobin model of money demand. Other things being equal, more inflation implies higher nominal interest rates, which means that people will "go to the bank" more times to avoid holding high money balances. Each of those trips to the bank has a cost of F.[5] Using (10.4.3) we can compute the total cost of trips to the bank as:

$$\begin{aligned} Cost &= NF \\ &= \sqrt{\frac{iYF}{2}} \\ &= \sqrt{\frac{(r+\pi)YF}{2}} \end{aligned} \tag{11.4.1}$$

At the times of high inflation in Argentina in the late 1980s, my dad would literally go to the bank twice a day, once around noon and once after work just to make sure that he had exactly enough money for the day's

[4] Technically, this is not quite right. The government can obtain seignorage revenue with zero inflation if the economy is growing. Exercise 11.3 asks you to work out how much.

[5] Sometimes this is known as "shoe leather cost": people wear out their shoes by walking to the bank all the time.

expenses and no more. That time spent dealing with the problem of how much money to hold has a real opportunity cost.

On the basis of a reasoning like this, Milton Friedman advocated keeping nominal interest rates at or very near zero, a policy known as the "Friedman Rule". The idea of the Friedman Rule is to eliminate the opportunity cost of holding money. In formula (11.4.1), having $i = 0$ would make the cost equal to zero, because it would mean that you don't ever need to go to the bank: since it has no opportunity cost, you can just hold all your wealth in money. Notice that in order to have $i = 0$, one would need to have $\pi = -r$. Since the real interest rate is usually positive, this means that implementing the Friedman Rule requires deflation. The Friedman Rule is usually seen as a theoretical extreme, more valuable for the underlying logic than as a concrete policy proposal.

Economists sometimes refer to "menu costs" as part of the cost of inflation. Sometimes there are real resources that need to be dedicated to put in place a change in prices. For instance, restaurants need to print new menus, shops need to print new signs, etc. When inflation is high this needs to be done more often, which is a real cost. Minimizing menu costs would require keeping inflation at zero, rather than running deflation as implied by the Friedman Rule. Even this would not eliminate menu costs: zero inflation means that the price index would stay constant, but the prices of individual goods would still move up and down a lot, and making those changes would incur menu costs. There is some disagreement among economists about whether it's plausible that menu costs are large.

Another cost of inflation is that it creates uncertainty about relative prices. In order to decide what to buy, people need to know the prices of different goods. But they don't look at all the prices of all the goods at the same time. Typically, a consumer just looks at the price of a few different goods at a time and relies on his knowledge of approximately how much stuff you can get for a dollar to assess whether prices of the goods he is considering are worth paying. High inflation makes it harder to keep track of how much a dollar is worth, which makes it harder to make the right consumption decisions. It's a little bit like trying to take measurements with a ruler that keeps changing size. Producers face the same problem: in order to decide what price to charge for the goods they sell, they need to keep track of how much a dollar is worth, and inflation makes this harder.

The government is not the only one to earn seignorage. Part of the money supply is made up of bank deposits, which either pay no interest or pay less than market rates, so they are earning seignorage as well. An additional source of costs of inflation is that it allows banks to earn more seignorage, leading to excessive entry into the banking industry. Exercise 11.7 asks you to compute how much seignorage banks earn.

Exercises

11.1 **The Elasticity of Money Demand**

Recall the Baumol-Tobin model of Chapter 10. Compute η, the elasticity of money demand with respect to GDP, in the money-demand function that arises from this model.

11.2 **The Quantity Theory**

The quantity theory of money assumes that money velocity is constant. Try to come up with a money-demand function that would imply that velocity is indeed constant. What does this money-demand function say about how money holdings depend on interest rates? What does it say about how households manage their trips to the bank?

11.3 **Seignorage with Zero Inflation and Growth**

Suppose an economy is growing at a constant rate g, money demand comes from the Baumol-Tobin model and the money multiplier is ω. Suppose the government wants to maintain zero inflation. How much seignorage revenue can it obtain as a fraction of GDP? In other words, find an expression for the level of:

$$\frac{M_t^B - M_{t-1}^B}{p_t Y_t}$$

that is consistent with zero inflation. How does this depend on g, F and ω? Why? [Hint: use the result from Exercise 11.1.]

11.4 **Growth in the Money Supply**

There are two otherwise identical economies. In both of them GDP is constant, prices are flexible and the real interest rate is constant. The rate of growth of the money supply in each of them is:

	Growth rate of the money supply	
Year	**Country A**	**Country B**
2009	3%	6%
2010	3%	5%
2011	3%	4%
2012	3%	3%
2013	3%	2%
2014	3%	1%
2015	3%	0%

How does inflation in the two countries in 2012 compare? Explain.

11.5 **Inflation Targeting**

(a) Suppose a Central Bank has decided it wants to keep inflation at exactly 2% every year. The economy is not growing and the real interest rate is 3%. Describe what the Central Bank must do to the money supply to achieve its objective.

(b) Suppose the real interest rate suddenly and unexpectedly falls from 3% to 1%. How should the Central Bank respond? Explain in detail.

11.6 **Seignorage with High Inflation**

Suppose that money demand is given by:

$$m^D\left(Y_t, i_{t+1}\right) = Y_t \cdot \left(1 + i_{t+1}\right)^{-\mu}$$

where μ is a parameter. This is sometimes known as the Cagan money-demand function. Equilibrium in the money market is given by:

$$M_t^S = Y_t \cdot (1 + i_{t+1})^{-\mu} p_t$$

The real economy is in a steady state so that:

$$Y_t = Y_{ss}$$
$$r_{t+1} = r_{ss}$$

The government obtains seignorage revenue S_t (in real terms) by expanding the monetary base. S_t is given by:

$$S_t = \frac{M_t^B - M_{t-1}^B}{P_t} \tag{11.4.2}$$

Let ω be the money multiplier, so that $M^S = \omega M^B$.

Suppose that the rate of growth of the money supply is constant at rate γ, i.e.:

$$M_t^S = (1 + \gamma) M_{t-1}^S \tag{11.4.3}$$

(a) What will be the rate of inflation?

(b) What will be nominal interest rate?

(c) Find an expression for the level of real money balances. How does it depend on γ? Why?

(d) Solve equation (11.4.3) for M_{t-1}^S and replace this in (11.4.2) to obtain an expression for seignorage revenue S_t in terms of γ and $\frac{M_t^S}{P_t}$. How does S_t depend on γ? How does it depend on $\frac{M_t^S}{P_t}$? Why?

(e) Replace the value of $\frac{M_t^S}{P_t}$ that you found in part (c) into the expression for S_t that you found in part (d) to obtain an expression for seignorage revenues S_t in terms of γ, Y_{ss}, r_{ss}, ω and μ.

(f) Assume the following parameter values:

$$Y_{ss} = 1$$
$$\mu = 4$$
$$r_{ss} = 0.0025$$
$$\omega = 5$$

Note that $\mu = 4$ is close to the value that has been empirically estimated in some high inflation countries when the time period is one month. To be consistent, r_{ss} should be interpreted as a monthly real interest rate and Y_{ss} as monthly GDP.

(g) Plot S_t against γ for values of γ between 0 and 0.8.

(h) What is the monthly rate of growth of the money supply that maximizes seignorage revenue? Call this γ^*.

(i) Why is seignorage revenue decreasing in γ for $\gamma > \gamma^*$?

(j) What monthly inflation rate does $\gamma = \gamma^*$ imply?

(k) What monthly nominal interest rate does $\gamma = \gamma^*$ imply?

(l) What yearly rate of inflation does $\gamma = \gamma^*$ imply?

(m) What is $\frac{M_t}{P_t}$ if $\gamma = \gamma^*$?

(n) What is the maximum revenue as a fraction of GDP that the government in this example can constantly collect from seignorage?

(o) Look at the data from Sargent (1982) on the German hyperinflation of the 1920s (you can find it at the book website). What was the maximum monthly inflation rate that Germany experienced? How does it compare to γ^*?

(p) Suppose that the government is increasing the money supply at a rate $\gamma = \gamma^*$, and in month t it credibly announces that (i) M_{t+1}^B will not be equal to $M_t^B(1+\gamma^*)$ the way it would have been if the policy had continued as usual, but it will be some other level $\bar{M}$ (maybe higher, maybe lower than $M_t^B(1+\gamma^*)$) (ii) from then on, it will set $M_{t+s}^B = \bar{M}$ for every $s \geq 1$ (i.e. the money supply will be constant).

(q) What will be the rate of inflation from period $t+1$ onwards?

(r) What will be the price level P_{t+1} in period $t+1$? (This will depend on the government's choice of $\bar{M}$)

Suppose the government sets $\bar{M}$ at the level that will ensure that $P_{t+1} = P_t$ (i.e. the level that will immediately stop inflation).

(s) What value of $\bar{M}$ will it need to choose? How does it compare to $M_t^B(1+\gamma^*)$?

(t) How much seignorage revenue will the government obtain in period $t+1$? How does this level of seignorage compare to what the government was getting every month before making this change?

(u) Look carefully at the timing of the end of hyperinflation in Germany? When exactly did the money supply stop growing? When exactly did hyperinflation stop? How can we make sense of this?

11.7 **Bank Seignorage**

Suppose that the demand for M1 money is given by $\frac{M}{pY} = A^{-\eta i}$ where M is the quantity of M1 money, pY is nominal GDP, i is the nominal interest rate and A and η are parameters. Households want to hold a fraction χ of their M1 money in the form of cash and a fraction $1-\chi$ in the form of checking accounts, which earn no interest. Banks earn seignorage by taking checking deposits and investing in assets that earn the nominal interest rate.

(a) Find an expression for the ratio of total seignorage earned by banks to GDP. Call this ratio s.

(b) Compute $\frac{\partial s}{\partial i}$. Why does this number depend on η?

(c) Look up data on M1 and its components in 2018. What is a reasonable value for χ? What was the value of $\frac{M}{pY}$? If the average nominal interest was 2%, how much seignorage did banks earn as a fraction of GDP?

(d) If $\eta = 0.2$, how much seignorage would banks earn as a fraction of GDP if the nominal interest rate went up to 3%?

11.8 **Real Interest Rates**

Define **realized** real interest rates r_{t+1} as:

$$r_{t+1} \equiv i_{t+1} - \pi_{t+1}$$

where i_{t+1} is the nominal interest rate between period t and period $t+1$ and π_{t+1} is the rate of inflation between period t and period $t+1$. If there is uncertainty about what the rate of inflation is going to be, this implies that there is uncertainty about what realized real interest rates will be. Define the **expected** real interest rate as:

$$r^E_{t+1} \equiv i_{t+1} - \mathbb{E}\left(\pi_{t+1}\right)$$

where $\mathbb{E}\left(\pi_{t+1}\right)$ is the expected rate of inflation.

Suppose there are two parties to a nominal lending contract: a borrower and a lender. Who benefits when **realized** real interest rates turn out to be higher than **expected** real interest rates? What monetary policy will each of them support?

11.9 **Money among Prisoners of War**

Read "The Economic Organisation of a P.O.W. Camp", which you can find at: https://www.jstor.org/stable/2550133. Pick one passage out of the article and explain how it relates to the models of money supply, money demand and inflation from Chapters 10 and 11.

PART V

Business Cycles

This part of the book looks at business cycles, relatively short-term movements in the aggregate economy.

In Chapter 12 we look at patterns in business cycle data to establish some of the basic facts of what business cycles look like. Then we turn to attempts at explaining why business cycles happen and study versions of two of the leading theories: real business cycle models and Keynesian models. We'll do them in reverse historical order. The real business cycle model was first proposed in the late 1970s and early 1980s, in part as a result of some economists' dissatisfaction with earlier Keynesian models. However, the model is a bit simpler so we do it first, in Chapter 13, leaving the New Keynesian model for Chapter 14.

Finally, in Chapter 15, we take a look at some of the policies that are used to try to manage the business cycle.

CHAPTER 12

Facts about Business Cycles

12.1 What are Business Cycles?

There is no unique definition of exactly what is meant by a "business cycle". Here is one traditional definition, proposed by Burns and Mitchell (1946):

> Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle.

Figure 12.1.1 shows the evolution of GDP in the US since 1947.

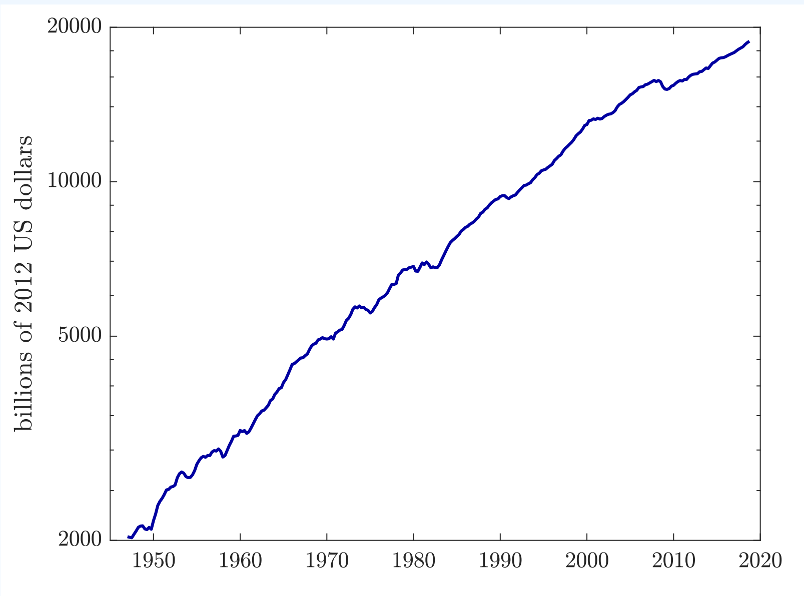


Fig. 12.1.1: *Real GDP in the US since 1947. Source: NIPA.*

As we know, the general trend is upwards, but it's far from a straight line. Often the term "business cycle" is used to refer to the wiggles in the trajectory of GDP and the movements in other economic variables that accompany them. The word "cycle" itself is a little bit misleading, since it evokes a regular, periodic, oscillation like that of a sine curve. The evolution of GDP is not like that, it's irregular. Sometimes GDP grows at a fairly steady rate for a long time; other times it expands rapidly and then falls steeply.

The terms "recession" and "expansion" are often used in the context of describing business cycles. Expansions are periods when GDP is growing; recessions are periods when GDP is shrinking. Sometimes a recession is defined more precisely as a period when GDP shrinks for two consecutive quarters, but not everyone adheres to that definition.

The NBER Business Cycle Dating Committee

There is a committee within the National Bureau of Economic Research dedicated to the task of declaring when recessions and expansions begin and end. The end of an expansion/beginning of a recession is called a "peak" and the end of a recession/beginning of an expansion is called a "trough". They don't have an exact rule of how they determine peaks and troughs (if they did, one wouldn't need a committee!) and look at a broad range of indicators, not just GDP. At some level, this labeling exercise is a bit absurd: there is more to be learned by looking at the entire data than by just having the labels "recession" and "expansion". On the other hand, it is sometimes useful to have a simple classification of which way economic activity is headed. Figure 12.1.2 shows the same data for GDP as Figure 12.1.1 with the NBER-designated recessions shaded in gray.

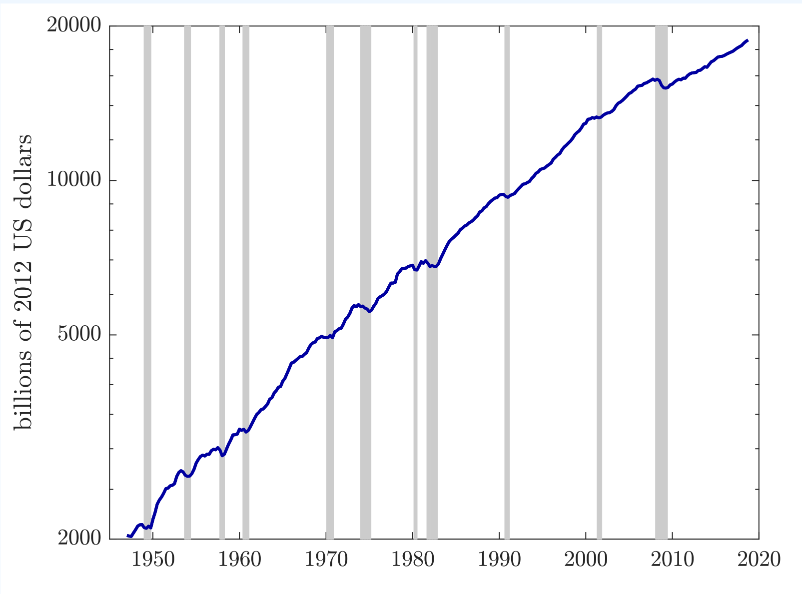


Fig. 12.1.2: *Real GDP in the US since 1947 and NBER recessions. Source: NIPA and NBER.*

The Hodrick-Prescott Filter

Hodrick and Prescott (1997) propose an algorithm for distinguishing a "cycle" from a "trend" in economic data. The idea is to separate out the long-run growth (the "trend") of any economic variable, such as GDP, from the shorter term deviations around that trend, which we will label a "cycle".

Suppose we observe a variable X_t from period $t = 1$ until period $t = T$. We are going to define an artificial variable $\hat{X}_t$ and call it the "trend" in X_t. We are going to want the trend to have the following properties:

1. It cannot be too far away from the actual variable X_t.
2. It has to move smoothly, i.e. the rate of growth in the trend should not change very much from one period to the next.

Mathematically, we are going to define the trend as the solution to the following problem:

$$\min_{\{\hat{X}_t\}_{t=1}^{T}} \sum_{t=1}^{T} \underbrace{\left(X_t - \hat{X}_t\right)^2}_{\substack{\text{Distance between trend} \\ \text{and actual variable}}} + \lambda \sum_{t=2}^{T-1} \underbrace{\left[\left(\hat{X}_{t+1} - \hat{X}_t\right) - \left(\hat{X}_t - \hat{X}_{t-1}\right)\right]^2}_{\substack{\text{Change in the trend's} \\ \text{growth rate}}} \tag{12.1.1}$$

where λ is a parameter. Figure 12.1.3 shows the trend in GDP using $\lambda = 1,600$, which is a standard value for quarterly data. Trend GDP ends up being a smoother version of actual GDP.

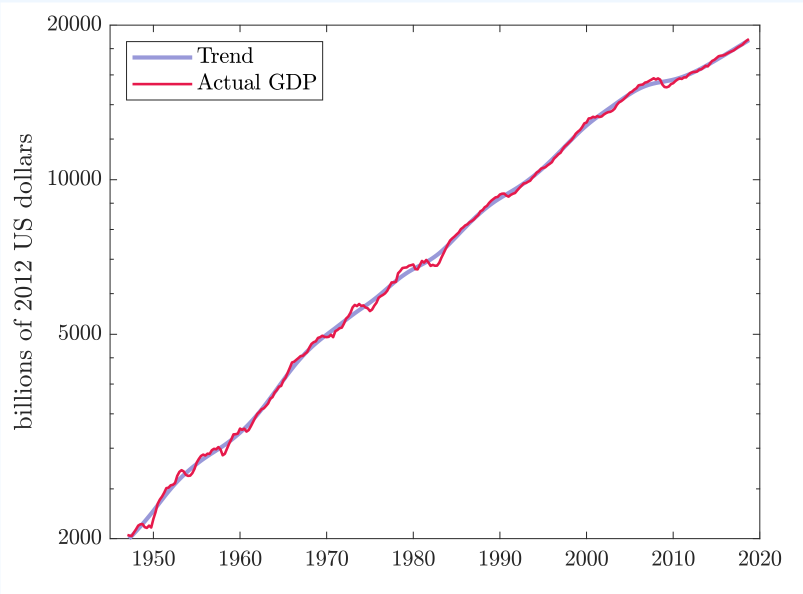


Fig. 12.1.3: *Real GDP in the US since 1947 and HP trend.*

Once we define a trend, the "cycle" or "cyclical component" is simply defined as

$$\tilde{X}_t \equiv X_t - \hat{X}_t$$

i.e. as the deviation of the variable from its trend. Figure 12.1.4 shows the cyclical component of GDP (expressed as a percentage of trend GDP, i.e. $\frac{\tilde{X}_t}{\hat{X}_t}$), compared again with NBER-defined recessions. The figure shows that what the NBER committee determines is not that different from what HP-filtering does: NBER-defined recessions are periods when then cyclical component of GDP moves down.

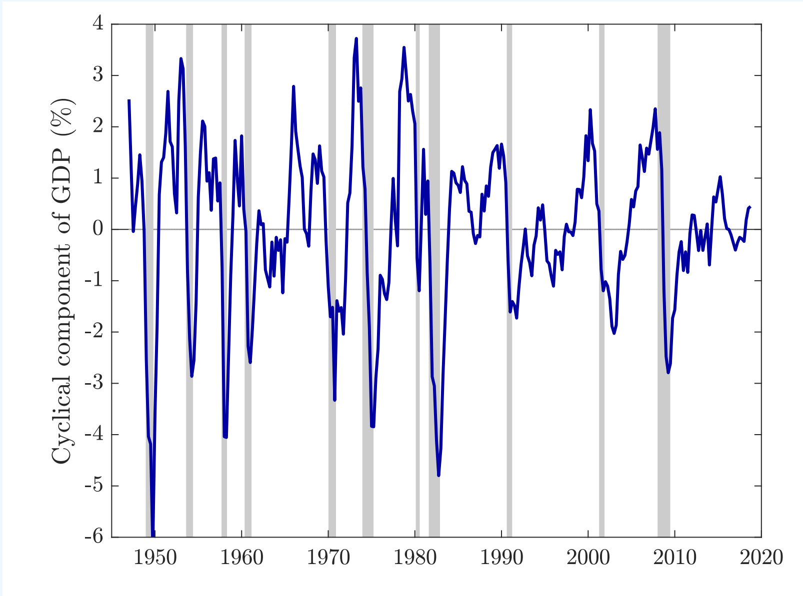


Fig. 12.1.4: *Cyclical component of real GDP in the US since 1947 obtained with HP filter.*

Note that different values of λ will result in different definitions of what the trend is and therefore different definitions of what the cyclical component of GDP is. Figure 12.1.5 shows the cyclical component of GDP for different values of λ. In expression (12.1.1), a high value of λ penalizes changes in the trend growth rate very heavily. As a result, the HP filter will make the "trend" close to a straight line and, as a result, allow the cyclical component of GDP to be large. Conversely, a low value of λ will result in a "trend" that changes quite a bit in order to stay very close to actual GDP. As a result, the implied cyclical component will be small.

This can matter. For instance, the standard value of λ implies that after the recession of 2008-2009 GDP returned to trend fairly rapidly and was back at trend by 2012 approximately. Mathematically, the reason is that after several years of slow growth, the HP filter infers that the trend has slowed down, so actual GDP is catching up to trend despite slow growth. In a sense, this level of λ imposes a limit on how much a recession can really last. Conversely, under a higher value of λ, the procedure insists that the "trend" continues to grow at close to its long-term average, so actual GDP continues to fall behind trend. We are going to use the HP filter with the standard value of λ to systematically describe the patterns that we observe in business cycles but it's worth bearing in mind that there is quite a bit of judgment going into how we construct the definition of the business cycle.[1]

[1]Hamilton (2018) argues that these and other problems make the HP filter completely useless.

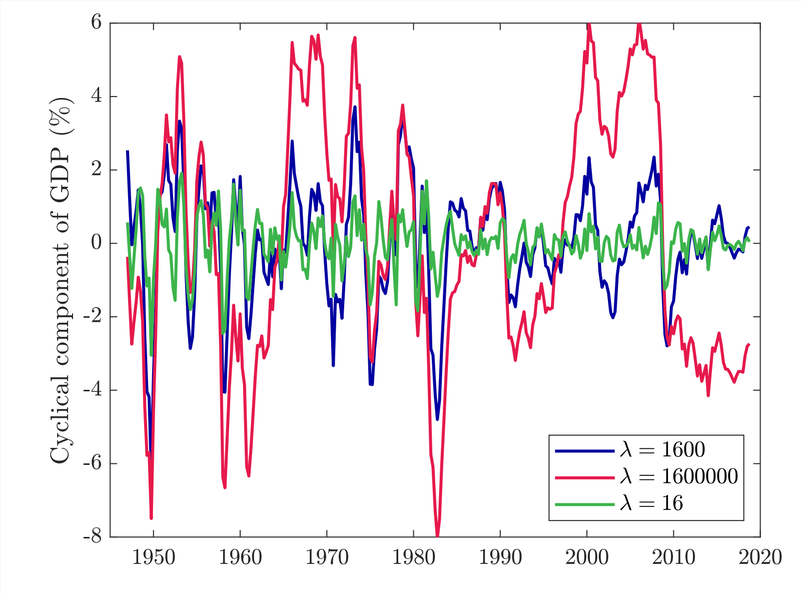


Fig. 12.1.5: *Cyclical component of real GDP in the US since 1947 obtained with HP filter for different values of λ.*

12.2 Patterns in Business Cycles

The Post-War US Business Cycle

Table 12.1 shows some of the patterns displayed by the US business cycle since 1947. It shows the cyclical patterns of several interesting macroeconomic variables. For each variable X_t, we follow these steps:

1. HP-filter the data and subtract the trend to compute the cyclical component $\tilde{X}_t$.

2. Compute the standard deviation of $\tilde{X}_t$.[2] This gives us a sense of how far away from its trend the variable X_t tends to be.

3. Compute the correlation between $\tilde{X}_t$ and the cyclical component of $\log(GDP)$.[3] This gives us a sense of whether X_t tends to move together with GDP, in the opposite direction, or with an unrelated pattern.

---

[2] The standard deviation of a variable X_t is defined as follows. T is the number of observations. $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$ is the mean. $Var(X) = \frac{1}{T}\sum_{t=1}^{T}\left(X_t - \bar{X}\right)^2$ is the variance. $\sigma_X = \sqrt{Var(X)}$ is the standard deviation.

[3] The correlation between two variables X and Y is defined as follows. $Cov(X,Y) = \frac{1}{T}\sum_{t=1}^{T}\left(X_t - \bar{X}\right)\left(Y_t - \bar{Y}\right)$ is the covariance. $\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$ is the correlation. It takes values between -1 and 1. $\rho_{XY} = 1$ means X and Y are exactly aligned; $\rho_{XY} = -1$ means X and Y are exactly aligned but move in opposite directions; $\rho_{XY} = 0$ means that movements in X and movements in Y go in the same direction and the opposite direction just as much.

Table 12.1: *Business cycle properties of macroeconomic variables. Y, C, I, G, M, X, total hours, TFP and real wages are measured in log scale so the units are comparable. All variables are detrended using an HP filter with $\lambda = 1,600$. Sources: NIPA for GDP and its components; BLS for labor market data including wages and for inflation; Fernald (2014) for TFP; Board of Governors of the Federal Reserve System for interest rates.*

Variable	Standard deviation	Relative standard deviation	Correlation with GDP
GDP	1.6%	1	1
Consumption	1.2%	0.74	0.78
Durable Goods	4.7%	2.93	0.60
Non-durable Goods	1.5%	0.95	0.58
Services	0.8%	0.54	0.57
Investment	7.3%	4.59	0.83
Government spending	3.3%	2.06	0.16
Exports	5.3%	3.31	0.42
Imports	5.0%	3.12	0.72
Total hours of work	1.8%	1.13	0.85
TFP (Solow residual)	1.3%	0.78	0.80
Real wages	0.7%	0.47	0.31
Unemployment rate	0.8%		−0.86
Inflation	3.2%		0.26
Nominal Interest Rate	1.1%		0.38

We are going to focus on the following facts that can be gathered from Table 12.1:

1. It is typical for GDP to be about 1.6% away from trend (either above or below).

2. Total consumption is less volatile than GDP. If we break it down into categories, consumption of durables is more volatile than GDP and consumption of nondurables and especially services less so.

3. Investment is much more volatile than GDP.

4. Consumption, investment, productivity, and hours of work are all highly positively correlated with GDP. Unemployment is highly negatively correlated with GDP.

5. Real wages are less volatile than GDP and only weakly positively correlated.

6. Inflation and nominal interest rates have weak positive correlations with GDP.

Figure 12.2.1 shows the cyclical components of GDP, total consumption, investment and total hours. They move up and down together, with investment moving much more than GDP, consumption a little bit less and hours about the same amount.

For now, we'll take these as the main facts about business cycles which we are tying to understand. We'll come back to some of the other patterns documented on Table 12.1 later on.

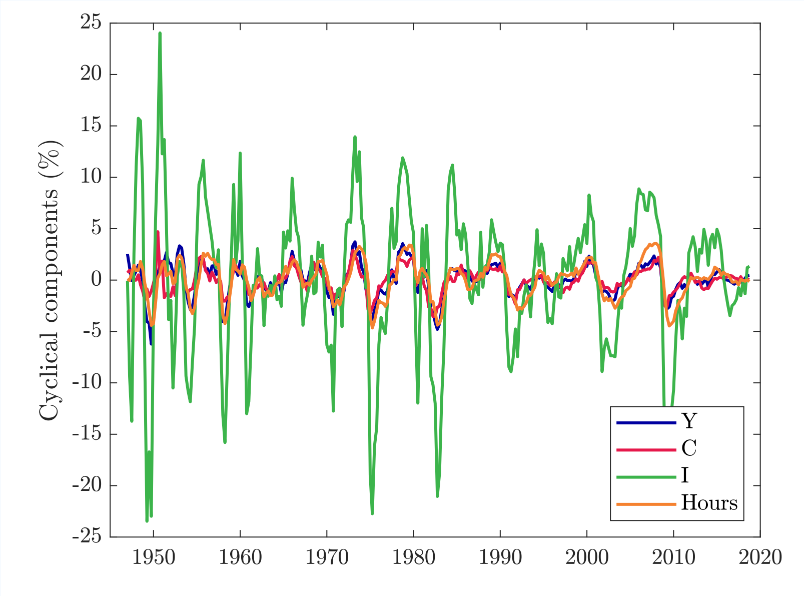


Fig. 12.2.1: Cyclical component of real GDP, consumption, investment and total hours.

The Phillips Curve

Phillips (1958) documented a relationship between between the rate of nominal wage increases and the unemployment rate in the UK between 1861 and 1957. He found that times of low unemployment were also times of fast wage growth. This pattern came to be known as the "Phillips Curve". There are several, slightly different, versions of the Phillips Curve. They all relate some measure of nominal price changes with some measure of economic activity such as unemployment or GDP growth. Probably the most standard version these days has inflation on the vertical axis and unemployment on the horizontal axis. Another version has the cyclical component of GDP on the horizontal axis instead of unemployment. Figure 12.2.2 shows how that relationship looks like for the US.

Overall, there is a mild negative relationship between inflation and unemployment. The slope of -0.34 on the left panel means that 1 percentage point higher unemployment is associated with 0.34 percentage points lower inflation. However, the association between the variables is quite noisy, their correlation is just -0.36. Similarly, the slope of 0.59 on the right panel means that 1% higher GDP is associated with 0.59 percentage points higher inflation. Again, the association is quite noisy, with a correlation of just 0.29. Since unemployment and the cyclical component of GDP are so highly correlated (as shown in Table 12.1), the two panels tell the same story: business cycle expansions are associated with higher inflation, but only mildly.

If we break down the relationship into different subperiods, as shown in Figure 12.2.3, an interesting pattern emerges. Until the mid-1960s, we observe the negative Phillips curve relationship. It somewhat noisy, but it's clearly visible. Between the mid-1960s and the late 1970s the relationship seems to break down, and inflation and unemployment become positively associated. Then, from the late 1970s until the mid-1980s, the relationship is again negative, and very strong. Finally, since the mid-1980s, the relationship is negative again, but very weak.

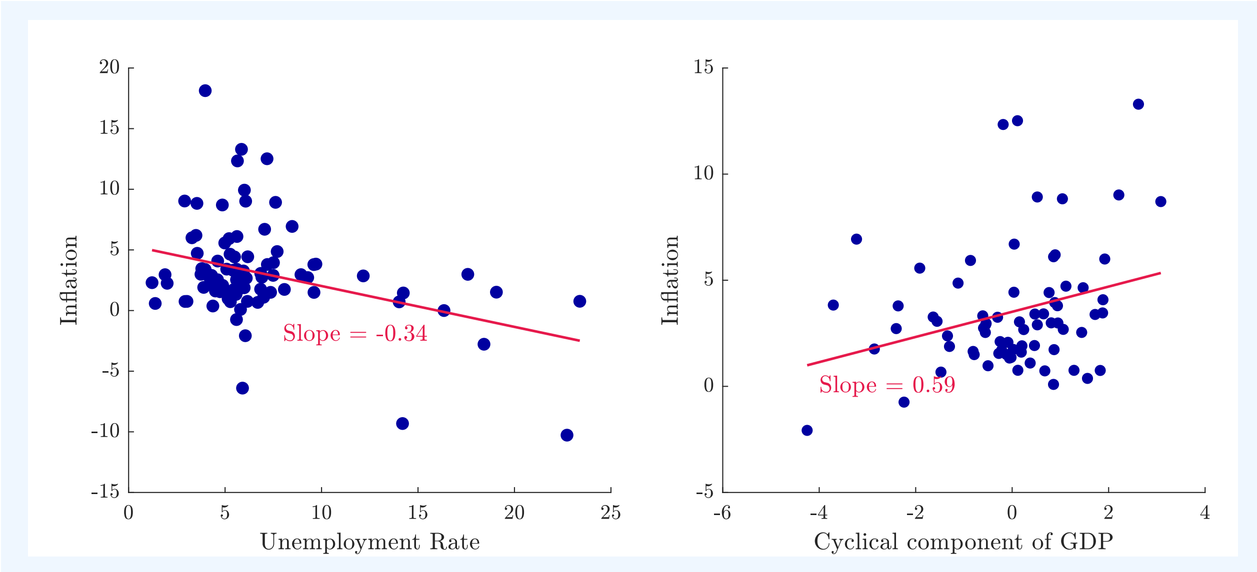


Fig. 12.2.2: *The Phillips Curve relationship in the US. Annual data for 1929-2018 on the left panel; annual data for 1947-2018 on the right panel. Source: NIPA and BLS.*

One of the features of business cycles that we going to try to understand is why the Phillips Curve relationship sometimes seems to hold and other times seems to not hold.

The Great Depression

Much of macroeconomics was originally motivated by trying to understand the Great Depression of the 1930s. Figure 12.2.4 shows some facts about what happened in the 1930s in the US.

Between its peak in 1929 and its trough in 1933, real GDP fell by about 27%, investment fell by more than 80%, unemployment rose from about 5% to over 20% and there was deflation, with prices falling by about 27%. Qualitatively, the Great Depression is consistent with the typical patterns of business cycles:

1. investment moved in the same direction as GDP but much more,

2. unemployment rose at the same time as GDP fell,

3. deflation coincided with high unemployment, as the Phillips Curve stipulates.

However, the magnitude of the movements was *much* higher than in typical business cycles: a typical recession will have GDP a couple of percentage points below trend; in the Great Depression the fall in GDP was a staggering 27%. How could something like this happen? What should be done to prevent it from happening again?

At a more theoretical level, an important question is whether the Great Depression was just another business cycle but much larger or whether it was a fundamentally different phenomenon.

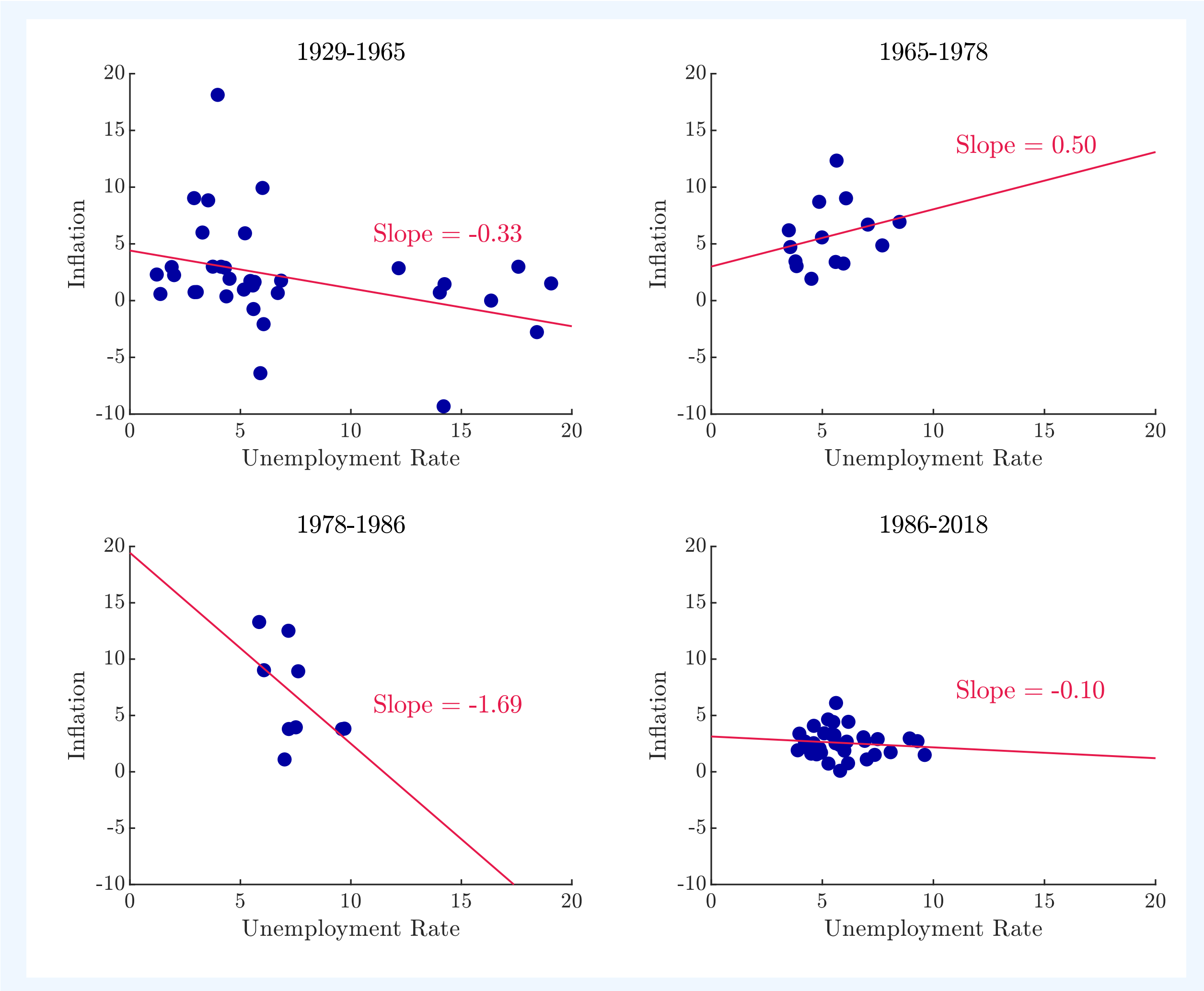


Fig. 12.2.3: Cyclical component of unemployment and CPI inflation in the US. Source: BLS.

12.3 Who Cares about the Business Cycle?

There has been an enormous amount of research into how the business cycle works. Part of the objective of this research is to find ways to make the economy more stable. Lucas (1987) asked the following question: suppose we could figure out a way to eliminate the business cycle altogether, how valuable would that be? He proposed answering the question by doing a version of the following calculation.

Suppose the representative household in the economy has standard preferences over consumption, given by:

$$\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}$$

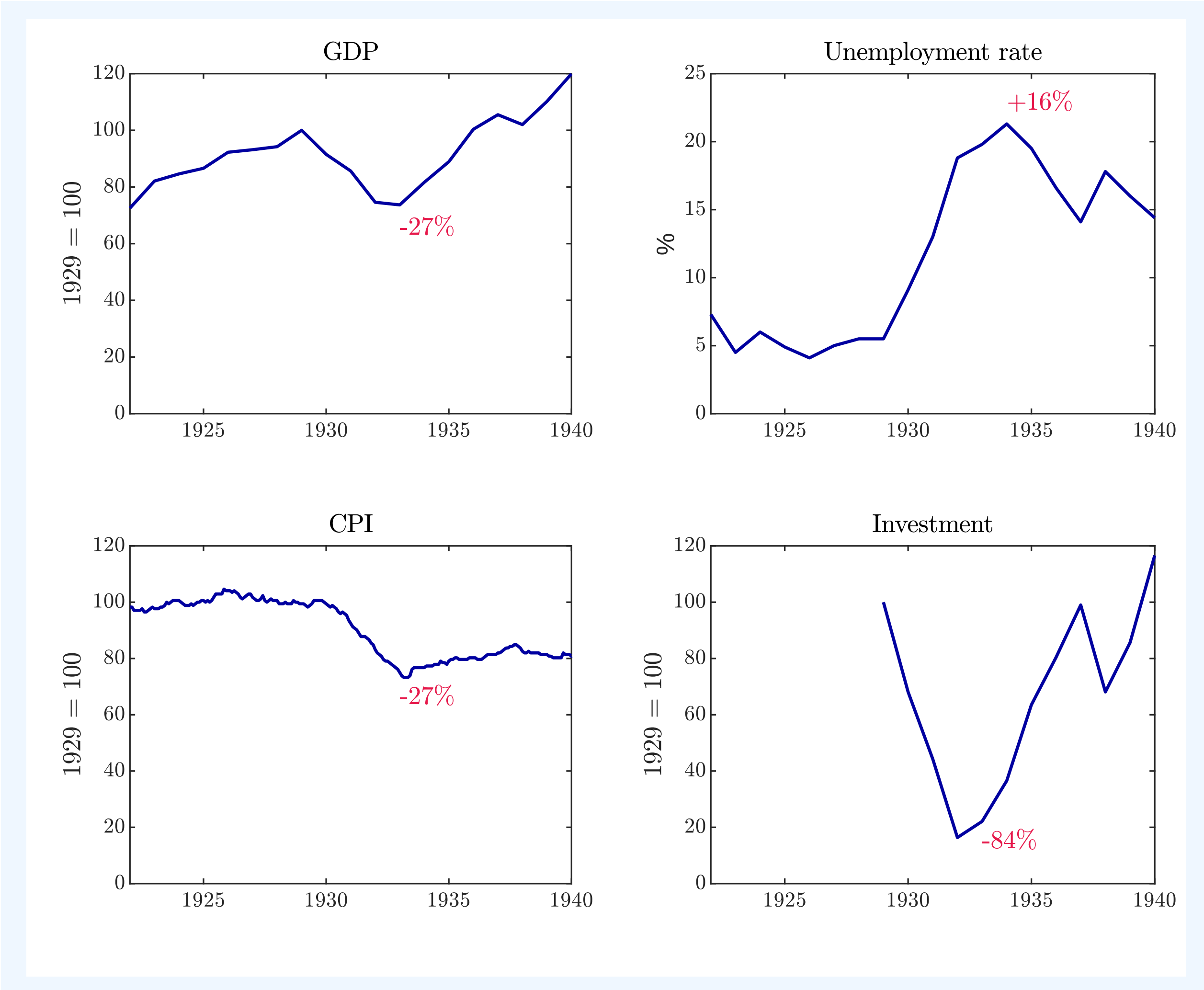


Fig. 12.2.4: The US economy during the Great Depression. Sources: Johnston and Williamson (2019) and NIPA for GDP, NIPA for investment, BLS for CPI and Smiley (1983) for unemployment.

Imagine that *trend* consumption is given by :

$$\hat{c}_t = c_0 (1 + g)^t$$

so that it grows at a constant rate g after starting at the level c_0 at some initial period 0. *Actual* consumption can differ from the trend consumption because of the business cycle. We are going to have the following representation of the business cycle. Half the time, the economy is in an expansion period so consumption is above trend and half the time the economy is in a recession with consumption below trend:

$$c_t = \begin{cases} \hat{c}_t\left(1+f\right) & \text{in an expansion} \\ \hat{c}_t\left(1-f\right) & \text{in a recession} \end{cases} \tag{12.3.1}$$

where f is a parameter that measures the amplitude of the business cycle.

We can compute the utility that the representative household will experience as a function of initial consumption c_0, the growth rate (denoted g) and the magnitude of business cycle fluctuations f:

$$\begin{aligned} W\left(c_0, g, f\right) &= \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} u\left(c_0\left(1+g\right)^t\left(1+f\right)\right) + \frac{1}{2} u\left(c_0\left(1+g\right)^t\left(1-f\right)\right)\right] \\ &= \frac{1}{2}\frac{c_0^{1-\sigma}}{1-\sigma}\left[\left(1+f\right)^{1-\sigma}+\left(1-f\right)^{1-\sigma}\right]\sum_{t=0}^{\infty}\beta^t\left[\left(1+g\right)^t\right]^{1-\sigma} \\ &= \frac{1}{2}\frac{c_0^{1-\sigma}}{1-\sigma}\frac{\left(1+f\right)^{1-\sigma}+\left(1-f\right)^{1-\sigma}}{1-\beta\left(1+g\right)^{1-\sigma}} \end{aligned}$$

In the first line, we are applying a version of the behind-the-veil-of-ignorance argument we first encountered in Chapter 2. In each period, the household might find itself either in an expansion or in a recession. The household computes how much utility it's going to experience in each case and then takes an average.

In order to compute the value of eliminating the business cycle, we are going to solve for λ in the following equation:

$$W\left(\lambda c_0, g, f\right) = W\left(c_0, g, 0\right) \tag{12.3.2}$$

What's the logic of equation (12.3.2)? It defines λ as the answer to the following question. Suppose someone offered two possibilities to the representative household: either multiply its consumption by some number λ each period (the left hand side) or keep average consumption the same but, by making $f = 0$, eliminate the business cycle (the right hand side). What is the value of λ that would make the household indifferent? Let's solve:

$$\frac{1}{2}\frac{\left(\lambda c_0\right)^{1-\sigma}}{1-\sigma}\frac{\left(1+f\right)^{1-\sigma}+\left(1-f\right)^{1-\sigma}}{1-\beta\left(1+g\right)^{1-\sigma}} = \frac{1}{2}\frac{c_0^{1-\sigma}}{1-\sigma}\frac{2}{1-\beta\left(1+g\right)^{1-\sigma}}$$

$$\lambda = \left(\frac{1}{2}\left(\left(1+f\right)^{1-\sigma}+\left(1-f\right)^{1-\sigma}\right)\right)^{\frac{1}{\sigma-1}} \tag{12.3.3}$$

Now let's plug in some actual numbers into (12.3.3). From Table 12.1 we know that the standard deviation of the cyclical component of $\log\left(c_t\right)$ is 1.2% of its trend level, so let's set $f = 0.012$. As we have seen before, there is a range of estimates for the value of σ, so we'll try a couple of different ones. Table 12.2 shows the results:[4]

[4] As we saw in Chapter 2, for $\sigma = 1$ we replace the utility function with $\log\left(c\right)$. Re-doing the steps that lead to (12.3.3) results in:

$$\lambda = \left(\left(1+f\right)\left(1-f\right)\right)^{-\frac{1}{2}}$$

Table 12.2: *The value of eliminating the business cycle, according to the Lucas (1987) calculation.*

σ	1	2	5	10
λ	1.00007	1.00014	1.00036	1.00072

The values of λ are all greater than 1. This means that the representative household views the elimination of the business cycle as equivalent, in utility terms, as an increase in consumption. The reason is that the household dislikes the variability in consumption that the business cycle entails, so it attains higher utility if this variability can be eliminated.

However, the the gain is tiny. Take, for instance, $\sigma = 2$. For this value we get $\lambda = 1.00014$. This means that the household is indifferent between eliminating the business cycle and obtaining an increase in consumption of 0.014%. To put in in dollar terms, consumption per person per year in the US is about $40,000; according to this calculation, eliminating the business cycle would be worth just $5.80 per year to the average person. Even for higher values of σ, which make the household dislike variability more, the numbers are still very small. For $\sigma = 10$, eliminating the business cycle is still worth only $28.80 per year to the average person.

Reasons why the Business Cycle Could Matter More

If Lucas's calculation is correct, then the business cycle is simply not a big deal, and we should probably devote a lot less intellectual energy to understanding it and a lot less policy effort to stabilize it. However, the calculation leaves out a lot of things, and some have argued that taking these into account would significantly alter the conclusion.[5]

One reason why stabilizing the business cycle might be more valuable than Lucas's calculation suggests is the way recessions interact with individual-level risk. Lucas's calculation assumes that there is a representative household whose consumption moves up and down exactly with aggregate consumption. If instead a lot of the volatility in aggregate consumption is concentrated in certain households (for instance, the ones that are prone to losing their jobs in recessions and experiencing long spells of unemployment), then the value *for these particular households* of stabilizing the business cycle would be much larger, and the overall behind-the-veil-of-ignorance value could be higher as well. Exercise 12.3 asks you to consider this possibility.

Lucas's calculation assumes that the best one could hope for with stabilization policy is to make consumption stable without changing the average level. Underlying this is a view that business cycles represent movements up and down around some normal level. An alternative view is that the peak of the business cycle is the "normal" level for the economy and business cycles are *downwards* deviations from this normal level. If this view is right, then in principle the ideal policies could make the economy always remain at its peak, raising average consumption as well as making it more stable, as illustrated in Figure 12.3.1. According to standard preferences, this is much more valuable.

As formula (12.3.3) makes clear, the value of eliminating business cycles depends on f, i.e. on how large these business cycles were to begin with. By historical and international standards, the postwar US economy was relatively stable, which results in a small value of f. Lucas's conclusion that *further* stabilization is not that valuable for an economy like the US does not imply that a very volatile economy would not benefit from

[5] Barlevy (2005) surveys many of these arguments.

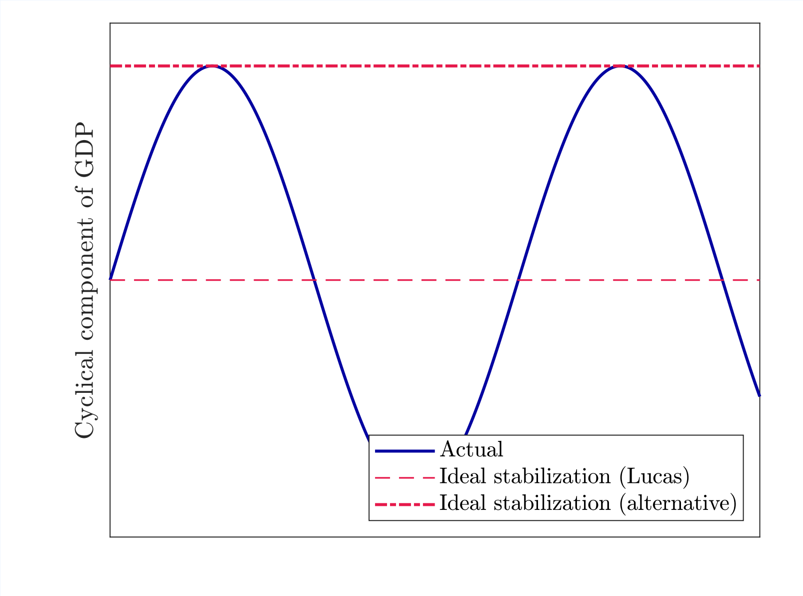


Fig. 12.3.1: *Different views on what stabilizing the business cycle might be able to achieve.*

policies that make it as stable as the US. Exercise 12.2 asks you to consider the value of stabilizing an economy where recessions of the magnitude of the Great Depressions happen often.

Finally, there might be factors that are not well captured by our models (the political repercussions of high unemployment, the stress of not knowing whether one's business will survive in a recession) that might make business cycles more important than what is captured by equation (12.3.3).

Exercises

12.1 **Business Cycles in Other Countries**

(a) Look up GDP accounts for a country other than the US. Download quarterly data for GDP, consumption and investment.

(b) Apply the HP filter to create trend and cyclical components. To do this, first transform the data by taking logarithms, i.e. compute $\log(Y)$, $\log(C)$ and $\log(I)$. Then apply formula (12.1.1) to compute a trend component for each log series. If you use Excel, you can find an add-in to do this at https://web-reg.de/webreg-hodrick-prescott-filter/; if you use Matlab, the command hpfilter does it for you. Finally, compute the cyclical component by subtracting the trend from the unfiltered log series.

(c) Compute the standard deviation of the cyclical component of GDP, consumption and investment.

(d) Compute the correlation between the cyclical components of consumption and investment with the cyclical component of GDP.

(e) How do the patterns compare with the US? What is similar and what is different?

12.2 **Stabilizing a Very Volatile Economy**

We are going to redo Lucas's calculation for an economy that keeps undergoing the Great Depression all the time. Look up the NIPA GDP data for 1929 (the peak before the Great Depression) and 1933 (the trough of the Great Depression). Imagine an economy where half the time consumption is at the level of 1929 and half the time consumption is at the level of 1933. What is the value for the representative household in this economy of complete consumption stabilization at the average level (which, incidentally, is pretty close to the level of consumption in 1931)? Try the following values of σ: 2, 5 and 10.

12.3 **Heterogeneity and the Value of Stabilization**

Imagine an economy that spends half the time in a recession and half the time in an expansion. In this economy there are two kinds of households: stable and volatile. Stable households consume c every period, so they are immune to the business cycle. Volatile households consume:

$$c^{Volatile} = \begin{cases} c\,(1+v) & \text{in an expansion} \\ c\,(1-v) & \text{in a recession} \end{cases}$$

Suppose that a fraction μ of the households in the economy is volatile and a fraction $1-\mu$ is stable.

(a) What is total consumption in the economy, in an expansion and a recession respectively?

(b) Suppose *aggregate* consumption is well described by equation (12.3.1). Given a value of μ, what value of v would make aggregate consumption have fluctuations of amplitude equal to f?

For the remaining questions, assume $\sigma = 5$, $f = 0.012$ and $\mu = 0.05$.

(c) What is the value for a stable household of eliminating the business cycle?

(d) What is the value for a volatile household of eliminating the business cycle?

(e) What is the value of eliminating the business cycle for a household that, behind the veil of ignorance, is not sure whether they are going to be stable or volatile?

12.4 **The Value of Growth**

Lucas (1988) argued:

> Is there some action a government of India could take that would lead the Indian economy to grow like Indonesia's or Egypt's?[6] If so, what exactly? If not, what is it about the "nature of India" that makes it so? The consequences for human welfare involved in questions like these are simply staggering: once one starts to think about them, it is hard to think about anything else.

In this exercise we'll try to make sense of what Lucas had in mind.

[6] This was back before the India started to grow fast. India's GDP per capita grew less than 1% per year in the 1970s, just over 3% per year in the 1980s and 1990s and over 5% per year since 2000.

(a) Solve for λ in the following equation:

$$W(\lambda c_0, g, f) = W(c_0, g', f)$$

with $g = 0.01,\ g' = 0.02,\ f = 0.01$ and $\sigma = 2$.

(b) What does the number you found represent?

CHAPTER 13

The Real Business Cycle Model

One of the leading theories of why business cycles happen is the so-called Real Business Cycle model. It has this name because, unlike the Keynesian-inspired theories we'll look at later, money and nominal variables play no role in it. The idea of the RBC model is to take the same model we used to think about long run growth and apply it to questions about the short run. In particular, we are going to maintain the assumption that all markets are competitive.

We'll proceed as follows. First, we'll write down a simplified version of our general equilibrium economy, which only contains the minimal ingredients that are needed to think about short-run problems. Second, we'll introduce "shocks", exogenous changes in some aspect of the economy. We'll then work out how the endogenous outcomes in the model economy (employment, output, etc.) react to these exogenous shocks. Finally, we'll compare the patterns that emerge from the economy reacting to shocks to the empirical patterns we found in Chapter 12 to assess whether we have a plausible theory of business cycles.

13.1 A Two-Period Model

We are going to look at an even-more-simplified version of the two-period, general-equilibrium economy we studied in Chapter 9.

We are going to assume that the economy starts off having no capital and the production function in period 1 only uses labor:

$$Y_1 = F_1(L)$$

Why make this assumption? At any point in time, the capital stock is the result of past decisions and cannot be changed, so we keep things simple by just not having it in the model.

Conversely, we are going to assume that in the second period, the production function only uses capital, i.e.:

$$Y_2 = F_2(K)$$

Why the difference? Remember, we are trying to understand behavior in period 1. The capital stock in period

2 is the result of investment decisions taken in period 1, so it's important to think about what governs these decisions. We don't have labor in the period-2 production function in order to have one less object to think about.

Since this economy has competitive markets and no externalities, we know that the First Welfare Theorem holds. Therefore, in order to find the competitive equilibrium we can just solve the problem of a fictitious social planner. We'll do that first and then we'll go back to thinking about how markets attain this outcome.

The planner solves:

$$\begin{gathered}\max_{c_1,c_2,l,L,K} u(c_1)+v(l)+\beta u(c_2)\\ s.t.\\ c_1+K=Y_1\\ Y_1=F_1(L)\\ c_2=F_2(K)\\ L=1-l\end{gathered} \tag{13.1.1}$$

Just like we had in Chapter 9, the planner's objective function is to maximize the utility of the representative household. The household gets utility from consumption and leisure in period 1 and from consumption in period 2, discounted. Since we have not included period-2 labor in the model, we also don't model the household's preferences for period-2 leisure. The first constraint is the GDP accounting identity: in a closed economy with no government, total output must be split between consumption and investment, and since the economy starts with no capital, investment and the period-2 capital stock K are the same thing. The second constraint says that output results from the production function, which uses labor. The third constraint says that in the second period all output is consumed (since it's the end of the world, there is no point in investing for later). The last constraint just says that the household's total time is divided between labor and leisure.

The first order conditions for this planner's problem are familiar:

$$\frac{v'(l)}{u'(c_1)}=F_1'(L)$$
$$u'(c_1)=\beta F_2'(K)\,u'(c_2)$$

These are just like equations (9.1.12) and (9.1.13) from Chapter 9. The first describes the tradeoff between consumption and leisure in period 1 and the second describes the tradeoff between consumption in period 1 and consumption in period 2. In both cases, the planner equates the household's marginal rate of substitution to the marginal rate of transformation.

From a mathematical point of view, we now have a system of 6 equations (4 constraints and 2 first order conditions) and 6 unknowns (Y_1,c_1, c_2, l, L and K). Solving this system of equations will tell us everything that's going to happen in our model economy. At some abstract level, we are done: we have succeeded in reducing an economic question into a mathematical question. We could in principle just solve this system of equations in a computer and it would tell us what's going to happen to the economy. However, we don't just want to compute a solution, we want the model to teach us some conceptual lessons as well. For this, we are going to have to do a bit more work.

We'll start by simplifying the system of equations a little bit by substituting out c_2 and l using the constraints to get:

$$Y_1 = F_1(L) \tag{13.1.2}$$

$$\frac{v'(1-L)}{u'(c_1)} = F_1'(L) \tag{13.1.3}$$

$$u'(c_1) = \beta F_2'(K)\, u'(F_2(K)) \tag{13.1.4}$$

$$Y_1 = (u')^{-1}\left[\beta F_2'(K)\, u'(F_2(K))\right] + K \tag{13.1.5}$$

Now we have a system of 4 equations in 4 unknowns: GDP Y_1, consumption c_1, investment K and total employment L. This system of equations can be a little bit hard to interpret, so we'll start by looking at them graphically. Notice that by setting up the equations this way, we have only two of the four endogenous variables showing up in each equation. For instance, equation (13.1.3) tells us about the relationship between c_1 and L, taking as given the values of all exogenous parameters. This makes it possible to plot each of these equations in a simple graph to help interpret what each of them says. Figure (13.1.1) shows each of these four relationships.[1]

The top left graph shows the production function (13.1.2). This shows a positive relationship between employment L and total output Y_1. This is simply due to the fact that producing more output requires more labor.

The bottom left graph shows the consumption-labor tradeoff equation (13.1.3). This defines a *negative* relationship between consumption c_1 and employment L. What does this mean in economic terms? The household must, at the margin, be indifferent between dedicating a unit of time to leisure or to market production. A higher level of L means that the marginal unit of time dedicated to production: (i) produces *less* output (due to diminishing marginal product of labor, i.e. lower $F_1'(L)$) and (ii) costs *more* in utility terms (due to diminishing marginal utility of leisure, i.e. higher $v'(1-L)$). Therefore the household will only be willing to supply this extra labor if the marginal value of the output it obtains is higher. Due to diminishing

[1] It sometimes helps a little bit to use specific functions to get a little more sense of what's going on, so let's set $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, $v(l) = -\theta\frac{(1-l)^{1+\epsilon}}{1+\epsilon}$ $F_1(L) = L^{1-\alpha}$ and $F_2(K) = K^{\alpha}$. This results in

$$Y_1 = L^{1-\alpha}$$
$$\frac{\theta L^{\epsilon}}{c_1^{-\sigma}} = (1-\alpha) L^{-\alpha}$$
$$c_1^{-\sigma} = \beta\alpha K^{\alpha-1-\sigma\alpha}$$
$$Y_1 = \left[\beta\alpha K^{\alpha-\alpha\sigma-1}\right]^{-\frac{1}{\sigma}} + K$$

If we set $\alpha = 0.5$, $\theta = 1$, $\beta = 1$ and $\sigma = \epsilon = 1$ this further simplifies to:

$$Y_1 = L^{0.5}$$
$$c_1 = \frac{1}{2} L^{-1.5}$$
$$c_1 = 2K$$
$$Y = 3K$$

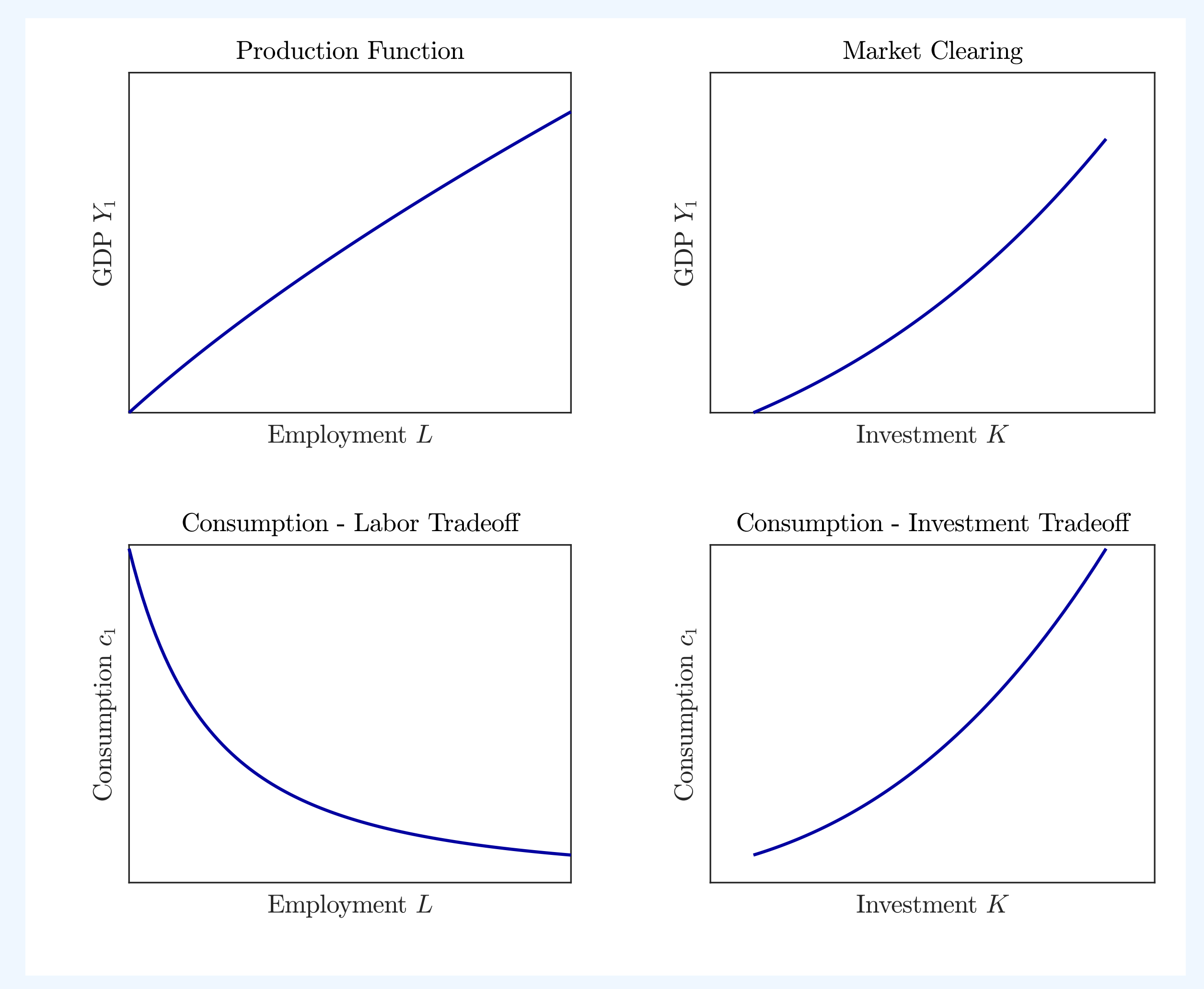


Fig. 13.1.1: *Equilibrium conditions in the RBC model.*

marginal utility, $u'(c_1)$ can only be higher if c_1 is lower. In other words, *other things being equal*, households will choose to work more if they feel poorer and so place a higher value on marginal consumption. Changes in the household's wealth will produce movements *along* the curve: a household that is wealthier chooses more consumption and less labor, and vice versa. Changes in the production technology will produce *shifts* of the curve: if the marginal rate of transformation between time and consumption changes, the household will choose a different amount of consumption to go along with any given amount of labor.

The bottom right graph shows the consumption-investment tradeoff equation (13.1.4). This defines a *positive* relationship between consumption c_1 and investment K. What does this mean? The household must, at the margin, be indifferent between dedicating a unit of output to consumption or to investment. A higher level of K means that the marginal unit of output dedicated to investment: (i) produces *less* period-2 output (due to diminishing marginal product of capital, i.e. lower $F_2'(K)$) and (ii) produces *less* period-2 marginal utility per unit of output (due to diminishing marginal utility of consumption, i.e. lower $u'(c_2)$). Therefore the household will only be willing to dedicate this extra output to investment if the marginal utility of present consumption is also lower. Due to diminishing marginal utility, $u'(c_1)$ can only be lower if c_1 is higher. In other words, *other things being equal*, households will choose to invest more only if they are also consuming more.

Finally, the top right corner shows the relationship between investment and output implied by equation (13.1.5). This is derived from the GDP accounting identity / market-clearing condition for period-1: $Y_1 = c_1 + K$. We know from (13.1.4) that there is a positive relationship between consumption and investment. Solving (13.1.4) for c_1 we obtain:

$$c_1 = (u')^{-1}\left[\beta F_2'(K)\, u'(F_2(K))\right]$$

and then replacing this in the market clearing condition $Y_1 = c_1 + K$ we obtain (13.1.5). This defines a *positive* relationship between GDP Y_1 and investment K. If investment is higher, the positive relationship between consumption and investment implies that consumption-plus-investment will be higher too. If households want to invest more and consume more, this is only possible if they also produce more.

We are going to be interested in solving the system of equations (13.1.2) - (13.1.5) in order to figure out how much output, consumption, investment and employment there's going to be in our model economy, and how these change in response to different things. There is more than one way to solve these equations. Figure 13.1.2 shows a way to solve the equations graphically.

1. Start from the top-left panel which shows the production function (13.1.2). If we guess some level for GDP Y_1, this immediately implies a level of employment L needed to produce this amount.

2. Now move to the bottom-left panel, which shows the consumption-labor tradeoff (13.1.3). Given the level of employment L, satisfying this condition implies a level of consumption c_1.

3. Now move to the bottom-right panel, which shows the consumption-investment tradeoff (13.1.4). Given the level of consumption c_1, satisfying this condition implies a level of investment K.

4. Finally circle back to see if our guess for GDP was correct by looking up in the top-right panel what level of GDP Y_1 is consistent with K. If it coincides with our original guess, then we have found a solution. Otherwise we need to adjust our guess.

One thing to notice is that we could have gone through the graphs in a different order: instead of 1-2-3-4 we could have done, for instance, 1-4-3-2. The only question is whether we go back to the original guess at the end. If the answer is yes, then all the equilibrium conditions are satisfied.

Figure 13.1.3 shows graphically what happens if the guess for Y_1 is not correct. High Y_1 implies high L (through the production function). High L implies low c_1 (through the labor-consumption tradeoff). Low c_1 implies low K (through the consumption-investment tradeoff. Low c_1 and K imply low Y_1 (through the condition $Y_1 = c_1 + K$), so the original guess does not satisfy the equilibrium conditions.

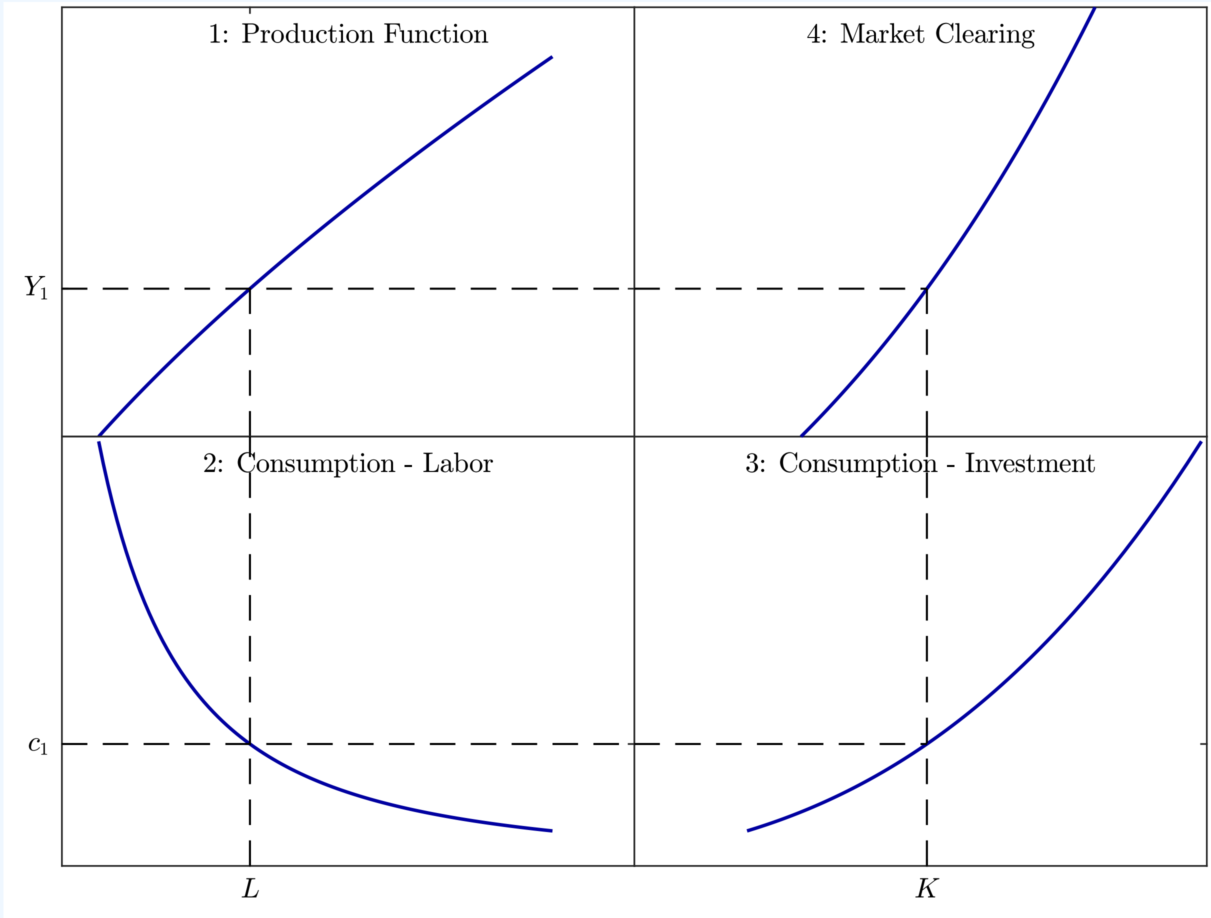


Fig. 13.1.2: *Graphical Solution of the RBC model.*

13.2 Markets

So far we have looked at our model economy as though it was administered by a social planner. The first welfare theorem justifies this approach: we know that competitive equilibrium allocations will coincide with the social planner's decisions. But we are also interested in understanding how markets bring about this outcome. In particular, we want to know what are the prices that are associated with the competitive equilibrium allocations.

These prices can be recovered from the first order conditions of the representative household and the representative firm, which we obtained in Chapter 9. Conditions (9.1.10) and (9.1.7) imply that the wage level must be equal to the marginal product of labor, and also to the marginal rate of substitution between consumption and leisure:

$$w = F_1'(L) = \frac{v'(1-L)}{u'(c_1)} \tag{13.2.1}$$

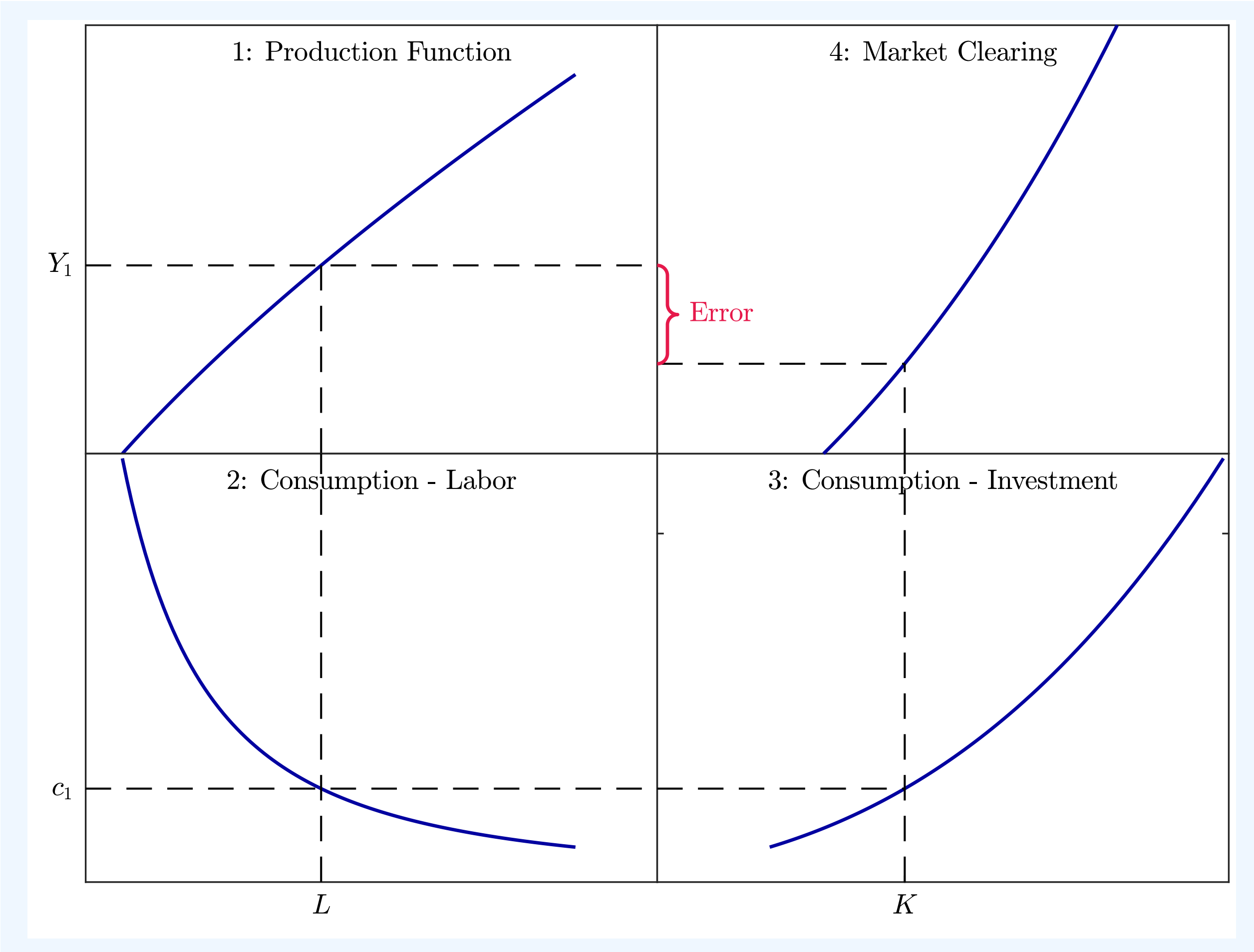


Fig. 13.1.3: *An incorrect guess for Y_1 results in an inconsistency.*

Condition (9.1.9) implies that the rental rate of capital must be equal to the marginal product of capital:

$$r^K = F_2'(K) \tag{13.2.2}$$

and since period 2 is the end of the world, it's as though we had $\delta = 1$, so conditions (9.1.11) and (9.1.13) imply that the real interest rate is:

$$1 + r = F_2'(K) = \frac{u'(c_1)}{\beta u'(c_2)} \tag{13.2.3}$$

Equations (13.2.1)-(13.2.3) allow us to think about how prices, in addition to quantities, behave in the RBC model, which will be useful both to understand how the model works and to compare it to evidence.

13.3 Productivity Shocks

A "productivity shock" is an unexpected change in the productive capacity of the economy, i.e. a change in the production function. Productivity shocks are one of the main shocks that have been proposed as possible drivers of business cycles. Let's use our model to think through how the economy would react to a productivity shock.

Let's imagine that there is an improvement in the production function, which goes from $Y_1 = F_1(L)$ to $Y_1 = AF_1(L)$, where A is some number greater than 1. This shock may represent a technological discovery, an improvement in some policy, or anything that makes the economy more productive. Let's assume that this shock is short-lived, so that there is no change in F_2.

The productivity shock affects the equilibrium conditions in two places:

$$Y_1 = AF_1(L)$$
$$\frac{v'(1-L)}{u'(c_1)} = AF_1'(L)$$
$$u'(c_1) = \beta F_2'(K)\, u'(F_2(K))$$
$$Y_1 = (u')^{-1}\left[\beta F_2'(K)\, u'(F_2(K))\right] + K$$

The shock affects the production function: for the same amount of labor, the economy gets more output, which it can distribute between consumption and investment. It also affects the labor-consumption decision because it affects the *marginal* product of labor and thus the marginal rate of transformation between time and consumption. Figure 13.3.1 shows how this affects the equilibrium.

The shock affects two of the four curves in the figure. The production function curve moves up: for any level of L, Y_1 moves up. Second, the consumption-labor curve moves to the right: dedicating one unit of time to production produces more goods than before, so *other things being equal*, the household should be willing to work more. The consumption-investment-output curves do not move, so any movement must be a shift *along* these curves.

Let's trace out the effect of this shock on the equilibrium. It will be useful to do it in the order 1-4-3-2. Start with panel 1. The most direct effect is that Y_1 rises. If we trace out the rise of Y_1 from panel 1 to panel 4, we see that higher Y_1 must necessarily imply higher investment K. Using panel 3, this implies higher consumption. If we then go to panel 2, we see that there are two opposing effects on L. First, the curve shifts to the right: holding c_1 fixed, the household would choose higher L. However, there is also a movement *along* the curve, because c_1 increases. This lowers L. In this example, the net effect is that L rises, but it's possible to construct examples where L falls.

What's going on economically? The economy has become temporarily more productive. The total effect of this on the economy is the result of the various ways in which the household wants to react to this. The representative household wants to:

1. Increase consumption, because more productivity makes the household feel wealthier.

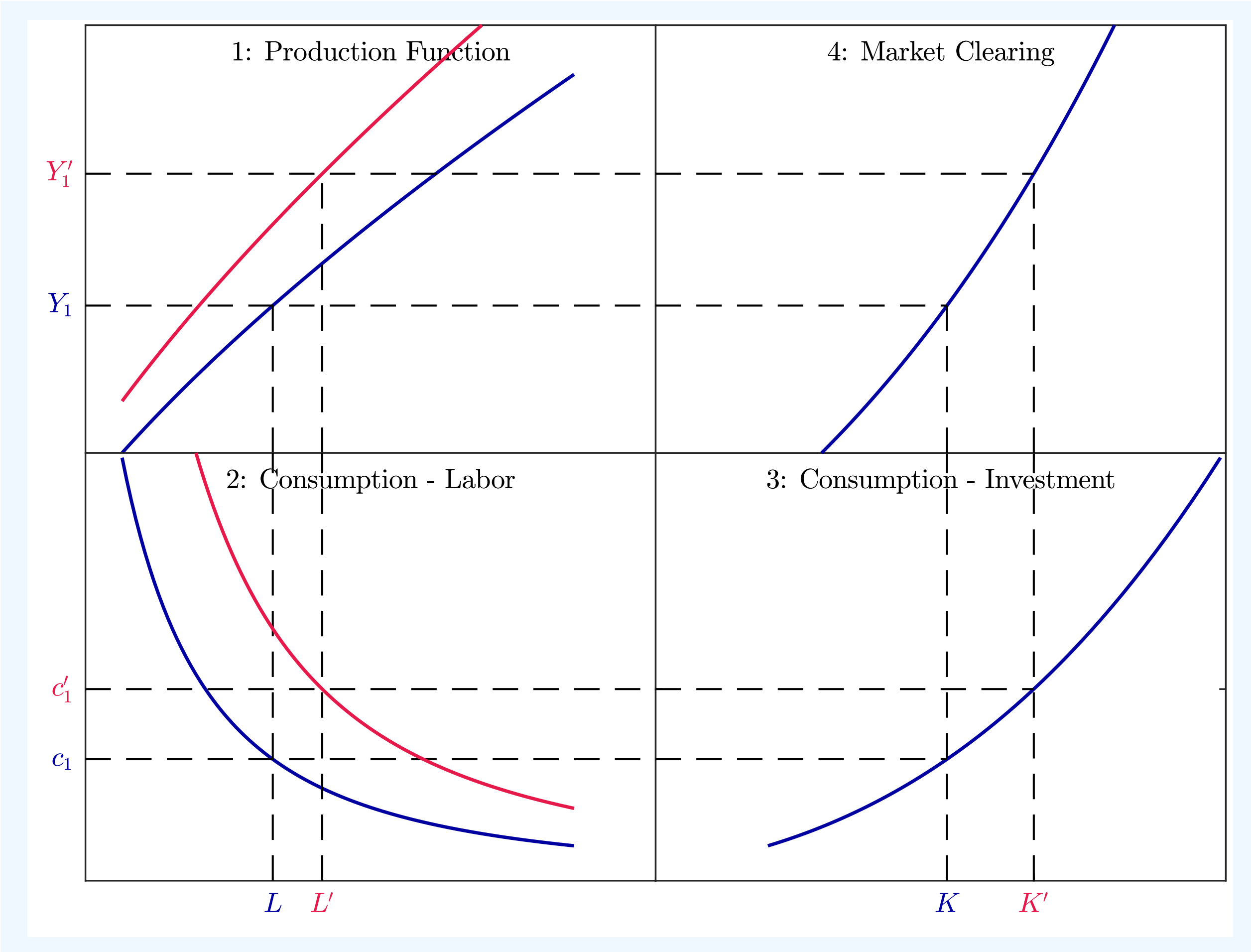


Fig. 13.3.1: *The economy's reaction to a productivity shock.*

2. Increase investment, in order to smooth out the increase in consumption between the present and the future.

3. Work more, to take advantage of the improved technology for transforming time into consumption goods.

4. Enjoy more leisure, because more productivity makes the household feel wealthier.

The first two forces are unambiguous: the household will react by increasing consumption and investment. The last two forces point in opposite directions. The substitution effect pushes the household to work more while the income effect pushes the household to work less. This is the same issue that came up in Chapter 7. If we imagine that the productivity increase is temporary, then the income effect is likely to be small and the substitution effect dominates, persuading the household to work more. This is the case depicted in Figure 13.3.1.

The RBC model with productivity shocks therefore offers us one possible account of how and why business cycles take place. If our model economy undergoes a productivity shock, then productivity, GDP, consumption, employment and investment will all move in the same direction. In this account, an economic expansion is a time when households choose to work more in response to temporarily high productivity, which, by equation (13.2.1), translates into temporarily high wages. Conversely, a recession is a time when households *choose* to work less in response to temporarily low wages.

13.4 Other Shocks

Just like we did with productivity shocks, we can also ask how the economy would react to other kinds of shocks.

Impatience Shocks

Let's imagine that households suddenly decide they really want to consume now rather than later. We'll see later that effects of this kind can be important in Keynesian models. We are going to model higher impatience as a fall in β: the value that the households place on the future falls, so they really want to prioritize the present.

In our system of equations, β enters equations (13.1.4) and (13.1.5) More impatient households place lower value on the future so, for any given level of consumption, the level of investment they choose is lower. Graphically, this means that the curves in panels 3 and 4 shift to the left.

The net effect is shown in Figure 13.4.1. Households consume more which, through equation (13.1.3), means they also work less. Less work means lower total output, and since consumption has risen, investment must fall. Overall, this does *not* look like the patterns in business cycles: consumption goes up while output, employment and investment go down. Empirically, all these variables tend to move together.

Optimism

Let's imagine that households become optimistic about the future. How would this affect the economy today? We are going to model optimism as a rise in *future* productivity, so that the production function for period 2 becomes:

$$Y_2 = AF(K)$$

where A is a number greater than 1. For our purposes, it doesn't really matter whether the optimism is justified, i.e. whether, once period 2 comes along, productivity does in fact rise. Since we are focusing on how the economy behaves in the short run, all we are interested is how the *expectation* of future productivity affects decisions in the present.

In our system of equations, A enters through the consumption-investment equation (13.1.4) and the market clearing condition(13.1.5), which contain the term:

$$\beta \quad \underbrace{AF_2'(K)}_{\text{increases with } A} \quad \underbrace{u'(AF_2(K))}_{\text{decreases with } A}$$

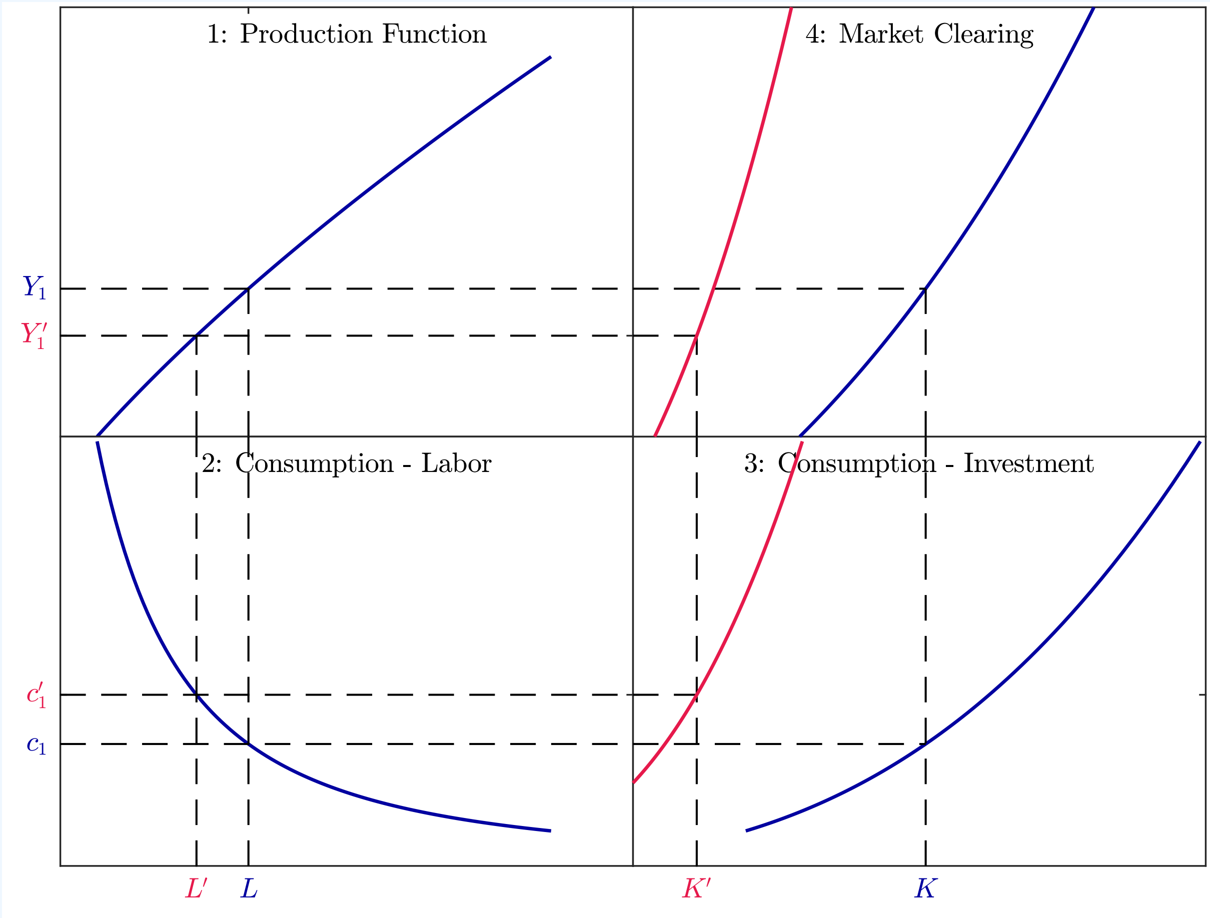


Fig. 13.4.1: *The economy's reaction to an increase in impatience.*

The net effect of A on the relationship between c_1 and K is ambiguous since A affects the consumption-investment decision in two opposite ways. First, higher period-2 productivity makes investment more attractive by raising the marginal product of capital. Other things being equal, this leads to more investment (and less consumption). But other things are not equal. For any given level of investment, higher period-2 productivity means higher period-2 consumption, and therefore lower marginal utility of period-2 consumption. This makes the household want to smooth consumption by raising period-1 consumption (and lowering investment). Figure (13.4.2) shows an example where the first effect dominates and investment rises but the net effect could go either way.

In order to liberate resources for investment, households end up consuming less. Since households are cutting back on period-1 consumption, their marginal utility of consumption rises, so households react by working more i.e. they move *along* the curve in panel 2. Since households work more, GDP rises.

This pattern captures some of the features the we observe empirically in business cycles: output, employment and investment move in the same direction. However, it misses one very important dimension:

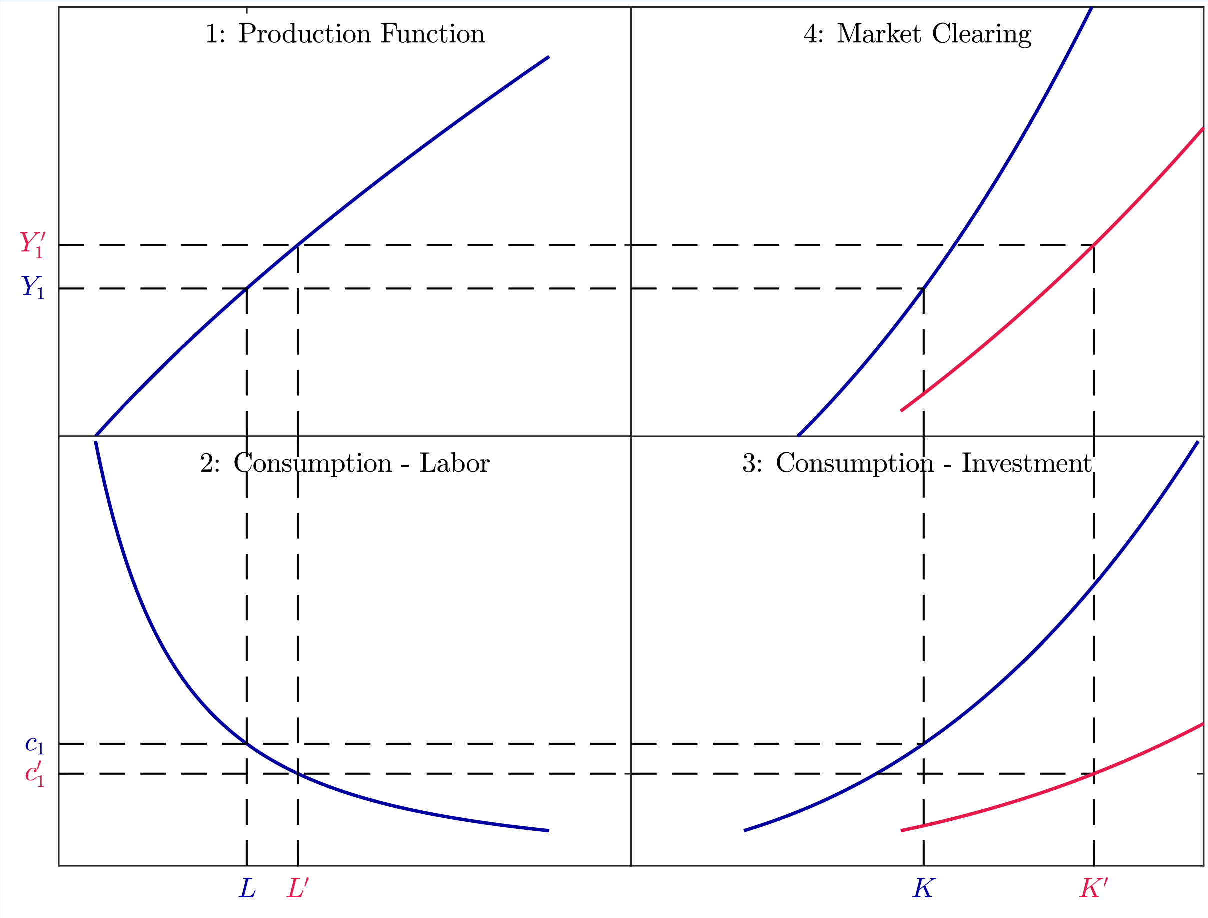


Fig. 13.4.2: *The economy's reaction to optimism about future productivity.*

consumption moves the wrong way. Since the pro-cyclicality of consumption is such a central fact of business cycles, we must conclude that this type of shock in an RBC model cannot be the main driver of business cycles.

Figure (13.4.2) focuses on a case where, other things being equal, higher A makes investment rise. Could it be that making the opposite assumption fixes the problem? No. In the case where the consumption-capital shifts to the left, consumption does indeed rise. But then movement along the labor-consumption curve implies that L falls, which means that GDP must fall, and investment falls as well. Again, we would have consumption moving in the opposite direction as everything else.

Laziness or Taxes

Let's imagine that households become lazy. More precisely, they change their relative preference for consumption and leisure. We are going to model an increased desire to enjoy leisure by saying that preferences change to:

$$u(c_1) + \theta v(l) + \beta u(c_2)$$

where θ is a number greater than 1.

In our system of equations, θ enters through the labor-consumption equation (13.1.3), which becomes:

$$\frac{\theta v'(1-L)}{u'(c_1)} = F_1'(L) \tag{13.4.1}$$

Graphically, this is a shift to the left of the labor-consumption curve in panel 2: for any given level of consumption, households want to supply less labor.

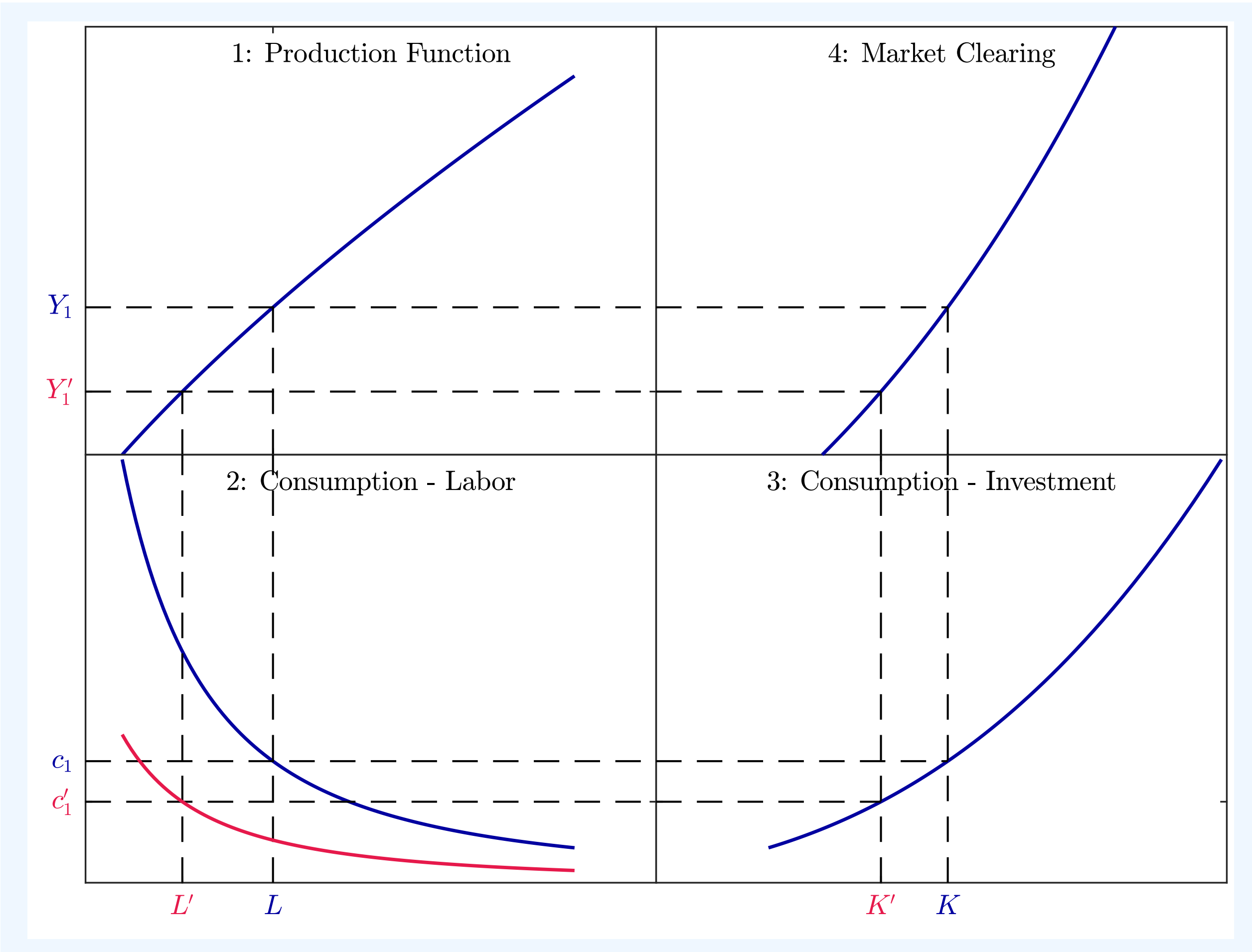


Fig. 13.4.3: *The economy's reaction to an increase in preference for leisure or an increase in income taxes.*

The net effect is shown in Figure (13.4.3). Households work less, which lowers output. This fall in output results in both a fall in consumption c_1 and a fall in investment K, i.e. a movement *along* the consumption-investment curve in panel 3.

This type of shock does produce a reaction that looks like a recession: employment, output, consumption and investment all fall together. On the other hand, the account of recessions that this type of shock leads to is not terribly satisfactory: recessions are what happens when everybody simultaneously decides that it's time to get some rest.

One alternative that is mathematically equivalent to an increase in laziness is an increase in taxes. Remember from equation (7.2.5) in Chapter 7 what happens when you introduce a labor-income tax into a consumption-leisure choice: the worker equates the marginal rate of substitution to the after-tax wage $w(1-\tau)$ rather than the pre-tax wage w. Since the representative firm is still equating the pre-tax wage to the marginal product of labor, equation (13.1.3) becomes:

$$\begin{aligned} \frac{v'(1-L)}{u'(c_1)} &= w(1-\tau) = F_1'(L)(1-\tau) \\ \Rightarrow \frac{\frac{1}{1-\tau}v'(1-L)}{u'(c_1)} &= F_1'(L) \end{aligned} \tag{13.4.2}$$

Comparing equations (13.4.1) and (13.4.2) we can see that for every laziness shock $\theta > 1$ there is a tax-rate shock $\tau > 0$ such that the effects of either of these shocks are equivalent.[2] While perhaps more appealing than pure changes in preferences, a theory of business cycles based on changes in tax rates has one empirical shortcoming: we simply don't change tax rates so often, and the changes we do make are not so strongly correlated with business cycles.

13.5 Assessing the RBC model

Why Some Kinds of Shocks Produce Business Cycles but not Others

As pointed out by Barro and King (1984), the failure of either impatience shocks or optimism to produce something that looks like a business cycle in an RBC model has a common source: the labor-consumption decision. As long as nothing changes how people trade off time against consumption goods (i.e. as long as nothing causes a *shift* of the labor-consumption curve in panel 2 of Figure 13.1.2) then anything that raises consumption must lower employment. After any change that does not affect the labor consumption tradeoff, the representative household either: (i) feels richer and thus consumes more and works less or (ii) feels poorer and thus consumes less but works more.

The reason why productivity shocks or changes in taxes could, in principle, produce a business-cycle-like reaction is that they affect how households perceive the tradeoff between dedicating time to work or to leisure. The other shocks we looked at do not affect this margin directly and therefore cannot make consumption and employment move in the same direction.

One maintained assumption in all of this analysis is that the labor market is competitive: households can choose, without any restrictions, how much labor to provide each period at the market wage. Therefore this

[2] To make the equivalence exact we need to assume that this is a pure tax-rate shock with no change in government spending, so that the government gives back all the revenue it collects in the form of a lump-sum transfer.

model can be used to think about employment but not about *un*employment, since in the model no one is ever unemployed. Later on we will consider the possibility that the labor market is not competitive, so households are not necessarily choosing how much to work every period. This will open up other possibilities of why employment and consumption can move together.

Quantitative Assessment

One way to assess whether the RBC is a good theory of business cycles (and, more generally, to assess models of the macroeconomy) is the following procedure, advocated by Prescott (1986):

1. Construct a full version of the model. In this chapter we have looked at a simplified two-period version of the model but in a full assessment one would use a model with an infinite horizon, labor and capital in every period, investment, depreciation, etc. and perhaps other ingredients as well.

2. Set values for the parameters of the model, based as much as possible on *micro*economic data. For instance, one would use data on household labor supply to set values for the parameters that govern preferences for consumption and leisure.

3. Do a growth-accounting exercise like we did in Chapter 4 with data from the US economy (or some other country) and find the Solow residuals for each period (usually one quarter). The point of this is that if we are trying to assess a model that says business cycles are the result of changes in productivity we need to have a sense of how large these changes in productivity actually are. Implicitly, we are assuming that measured Solow residuals are accurate measures of exogenous technological shocks.

4. Simulate how the model economy would respond to the types of technological shocks extracted in the previous step.

5. Measure the behavior of the variables of interest in the model economy (GDP, employment, etc.) and compare it to the same measurements taken from the real economy.

Applying this procedure, Kydland and Prescott (1982) and others found that the model RBC economy could produce about $\frac{2}{3}$ of the volatility of GDP observed in the real data. Furthermore, the model economy could reproduce the main correlations that define business cycles: productivity, output, employment, consumption and investment all move together, investment is more volatile than GDP and consumption less so. They concluded that, while not perfect, the RBC model is a satisfactory approximation to how the real economy behaves.

Policy Implications

If the RBC model is correct, then the implications for macroeconomic policy are profound. In the RBC model, the First Welfare Theorem holds. This means that no social planner, no matter how unrealistically powerful, could improve upon what the market economy is doing. In particular, nothing should be done to prevent or stabilize business cycles. Business cycles are just the efficient way for the economy to respond to changes in productivity. When productivity rises, it's a good time to produce goods, so households *should* work more; when productivity falls, it's a bad time to produce goods, so households *should* enjoy more leisure. Any attempt to stabilize employment or output, even if it could succeed, would reduce the welfare of the representative household.

When it was first proposed, this conclusion ran very counter to decades of thinking about macroeconomic policy, which had made stabilizing the business cycle one of its priorities. At the very least, the model forces us to ask harder questions. The model proves that observing economic fluctuations does not *imply* that something is wrong and needs to be fixed. Economic fluctuations can be perfectly consistent with a world in which nothing is wrong. Therefore any argument for trying to stabilize the business cycle must first make the case of why such stabilization is desirable. We'll come to some of these arguments later on.

Furthermore, if the model is correct, conventional macroeconomic policy might not work anyway. One of the main means by which policymakers attempt to stabilize the economy is by using monetary policy (we'll come to some of the reasons for this later on). But the RBC is a completely *real* model: there is no room for monetary policy because there is no money at all in the model. Hence, if the RBC model is right, it would be a good idea to close central banks, or at least to drastically limit what they do. According to the model, anything they do is irrelevant and, if it were to have an effect, it would be counterproductive.

Criticisms

One of the earliest criticisms made of the RBC model is that many economists just don't find the mechanism for generating recessions plausible.[3] If we interpret productivity literally as the result of technological progress, some economists don't think it's plausible that the rate of technological progress can fluctuate so much over the course of a year or two to result in the types of business cycle movements we observe. Furthermore, the example in Figure 13.3.1 shows what happens after an *improvement* in productivity. We at least know that technology does improve over time, even if we disagree about how smooth this progress is. In order for the model to produce a recession, there needs to be a *fall* in productivity. What exactly is this fall supposed to represent? Can an economy, from one year to the next, lose the ability to successfully employ technologies that it used the year before? Most defenders of the RBC model argue that productivity shocks should be interpreted less literally. They argue that other things like changes in regulations or specific problems at individual large companies can make the economy behave *as if* it had experienced technological regression.

Another criticism of the model has to do with the parameter values that one needs to use in order to get the model to work quantitatively. In particular, there is much disagreement as to the right numbers to use to describe household's willingness to substitute between consumption and leisure, which in turn governs the elasticity of labor supply.[4] For the RBC model to produce large fluctuations in employment, the elasticity of labor supply needs to be high. Remember, all the changes in employment in the model are the result of households willingly changing how much they work. In order to produce the changes in employment that we observe, households must respond strongly to changes in wages. Many economists regard the elasticity usually used in the RBC model as contradictory with microeconomic evidence on household behavior. Much of the debate is about the right way to derive the response of the aggregate labor supply on the basis of individual

[3] Many of these criticisms can be found in Summers (1986).

[4] We already encountered this issue in Exercise 7.5 when we thought about the response of labor supply to different tax rates in Europe and the US.

labor supply.

The simulations of the RBC model often focus on how quantities move: output, employment, investment, etc. But, using equations (13.2.1) and (13.2.3), one can also look at what the model implies for how prices (wages and interest rates) move. One criticism of the model is that it doesn't fit the behavior of prices as well as it fits the behavior of quantities. In particular, the model implies that wages should be more variable than we actually observe. The evidence in presented in Table 12.1 indicates that real wages do not co-move strongly with the business cycle. But it's also possible that standard ways of measuring wage movements underestimate how much wages actually move.

In Chapter 7 we made a distinction between being unemployed (not working but looking for work) and being out of the labor force (not working and not looking for work). In the RBC model, there is never any unemployment, since the labor market is competitive and everyone who wants to work finds a job. The model can therefore provide a theory of changes in employment but not of changes in unemployment. Many economists consider changes in *un*employment as a central feature of business cycles and therefore consider the RBC unsatisfactory on those grounds.[5]

Another line of criticism of the RBC model has to do with the practice of treating measured Solow residuals are exogenous productivity. One of the reasons why this might be inaccurate is that mismeasured capacity utilization can contaminate measurement of the Solow residual. To see why that is, imagine a restaurant at a time when business is slow. The restaurant still has its usual level of capital (the building, the kitchen equipment, etc.) and all its employees. However, it is producing fewer meals than usual because customers are not showing up. If we go back to equation (5.4.2) that describes how one would construct a Solow residual, we'll see that for this particular restaurant:

$$g_Y = \text{Capital Share} \times 0 + \text{Labor Share } \times 0 + \text{Solow Residual}$$

Since neither the labor nor the capital it employs has changed, this accounting procedure would attribute all the change in the number of meals the restaurant produces to lower productivity. At a literal level, this is not wrong: the restaurant *is* being less productive by producing fewer meals with the same amount of labor and capital. However, this lower productivity is endogenous, it's the result of whatever it is that is causing business to be slow so that factors of production are not being fully utilized, not of the restaurant having become technologically worse at producing meals. Later on we'll study models where business can be slow for the overall economy, not just a single restaurant. If these models are right, then the practice of treating measured Solow residuals as exogenous technological changes is inaccurate.[6]

[5] We saw one theory of unemployment in Chapter 7, based on modeling the search process by which firms find workers and workers find jobs. Inserting this type of model of unemployment into an RBC model can account for why there is unemployment but has a hard time in getting unemployment to change very much with productivity shocks. See Shimer (2005) for a discussion of this point.

[6] Basu et al. (2006) attempt to correct for this by constructing a "utilization adjusted" measure of TFP.

Exercises

13.1 **Prices in the RBC Model**

Consider the two-period RBC model summarized by equations (13.1.2)-(13.1.5).

(a) Suppose there is a temporary, positive, productivity shock in period 1 only. What happens to the wage w and the real interest rate r? Explain in words why it is that each of the prices change.

(b) Suppose there is temporary increase in households preference for leisure. What happens to the wage w and the real interest rate r? Explain in words why it is that each of the prices change.

13.2 **Stabilizing a Real Business Cycle**

Suppose the economy is well described by the Real Business Cycle model. The economy has just had a *positive* shock to aggregate productivity.

(a) Suppose the government wants GDP to remain the same as it would have been without the productivity shock. Describe one possible policy instrument that could be used for this. Explain why it works.

(b) Suppose instead that the government wants employment to remain the same as it would have been without the productivity shock. Does the policy need to be applied more intensely or less intensely? Why?

13.3 **Good News**

Consider the two-period RBC model summarized by equations (13.1.2)-(13.1.5). Suppose that the representative household develops a very particular form of optimism about the future. It believes that the period-2 production function is:

$$Y_2 = F_2(K) + A$$

where A is some positive number.

(a) How does this type of optimism differ from the one we looked at in Section 13.4?

(b) What will happen to employment, output, consumption, investment, the wage and the real interest rate?

(c) Does this type of optimism produce something that looks like a business cycle? Why or why not?

13.4 **Government Spending**

Consider an economy that is well described by the RBC model with one modification: there is a government that spends G in period 1. It pays for this by collecting lump-sum taxes from the representative household (Ricardian equivalence holds, so we don't need to specify when the government collects these taxes). G does not enter the representative household's utility function.

(a) Where in the system of equations (13.1.2)-(13.1.5) would G show up? Explain.

(b) Suppose there is an increase in G. What happens to output, consumption, employment and investment? What happens to wages and interest rates? Show this graphically and explain in words what is going on.

13.5 **Fish**

In the town of New Oldport in Maine, the main economic activity is fishing. Crews sail out to New Oldport Bay in the morning and return in the evening with their catch. It's a very unstable occupation because it depends on the migration patterns of cod and bass, which are erratic. Sometimes (when a large school of fish is passing by) fish are plentiful in the New Oldport Bay; otherwise they are scarcer. Sonar tracking gives residents a reasonably reliable daily estimate of the number of fish in the bay. On average, large schools of fish come by the Bay a few days a year.

(a) Write down two production functions for the New Oldport economy: one for when fish are plentiful and one for when they are not.

(b) Write down the problem of a household that resides in New Oldport and has to decide how much to consume and how much to work each day. Derive first-order conditions (you may skip steps if you want).

(c) Suppose you compare a day when fish are plentiful and one when they are not. On what day will workers choose to work longer hours?

(d) Suppose in 2018 the migration patterns of cod and bass change such that now fish in the New Oldport Bay are always plentiful. How will the number of hours that households choose to work on a typical day in 2018 compare to the number of hours they worked on a plentiful-fish day in 2017? How will they compare to the *average* daily number of hours in 2017?

CHAPTER 14

The New Keynesian Model

14.1 A Historical and Methodological Note

In this chapter we'll study a simple version of the so-called New Keynesian model. For a long time, starting around the 1930s, the ideas of Keynes (1936) were the dominant way to think about business cycles. We'll be more precise about this, but broadly speaking the main idea in Keynes' work was that output and employment could fall short of their normal level due to a lack of demand.

Keynes himself made his arguments in prose, rejecting the mathematical formulations that are preferred nowadays. As a result, his writings can be a little bit hard to interpret. The so-called IS-LM model, first proposed by Hicks (1937), is one popular mathematical representation of Keynes' ideas. There is some debate as to whether the IS-LM model accurately represents what Keynes really meant. The best answer is probably "who cares?": the merit of the model or lack thereof must be judged on its own and not on the basis of its faithfulness to Keynes.

The RBC model that we looked at in the last chapter, developed around the late 1970s, was a departure from the IS-LM model on two dimensions. First, a methodological difference. The IS-LM model started by describing relationships between aggregate variables (such as total investment and total consumption) without being totally precise about how those relationships came about. In contrast, everything that happens in the RBC model is a result of households and firms making decisions in a way that is explicitly modeled. This type of model is sometimes known as a "microfounded" model, because it is built on an explicit microeconomic model of decision-making. Arguably, such an approach makes models clearer because it allows one to keep track of where each result comes from. It also allows one to test the model against a wider range of evidence, testing not just the implications of the model for aggregate variables but also its implications for microeconomic outcomes. Finally, since one of the building blocks of a microfounded model is a utility function, the model can be used to evaluate the welfare of the household in the model in an internally consistent way.

The second difference between the RBC model and the earlier IS-LM model is substantive. In the Keynesian/IS-LM account, recessions happen because things go wrong with markets, which implies that perhaps the government should do something about it. Later on we'll think more about exactly what it is that goes wrong and what the government can do to fix it. In the RBC model, instead, business cycles do not reflect any failure of the market economy and the government should not do anything about them, even if it could.

The New Keynesian model that we'll study in this chapter is, like the RBC model, a microfounded model where everything that happens results from decisions that are explicitly modeled. The "New" in New Keynesian comes from this methodologically more modern way of thinking. In contrast to the RBC model, it is a model economy where things do go wrong in specific ways. The "Keynesian" in New Keynesian come from the fact that the specific types of market failures in the model are very similar in spirit to the earlier generation of Keynesian models.

The starting point for our New Keynesian model will be the RBC model summarized by equations (13.1.2)-(13.1.5). We'll then add three ingredients: monopoly power, sticky prices and the theory of money markets we developed in Chapters 10 and 11. Adding these ingredients will lead to a model of the economy made up of two equations that we'll label "IS" and "LM", i.e. a version of the IS-LM model.

14.2 Monopoly Power

The Microeconomics of Monopoly Power

Let's review from microeconomics how a monopolistic firm sets its price. Suppose the firm faces a demand function $q(p)$, meaning that it will sell q units if it sets the price p. Let the total cost of producing q units be $c(q)$. Then the firm's problem is:

$$\max_{p} \underbrace{pq(p)}_{\text{Total Revenue}} - \underbrace{c(q(p))}_{\text{Total cost}}$$

The first-order condition for this problem is:

$$\begin{aligned} pq'(p) + q(p) - c'(q(p))\,q'(p) &= 0 \\ \Rightarrow \underbrace{p + \frac{q(p)}{q'(p)}}_{\text{Marginal Revenue}} &= \underbrace{c'(q(p))}_{\text{Marginal cost}} \end{aligned}$$

The monopolist will equate marginal revenue from selling an extra unit to the marginal cost of producing it. Note that since $q'(p) < 0$, marginal revenue is below the price. In order to sell an extra unit the monopolist needs to lower the price; the price cuts on all the units it was going to sell anyway subtract from what it earns on the last unit.

We can further rearrange the first order condition to get:

$$\begin{aligned} p - c'(q(p)) &= -\frac{q(p)}{q'(p)\,p}p \\ &= \frac{1}{\eta}p \\ p &= \underbrace{c'(q(p))}_{\text{Marginal cost}} \underbrace{\frac{\eta}{\eta-1}}_{\text{Markup}} \end{aligned} \tag{14.2.1}$$

where $\eta \equiv -\frac{q'(p)p}{q(p)}$ is, by definition, the elasticity of demand. Since $\frac{\eta}{\eta-1} > 1$, formula (14.2.1) says that the monopolist will set its price above marginal cost. The difference between price and marginal cost is called a "markup". Formula (14.2.1) also tells us that the markup will depend on the elasticity of demand faced by this firm. If demand is very elastic (i.e. η is very high), then the markup will be small. In the limit of $\eta = \infty$, we are back to the case of perfect competition, where the firm sets price equal to marginal cost.

Figure (14.2.1) illustrates this principle, showing how two different monopolists with different demand elasticities set their price. In the left panel, the monopolist faces a relatively inelastic demand. This means that in order to sell an additional unit it needs to lower the price a lot. As a result, the marginal revenue curve is far below the demand curve. The monopolist chooses a low quantity and a large markup. In the right panel, the monopolist faces a very elastic demand curve, so it sets a lower markup.

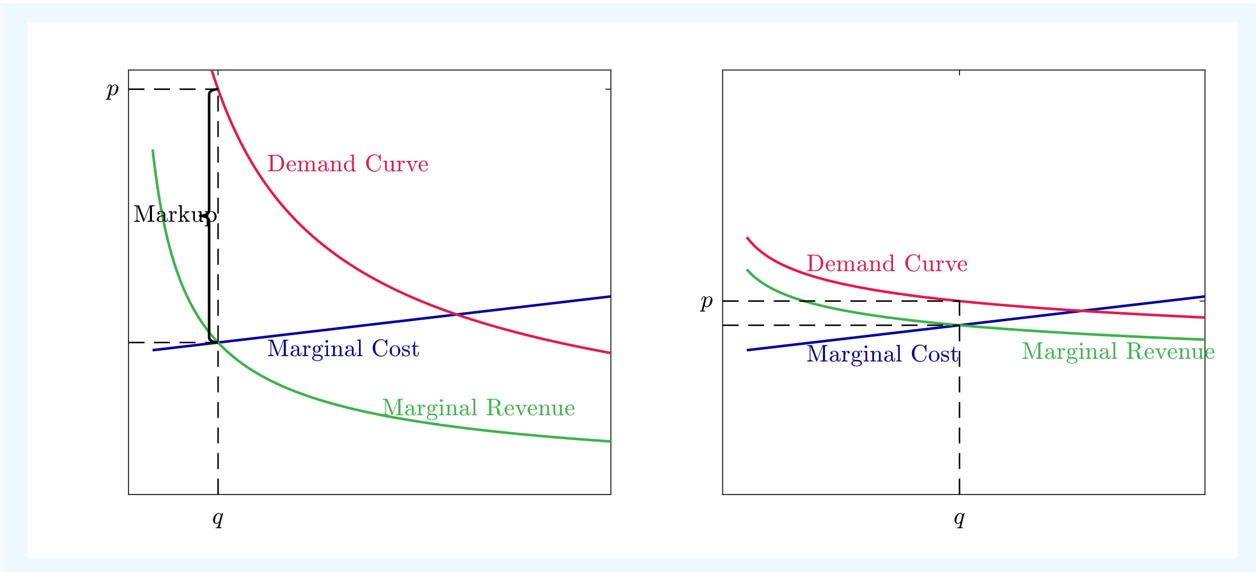


Fig. 14.2.1: *The monpolist's price-and-quantity decision.*

Introducing Monopoly Power into the RBC Model

We are going to assume that markets are *not* perfectly competitive. There are many ways for markets to be not fully competitive and we are going to model a simple one. Imagine that instead of selling their labor to the representative firm in a competitive labor market, each individual worker operates their own small firm. This small firm can produce output in period 1 with the same production function we had before:

$$Y_1 = F_1(L)$$

Each of these small firms produces a slightly different good, and this differentiation gives them some market power. We are going to assume that all these small firms are symmetric and face identical demand curves. Let η denote the elasticity elasticity of demand that each of them faces. Therefore we know that they are going to set price at marginal cost times a markup $\frac{\eta}{\eta-1}$. What exactly is the marginal cost of production for one of these monopolistic producers? The marginal cost is the answer to the question: "if I tell the producer to produce one more unit of his specific good, what do I have to give him in compensation to leave him exactly indifferent?" Let's construct an answer to this question.

$F_1'(L)$ is the marginal product of labor. This means that one marginal unit of labor will produce $F_1'(L)$ additional units of output. Therefore, producing one extra unit of output will require $\frac{1}{F_1'(L)}$ additional units of labor. Since $v'(1-L)$ is the marginal utility of leisure and supplying additional labor reduces leisure, supplying enough labor to produce one marginal unit of output costs the household $\frac{1}{F_1'(L)}v'(1-L)$ units of utility. How much extra consumption would the household need to receive to be exactly compensated for the disutility of supplying this extra labor? If $u'(c_1)$ is the marginal utility of consumption and the household needs to receive $\frac{1}{F_1'(L)}v'(1-L)$ extra units of utility, then it requires:

$$\text{Real Marginal Cost} = \frac{v'(1-L)}{u'(c_1)\,F_1'(L)} \tag{14.2.2}$$

We sometimes also want to express this cost in nominal terms, i.e. how many *dollars* does the worker require in order to compensate the marginal disutility cost of producing an extra unit of output. Let p be the price of the average good produced by all the different small firms. Then the nominal marginal cost is:

$$\text{Nominal Marginal Cost} = p\frac{v'(1-L)}{u'(c_1)\,F_1'(L)}$$

How will an individual small firm set its price? It will just set a markup of $\frac{\eta}{\eta-1}$ over the marginal cost. In nominal terms, this means that firm i will set its price p_i at:

$$p_i = p\frac{v'(1-L)}{u'(c_1)\,F_1'(L)}\frac{\eta}{\eta-1} \tag{14.2.3}$$

But we know that all of these firms are symmetric: they all have the same production function and face the same elasticity of demand. It stands to reason that they will all set the same price, so the average price p will be the same as the price of any one of them: $p_i = p$ for all i. Using this in (14.2.3) implies:

$$1 = \frac{v'(1-L)}{u'(c_1)\,F_1'(L)}\frac{\eta}{\eta-1}$$

$$\Rightarrow \frac{v'(1-L)}{u'(c_1)} = F_1'(L)\,\frac{\eta-1}{\eta} \tag{14.2.4}$$

Contrast equation (14.2.4) with condition (13.1.3) from the RBC model. The equations are the same except for the term $\frac{\eta-1}{\eta} < 1$. In fact, equation (14.2.4) is exactly equivalent to (13.4.2), which describes how the economy reacts to an increase in taxes. Monopoly power ends up having the exact same effect as a tax on labor income, with a tax rate of $\tau = \frac{1}{\eta}$.

What's going on? Each worker is acting as a monopolist: reducing supply in order to maintain a high price. But since they are all doing it at the same time, none of them succeeds in raising the relative price of their own product and the aggregate effect is just that they all reduce supply relative to what would happen in a competitive market. Notice that as $\eta \to \infty$, the term $\frac{\eta-1}{\eta}$ converges to 1 and we are back to competitive markets.

In terms of how the model economy reacts to various shocks, having monopoly power in the model does not make too much difference. We just start from a situation where the labor-consumption curve is shifted to the left, as in Figure (13.4.3). Any further shock will have the same consequences that they do in the plain RBC model.

One major difference that monopoly power *does* make is that the First Welfare Theorem no longer holds. In an economy where there is monopoly power, an all-powerful social planner *would* like to implement a different outcome compared to what the market equilibrium is bringing about. In particular, a social planner would like to undo the effects of monopoly power. The planner would like to get workers to work a bit more, knowing that the value of the goods they'd produce exceeds the disutility that they would incur.

In discussions of economic policy it is often taken as obvious that increasing employment ("creating jobs") is desirable. The First Welfare Theorem is a useful reminder that the desirability of higher employment is *not* obvious: if markets are competitive, increasing employment will lower welfare because the value of additional consumption does not make up for the loss of leisure. Conversely, this model with monopoly power gives a precise sense in which the commonly-held view is correct: if there is monopoly power, the value of additional consumption *does* make up for the loss of leisure and higher employment *is* desirable.

14.3 Sticky Prices

How Monopolists Want to Adjust Prices

Suppose, that, for whatever reason, the demand curve faced by one of our monopolistic producers shifts up. How should we expect the firm to adjust its price? We know that the firms wants to maintain a markup of $\frac{\eta}{\eta-1}$ over marginal cost, and let's assume that the *elasticity* of demand η hasn't changed. Then the firm will want to change its price in proportion to the change in its marginal cost. Figure 14.3.1 shows an example of this, where demand rises but the elasticity doesn't change. The monopolist increases the quantity it supplies; since marginal costs rise, then in order to maintain a constant proportional markup, the monopolist raises its price.

Note that the key to why the monopolist raises its price is that its marginal costs are increasing. If marginal costs were constant, the monopolist would react to the increase in demand by increasing output but keeping the price constant. How do we know that marginal costs are indeed increasing? In our model of owner-operated monopolistic firms, we can see this follows from equation (14.2.2). There are three effects, all pointing in the same direction:

1. Diminishing marginal product of labor (decreasing $F_1'(L)$). As the worker works more, it requires more additional labor to produce each successive unit of output.

2. Diminishing marginal utility of leisure ($v'(l)$ decreasing in l and therefore $v'(1-L)$ *increasing* in L). As the worker works more, each additional unit of leisure that it gives up is more valuable than the previous one because there are fewer of them left. A tired worker finds additional work especially unpleasant.

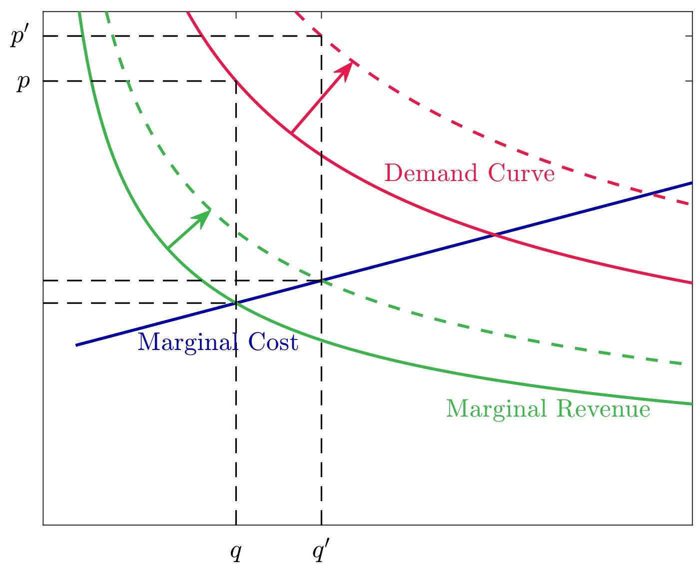


Fig. 14.3.1: *The monpolist's reaction to higher demand.*

3. Diminishing marginal utility of consumption (decreasing $u'(c_1)$). As the worker works more and earns more income, he increases consumption; this makes him place lower value on the additional units of consumption he'll be able to afford by working more.

To see more the forces at work more concretely, imagine that our worker-firm is a freelancer that sets his own prices, like a wedding photographer, a plumber or a maths tutor. He reacts to higher demand by saying: "I'm getting a bit tired of working so much, so I'll raise my prices. I know this will limit how much my business expands, which will cut into my total income, but I'm doing fine and the extra time off will make it worthwhile". Conversely, he would react to lower demand by saying: "I'm not getting enough work and I'll have trouble paying my bills, so I'll lower my prices to get more clients and increase total revenue. It'll mean more work than if I kept my prices as they are, but since business is slow I have a lot of free time so I don't mind".

Introducing Price Stickiness

The key assumption in the New Keynesian model is that prices are inflexible. Ultimately, it is an assumption about the timing of decisions. We are going to assume that firms set prices before knowing exactly what's going to happen in the macroeconomy. Once a firm has set its price, we'll assume that it's inflexible: the firm cannot change it when it faces a change in demand for its product. Sometimes this is referred to as prices being "sticky".

There are many reasons why prices might be sticky. There could be contracts such as rental contracts, service agreements or union pay scales that prevent price changes for a specified period. There could be costs to physically implementing price changes, for instance printing new menus (these types of costs are sometimes known as “menu costs”). It could be that it takes time for firms to realize that circumstances have changed in a way that would make them want to change prices, perhaps because they are rationally choosing not to pay too much attention to macroeconomic news.

Price stickiness might not seem like a big deal but it has major consequences for how the economy as a whole behaves. Let’s start by seeing how a monopolist whose price is sticky reacts to a change in demand for its product. Figure 14.3.2 shows the same increase in demand as Figure 14.3.1 for a firm whose price is sticky. If the firm could change prices when demand increases, it would choose price p^F (the F stands for “flexible”) in order to maintain a constant markup, and quantity would increase up to q^F. This is exactly what Figure 14.3.1 shows. If instead the price is stuck at p^S (the S stands for “sticky”), then the producer would produce enough to satisfy all the demand he faces at price p^S, so quantity would increase all the way to q^S. The fact the prices don’t react means that the quantity produced reacts *more* than it would under flexible prices.

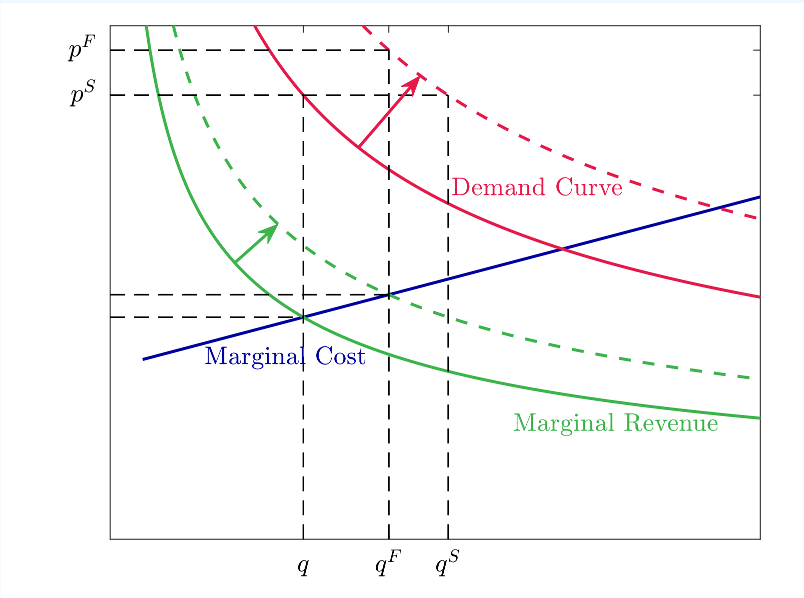


Fig. 14.3.2: *A monopolist with sticky prices.*

Going back to the example of the freelancer, imagine he has already advertised his prices for the year and, since a lot of people have seen them, he cannot easily change them. He then reacts to demand by simply adjusting how hard he works: if more clients want to hire him, then he works more; if fewer clients want to hire him, he works less and enjoys more leisure. Notice that even though higher demand makes the freelancer tired, he does not want to turn away clients, because for each additional unit, it’s still true that price is higher than marginal cost, so he’s happy to supply it.[1]

[1] If demand were to rise much more then there would reach a point where marginal costs are so high (when the worker is *very* tired) that the worker-firm with sticky prices would want to turn away clients, but we’ll assume we are not at that point. See Exercise 14.1.

14.4 Putting Everything Together

Now we are going to put all the ingredients together to work out how our model economy will behave. We did the same thing when we studied the RBC model, but there we had the advantage that the First Welfare Theorem applied. Therefore we could simply imagine that a social planner was choosing the allocation and study what the planner's decisions looked like. Here the First Welfare Theorem doesn't hold so we need to be careful about who makes each decision and how they all fit together.

Household, Firm and Government Decisions

1. First of all, each individual small firm sets a price, understanding that the price will be sticky. They would like to follow the pricing rule (14.2.3) but at the time they set the price they don't exactly know what c_1 and L are going to be, so they can't always get it exactly right. Since all firms are symmetric, we'll assume they all set the same price and call it p_1. Later on we are going to think more carefully about how firms choose p_1. For now we are going to take it as given because by the time interesting things start to happen in the economy, p_1 has already been chosen.

2. The government chooses a level for the money supply M^S, and possibly other policies as well.

3. The household makes a consumption-saving decision. These decisions will satisfy the usual Euler equation (6.2.8):
$$u'(c_1) = \beta(1+r)u'(c_2)$$

4. Each firm will make an investment decision for period 2. These will satisfy the investment-demand condition (8.3.2). Since period 2 is the last period we set $\delta = 1$ so this reduces to:
$$F_2'(K) = 1 + r \tag{14.4.1}$$
Let:
$$K(r) \equiv (F_2')^{-1}(1+r) \tag{14.4.2}$$
The function $K(r)$ comes from solving (14.4.1) for K. It simply tells us the level of investment as a function of the real interest rate. As we know from Chapter 8, this will be a decreasing function: higher interest rates lower the NPV of investment projects so, other things being equal, fewer of them will be undertaken.

   Notice that we are describing savings and investment as separate decisions. We know from our basic GDP accounting that in a closed economy savings and investment are equal by definition. We'll see below what makes this equality hold despite the decisions being taken separately.

5. The household decides how much money to hold. Its real money demand is:
$$m^D(Y_1, i)$$

Notice that investment decisions are governed by the *real* interest rate because investment transforms real goods today into real goods tomorrow. Instead, money demand is governed by the *nominal* interest rate, because the opportunity cost of holding money depends on the nominal interest rate.

6. In period 2 the government chooses a new level for the money supply. We are not going to model the period-2 money demand explicitly. We'll just assume that the price level in period 2, denoted p_2, depends on the money supply in period 2 in some way. We are going to take p_2 as exogenous for now, but we'll see that expectations about p_2 will play an important role. Since we have also taken p_1 as exogenous, we are effectively treating inflation $\pi \equiv \frac{p_2}{p_1} - 1$ as exogenous. The reason we care about p_2 and inflation is that in the model we have both nominal and real interest rates playing a role, so we need to connect them. Recall that by definition $r = i - \pi$.

Market Clearing Conditions

There are several markets in the model: markets for goods in periods 1 and 2, a market for labor and a market for money. Let's think about what it means for each of them to clear.

1. Goods in period 1. Denote GDP in period 1 by Y_1. Goods produced in period 1 can be used either for consumption or for investment, so market-clearing requires:

$$Y_1 = c_1 + K$$

   This is the basic GDP identity (1.1.1) for a closed economy with no government. We can also rearrange it as:

$$\underbrace{Y_1 - c_1}_{\text{Saving}} = \underbrace{K}_{\text{Investment}}$$

   By imposing market clearing for goods we are making sure that saving equals investment even though they are separate decisions.

2. Goods in period 2. Since period 2 is the last period, the household consumes all the output produced in period 2, so:

$$c_2 = F_2(K)$$

   This was also true in the RBC model.

3. Labor. Here we need to remember what we assumed about how the owner-operated firms behave. We assumed that the workers who run the firms work exactly as much as they need to meed demand for their product. Total demand includes demand for consumption and demand for investment, so it's equal to Y_1. The production function $Y_1 = F_1(L)$ implies that in order to produce Y_1, each worker has to supply:

$$L = F_1^{-1}(Y_1) \tag{14.4.3}$$

   units of labor. Note that this is very different from how the quantity of labor is determined in the RBC model. Here the worker is not optimizing how much to work. Instead, he works exactly as much as the demand for his product requires. He does not work more because no one would want to buy the extra units he would end up producing.

4. Money. The money-market clearing condition is just given by (14.4.9).

We are going to condense the four market-clearing conditions into just two, known as the IS and LM equations. First, we are going to ignore the labor market clearing condition. Why? Because we know that L will just respond to Y_1 according to (14.4.3), so if we determine the level of Y_1 we can always figure out L very easily. Second, we are going to use the goods market clearing conditions for both periods, together with the Euler equation, to derive an equation known as the IS equation. IS stands for "Investment=Savings", i.e. market clearing for period 1 goods. Finally, we are going to re-name the money market clearing condition as the LM equation. LM stands for "Liquidity=Money", where "liquidity" is another way of saying "money demand", so "Liquidity=Money" just means market clearing for money.

Both the IS equation and the LM equation are going to be relationships between between GDP Y_1 and the nominal interest i. These are two *endogenous* variables, and in the end we are going to find the values of Y_1 and i that are consistent with *both* IS and LM. This is similar to how we think of supply and demand in microeconomics: supply and demand each define a relationship between price and quantity, both of which are endogenous variables, and the equilibrium is the price and quantity that is consistent with both supply and demand.

The IS Relationship

Start from the Euler equation:

$$u'(c_1) = \beta(1+r)\,u'(c_2) \tag{14.4.4}$$

Replace c_1 and c_2 using the market clearing conditions for goods in each period:

$$u'(Y_1 - K) = \beta(1+r)\,u'(F_2(K))$$

Replace K using the investment-demand condition

$$u'(Y_1 - K(r)) = \beta(1+r)\,u'(F_2(K(r)))$$

Finally, replace r by $i-\pi$:

$$u'(Y_1 - K(i-\pi)) = \beta(1+i-\pi)\,u'(F_2(K(i-\pi))) \tag{14.4.5}$$

Equation (14.4.5) is the IS equation, the first of our two equations relating Y_1 and i. Let's first see mathematically which way the relationship goes and then try to interpret what it means in economic terms. Let's rewrite (14.4.5) as:

$$\varDelta = u'(Y_1 - K(i-\pi)) - \beta(1+i-\pi)\,u'(F_2(K(i-\pi))) = 0$$

and then take the derivative with respect to Y_1 and i:

$$\frac{\partial \Delta}{\partial Y_1} = \underbrace{u''\left(Y_1 - K\left(i-\pi\right)\right)}_{(-)} < 0 \tag{14.4.6}$$

$$\frac{\partial \Delta}{\partial i} = -\underbrace{u''\left(Y_1 - K\left(i-\pi\right)\right)}_{(-)}\underbrace{K'\left(i-\pi\right)}_{(-)} - \beta\underbrace{u'\left(F_2\left(K\left(i-\pi\right)\right)\right)}_{(+)} \tag{14.4.7}$$

$$-\beta\left(1+i-\pi\right)\underbrace{u''\left(F_2\left(K\left(i-\pi\right)\right)\right)}_{(-)}\underbrace{F_2'\left(K\left(i-\pi\right)\right)}_{(+)}\underbrace{K'\left(i-\pi\right)}_{(-)} < 0 \tag{14.4.8}$$

The signs of the derivatives come from the fact that:

- the marginal utility of consumption is positive: $u'(\cdot) > 0$,
- the marginal utility of consumption is decreasing: $u''(\cdot) < 0$,
- the marginal product of capital is positive: $F_2'(\cdot) > 0$, and
- investment is decreasing in the interest rate: $K'(\cdot) < 0$.

Equations (14.4.6) and (14.4.8) imply that IS imposes a negative relation between Y_1 and i. Increases in either of these variables make Δ go down, so in order to maintain $\Delta = 0$, if Y_1 goes up then i must go down, and vice-versa. Formally, the Implicit Function Theorem says that:

$$\frac{di}{dY} = -\frac{\frac{\partial \Delta}{\partial Y_1}}{\frac{\partial \Delta}{\partial i}} < 0$$

What is this telling us in economic terms? At the heart of the IS equation is the assumption the producers respond to demand: if demand for their product goes up, they just produce more. The IS relationship follows from answering the following question: what will aggregate demand (including demand-for-consumption and demand-for-investment) be for each possible level of nominal interest rates?

Let's start with investment demand. Equation (14.4.2) says that investment demand will be a decreasing function of r: at higher interest rates, fewer investment projects look attractive, so there is less investment. Since $r = i - \pi$ and we are taking π as exogenous, this means that K depends negatively on i.

Turn now to consumption demand. Let's try to figure out c_1 by studying the Euler equation (14.4.4). This tells us that c_1 depends:

- negatively on r, taking c_2 as given. This is the intertemporal substitution motive. If r is high, present goods are expensive relative to future goods so the household will consume fewer of them. Since $r = i - \pi$, this means that c_1 depends negatively on i.
- positively on c_2, taking r as given. This is the consumption-smoothing motive. If the household expects more consumption in the future, it will smooth this out by consuming more in the present. But we know that c_2 will be equal to Y_2, which is a function of K, which depends negatively on i. Therefore indirectly this also makes c_1 depend negatively on i.

This means that there are two channels by which higher interest rates induce lower consumption. First, higher interest rates persuade households to tilt their consumption pattern away from present consumption by making present goods expensive. Second, by lowering investment they lower expectations of future consumption, which persuades households to lower consumption in every period.

Therefore, in our model, both investment demand and consumption demand (and therefore aggregate demand) are decreasing in the interest rate. If we plot the IS equation, it will look like a downward-sloping curve.

The LM Relationship

We are going to import the theory of money markets we developed in Chapters 10 and 11 into our model of how the economy works. If you recall from Chapter 11, the so-called classical view states that money is neutral, i.e. it has no effect on real quantities. We also saw that under this view, changes in the money supply translate immediately into proportional changes in the price level. Therefore the classical view is incompatible with sticky prices. If prices are sticky, money will *not* be neutral and we need to think about how the money market interacts with the real economy.

The LM relationship is just a re-naming of the money market clearing condition:

$$M^S = m^D\left(Y_1, i\right) \cdot p_1 \tag{14.4.9}$$

Remember, we are treating the price level p_1 as exogenous because prices are sticky, and M^S is exogenous as well because it is chosen by the government. Therefore (14.4.9) also gives us a relationship between Y_1 and i, which are linked because both are arguments of the money-demand function $m^D\left(Y_1, i\right)$. Recall from Chapter 10 that the money demand is increasing in Y_1 (more transactions require more money) and decreasing in i (higher interest rates make households hold lower money balances). Since Y_1 and i move money demand in *opposite* directions, maintaining money-market clearing requires that they move in the *same* direction. Formally, the implicit function theorem says that:

$$\frac{dY}{di} = -\frac{\frac{\partial m^D}{\partial i}}{\frac{\partial m^D}{\partial Y_1}} > 0$$

Therefore the money-market / LM equation imposes a positive relationship between Y_1 and i.

What is this telling us in economic terms? Suppose GDP is higher. People will want to carry out a lot of transactions, so they will want to hold a lot of money to do this. But, by assumption, the money supply is fixed, so it's impossible for all of them to hold more money at the same time. So the opportunity cost of holding money (the nominal interest rate) rises until people are content with holding exactly M^S units of money.

How exactly does this adjustment take place? In the background, there is a market where people exchange interest-bearing assets like bonds for non-interest-bearing money. If people want to carry out a lot of transactions, they will be trying to sell bonds in exchange for money, so what happens is that the price of bonds falls. If you recall the basic Present Value formula from Chapter 8, then for an asset like a bond that promises a fixed future payment, a fall in the price is the same thing as a rise in the interest rate.

IS-LM

If we put the IS equation and LM equations together, we have:

$$u'\left(Y_1 - K\left(i-\pi\right)\right) = \beta\left(1+i-\pi\right)u'\left(F_2\left(K\left(i-\pi\right)\right)\right) \tag{14.4.10}$$

$$M^S = m^D\left(Y_1, i\right)\cdot p_1 \tag{14.4.11}$$

By solving this pair of equations (with M^S, p_1 and π taken as exogenous) we can jointly figure out the level of output and interest rates, as shown in Figure 14.4.1. The figure shows the downward-sloping IS relationship and the upward-sloping LM relationship. The point Y_1, i is the only combination of Y_1 and i that satisfies both equations. Since the model predicts that both the IS and the LM equations hold, it predicts what the levels of GDP and interest rates are going to be.

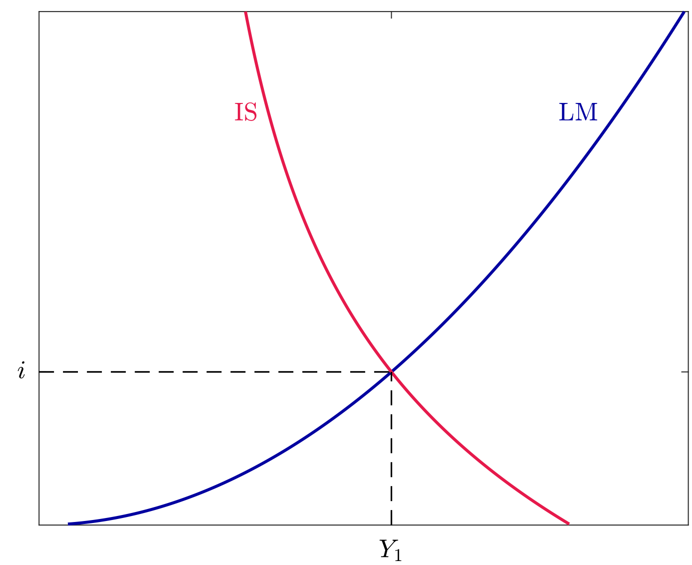


Fig. 14.4.1: *Equilibrium in the IS-LM representation of the New Keynesian Model.*

14.5 Shocks

Let's see how the economy would respond to various shocks. We'll start by analyzing the same shocks that we looked at in Chapter 13 and then we'll look at other possibilities.

Productivity Shocks

Suppose we look at the same kind of productivity shock that we looked at in Chapter 13: the production function goes from $Y_1 = F_1(L)$ to $Y_1 = AF_1(L)$ with $A > 1$. Other things being equal, how would this affect the economy?

If you look at equations (14.4.10) and (14.4.11), you'll notice that the production function does not appear anywhere. What's going on? In this model, output is demand-determined: We have assumed that firms expand and contract output to meet demand, and workers work however much it takes to satisfy this demand. An increase in productivity means that the economy *can* produce more output but not that it *will* produce more output. Output will only increase if there is more demand, and this shock does not affect demand directly.

Sometimes economists make the distinction between "supply" effects and "demand" effects. These terms are often used imprecisely but in this example they are reasonably clear. Higher productivity is a supply effect: more can be produced with the same inputs. However, with no change in demand there will be no change in quantity.

Instead, what will happen is that employment will fall. We know this from equation (14.4.3). Since workers are able to produce more output per hour but total demand hasn't changed, they will get their work done in fewer hours. Gali (1999) analyzed evidence that suggested that this is indeed what happens, spurring a large debate on whether the evidence was interpreted correctly.

Impatience Shocks

Suppose that β falls: households really want to consume now rather than later. β enters the IS equation (14.4.10). It shifts the right-hand side down, shifting the entire curve up and to the right, as shown in Figure 14.5.1.

When people become impatient, they want to consume more now. All the purchases they make induce producers to increase production, leading to an increase in output and employment. At the same time, the increase in output requires more transactions, which increases the demand for money. Since we are holding the money supply constant, this means that the interest rate rises to clear the money market: a movement *along* the LM curve. In turn this higher interest rate lowers investment.

Notice how different this is from what happens in the RBC model. In the RBC model, an increase in impatience makes consumption go up but employment go down, because there is no way to escape the logic of equation (13.1.3):

$$\frac{v'(1-L)}{u'(c_1)} = F_1'(L)$$

which implies that if households want more consumption then, other things being equal, they want more leisure as well. In the New Keynesian model this condition doesn't hold because workers are not choosing how much they work in response to shocks in a utility-maximizing way: they just accommodate the level of demand they face.

Still, as a theory of why we observe business cycles we haven't quite nailed it. Impatience shocks make output, employment and consumption move together, but investment moves the opposite way, so they cannot

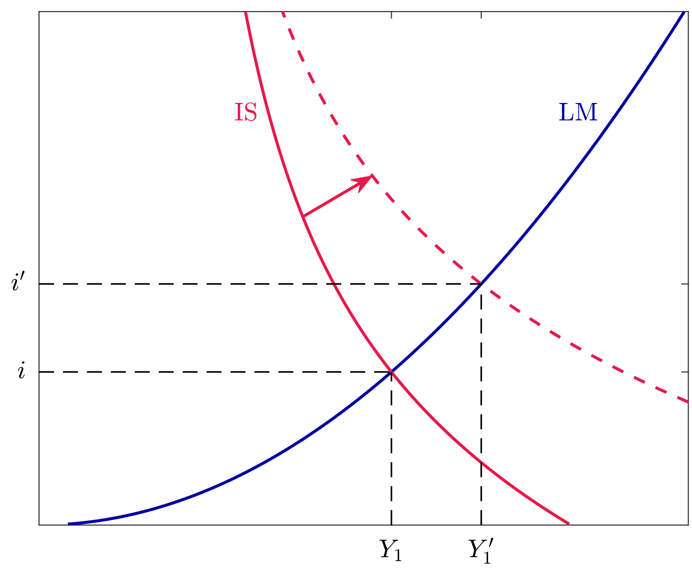


Fig. 14.5.1: *The effects of an increase in impatience (fall in β).*

be the whole story.

Optimism

Suppose households predict an increase in future productivity: the production function for period 2 goes from $Y_2 = F_2(K)$ to $Y_2 = AF_2(K)$ with $A > 1$. Now the investment function (14.4.2) becomes:

$$K(r, A) \equiv (F_2')^{-1}\left(\frac{1+r}{A}\right)$$

which means investment is increasing in A. The IS equation is now:

$$u'(Y_1 - K(i-\pi, A)) = \beta(1+i-\pi)\, u'(AF_2(K(i-\pi, A)))$$

It's easy to trace out that higher A lowers the right hand side and raises the left hand side, which has the effect of shifting the IS curve to the right, as shown in Figure 14.5.2.

What's going on? There are two effects, both going in the same direction. The first effect is through investment. If people believe that productivity will improve, then they expect the marginal product of capital to be higher. This means that a lot of investment projects are worth doing. All the resources needed to carry out these investment projects have the direct effect of increasing demand for output.

The second effect is through consumption. If productivity in the future will be higher, then future output and therefore future consumption will be higher. In addition, the fact that investment increases reinforces the effect. Households want to smooth out this anticipated future consumption by consuming more in the present. This adds to the increase in current demand.

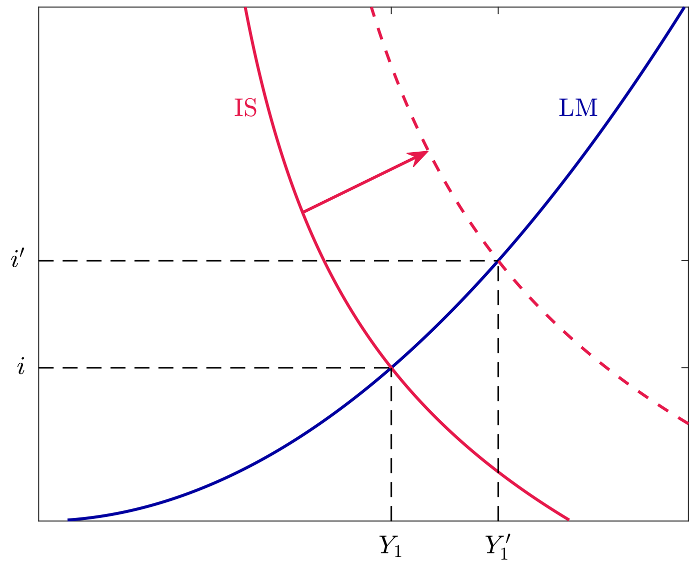


Fig. 14.5.2: *The effects of optimism (an increase in future productivity).*

As in the previous example, this requires an increase in the interest rate, i.e. a movement along the LM curve, in order to clear the money market.

This offers us a possible account of business cycles that fits the basic facts: output, consumption, employment and investment all move in the same direction at the same time. Unlike the example with impatience, a wave of optimism makes investment move together with consumption because optimism directly affects the perceived profitability of investment projects. Unlike optimism shocks in the RBC model, a wave of optimism in the New Keynesian model gets employment to move together with consumption by getting rid of the labor-consumption choice (13.1.3).

Interestingly, this type of effect may be close to what Keynes originally had in mind:

> "a large proportion of our positive activities depend on spontaneous optimism rather than mathematical expectations, whether moral or hedonistic or economic. Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as the result of animal spirits—a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities."

Laziness or Taxes

Suppose, like we did in Chapter 13, that there is an increase in labor-income taxes (or equivalently, households' preference for leisure increases), so that equation (14.2.4) becomes:

$$\frac{v'(1-L)}{u'(c_1)} = F_1'(L)\,\frac{\eta-1}{\eta}(1-\tau)$$

Other things being equal, how would this affect the economy?

If you look at equations (14.4.10) and (14.4.11), you'll notice that neither preferences for leisure, taxes nor the production function appear anywhere. As with productivity shocks, this would have no effect on output or interest rates because there is no effect on demand.[2] Taxes certainly affect worker's incentives to choose between work and leisure but we have assumed that, due to sticky prices, they are not responding to those incentives anyway, so they have no effect.

Changes in the Money Supply

Suppose the government decides to increase the money supply M^S. This increases the left-hand side of the LM equation (14.4.11), which results in a shift to the right in the LM curve, as shown in Figure 14.5.3.

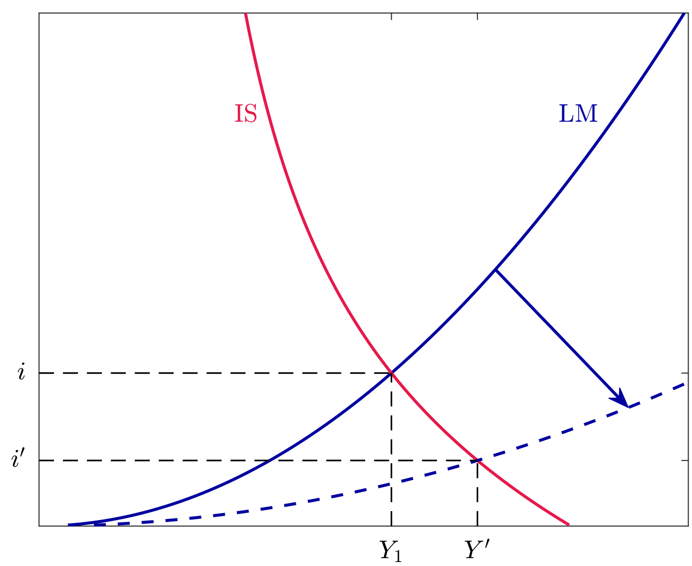


Fig. 14.5.3: *The effects of an increase in the money supply.*

This change in monetary policy leads to higher output and lower interest rates. What's going on? The money supply has increased, so something has to adjust for the money market to clear. Given the money-demand function $m^D(Y_1, i)$ and sticky prices, two things could happen to persuade households to hold the extra money that has been created. One possibility is that GDP rises, so that households need the extra money to carry out extra transactions. The other possibility is that the interest rate falls, so that households face a lower opportunity cost of holding money. The shift of the LM curve down and to the right shows the new combinations of Y_1 and i that are consistent with market clearing.

The IS curve hasn't shifted but there is a movement *along* the curve. A lower interest rate means that more investment projects are worth doing, so investment rises. Furthermore, present consumption rises both

[2] How come an increase in taxes has no effect on demand? Doesn't the fact that the government takes away part of people's income reduce their consumption? In the background, what's going on is Ricardian equivalence. We have assumed that the government increases taxes but not spending, so households rationally perceive that whatever the government is taking away it will give back, either at the same time or in the future.

to smooth out the higher future consumption and because lower rates have made present consumption cheaper. As a result of the increase in demand, output and employment rise.

This gives us another possible source of business-cycle-like movements: changes in the money supply can make output, employment, consumption and investment move in the same direction. It is an example of how money is *not* neutral in the New Keynesian model: changes in the supply of money can have effects on real variables. Friedman and Schwartz (1963) argued that changes in the money supply were one of the main sources of US business cycles. This led them to the conclusion that monetary policy should aim to keep the money supply as steady as possible to avoid causing business cycles, a point of view that came to be known as *monetarism*.

Changes in Expected Inflation

Suppose that people start to believe that prices are going to rise between $t = 1$ and $t = 2$, perhaps because of news about what future monetary policy is going to be like. In our model, this is captured by an increase in π, all else being equal. π enters the IS equation (14.4.10); higher π raises the right hand side and lowers the left hand side, which has the effect of shifting the IS curve to the right, as shown in Figure 14.5.4. Output and the interest rate both rise.

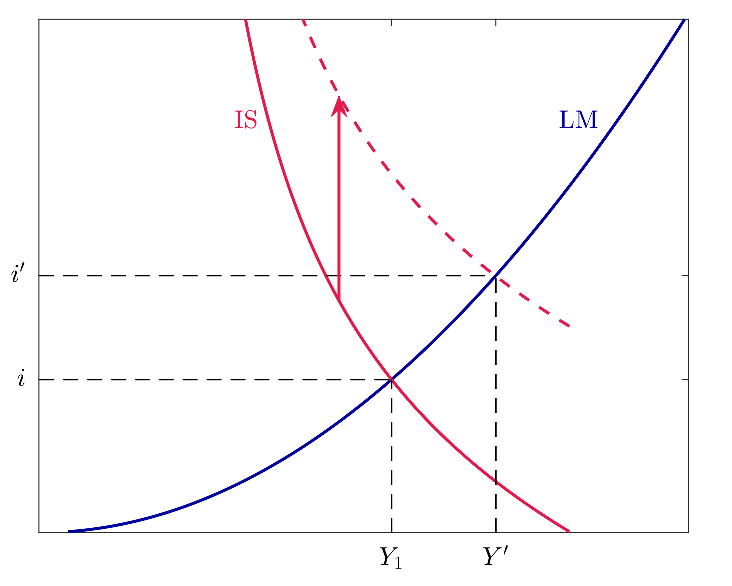


Fig. 14.5.4: *The effects of an increase in expected future inflation.*

More precisely, higher π shifts the IS curve *up*. To understand why, recall that what matters for the IS relationship is the *real* interest rate $r = i - \pi$, because this is what governs both investment and consumption decisions. What higher inflation does is change the relationship between nominal and real rates. For any given *nominal* rate i, higher inflation implies a lower *real* rate $r = i - \pi$, so it's like shifting the axis that governs the IS relationship. In equilibrium, even though the nominal rate goes up, the real rate goes down, so investment and consumption rise.

14.6 Simplified Versions of IS-LM

For some purposes it is useful to look at special cases of the IS and LM relationships in order to have fewer things to think about. In this section we'll consider some examples.

Exogenous Investment

Imagine that the period-2 production function is:

$$F_2(K) = \min\left\{AK, A\bar{K}\right\} \tag{14.6.1}$$

Under this function, the marginal product of capital is equal to A for levels of investment up to $\bar{K}$ and then drops to zero. If A is sufficiently high, this means that investment will always be equal to $\bar{K}$, no matter what the interest rate is, so GDP and consumption in period 2 will be $\bar{Y}_2 = A\bar{K}$. By doing this, we are effectively treating investment as exogenous, so the only thing left to be determined is consumption. Replacing $K = \bar{K}$ and $c_2 = Y_2$ into the IS equation (14.4.5), we get a closed form expression for i as a function of Y.

$$i = \frac{u'\left(Y_1 - \bar{K}\right)}{\beta u'\left(\bar{Y}_2\right)} - 1 + \pi \tag{14.6.2}$$

This simplified IS relationship just captures the intertemporal substitution effect of interest rates on consumption. c_1 is decreasing in the interest rate, and therefore so is $Y_1 = c_1 + K$. This gives us the negative IS relationship between interest rates and GDP. The basic economic message hasn't changed: total demand decreases with the interest rate. We have just reduced the model to its minimal essential economic ingredients.

The Old Keynesian Model

The model we have been studying is a version of the New Keynesian model. The "New" comes from the fact that it's built up from explicit microeconomic foundations. The traditional ("Old") Keynesian model has many ingredients in common. The main underlying assumptions are mostly the same (even though Old Keynesian models sometimes didn't state them precisely), and it's possible to summarize an Old Keynesian model with IS-LM equations, just like we did with the New Keynesian model. The main difference between the "New" and "Old" Keynesian models is that the Old Keynesian model is built on a less fancy theory of consumption. For some purposes, this makes a big difference.

Imagine that instead of assuming that consumption is a result of intertemporal optimization, we just proposed the Keynesian consumption function that we looked at in Chapter 6: consumption depends on current income:

$$c_1 = c\left(Y_1\right) \tag{14.6.3}$$

where $c(\cdot)$ is some function. With this consumption rule, households are not looking at the future when deciding how much to consume and they are also not looking at the interest rate.

As we saw in Chapter 6, the quantity $c'\left(Y_1\right)$ is known as the marginal propensity to consume, and is the answer to the following question: if income goes up by one dollar, by how many dollars does consumption go up? According to the intertemporal consumption theory we looked at in Chapter 6, consumption depends on the present value of the household's income. The income of one individual period is only a small part of this, so the marginal propensity to consume should be low (see Exercise 6.8). Johnson et al. (2006) and Parker et al. (2013) looked at how households responded to temporary tax rebates implemented in 2001 and 2008. They found that consumption reacted more than would be predicted by the pure intertemporal model. It is possible that some households just have a simple budgeting rule that says how much they'll consume as a function of their after-tax income, ignoring anything else. Also, some households may be borrowing-constrained and just consume as much as they can. Either of these assumptions could justify something like (14.6.3).

If consumption follows (14.6.3), the IS equation follows directly from the period-1 goods market clearing condition:

$$Y_1 = c\left(Y_1\right) + K\left(i - \pi\right) \tag{14.6.4}$$

Qualitatively, this IS equation is not that different from the New Keynesian IS equation. It also relates output Y_1 and the nominal interest i. Let's check that it is indeed downward-sloping. Restate (14.6.4) as:

$$\Delta = Y_1 - c\left(Y_1\right) - K\left(i - \pi\right) = 0$$

and then take the derivative with respect to Y_1 and i:

$$\begin{aligned}\frac{\partial \Delta}{\partial Y_1} &= 1 - c'\left(Y_1\right) > 0 \\ \frac{\partial \Delta}{\partial i} &= -\underbrace{K'\left(i - \pi\right)}_{(-)} > 0\end{aligned}$$

so, using the Implicit Function Theorem:

$$\begin{aligned}\frac{di}{dY_1} &= -\frac{\frac{\partial \Delta}{\partial Y_1}}{\frac{\partial \Delta}{\partial i}} \\ &= \frac{1 - c'\left(Y_1\right)}{K'\left(i - \pi\right)} < 0\end{aligned}$$

However, the Old and New Keynesian IS relationships do have different implications for some important questions. Section 15.1, for instance, studies how fiscal policy works differently depending on what version of the IS curve applies.

14.7 Partially Sticky Prices and the Phillips Curve

So far we have made the extreme assumption that all producers set their prices in advance, so prices are perfectly sticky. Let's now consider an intermediate case where prices are somewhat sticky but not perfectly so, and ask how the economy would behave in this case. There are two reasons to analyze this intermediate case. First, both the assumptions of perfectly flexible prices and perfectly sticky prices are extreme, so the intermediate case is probably more empirically relevant. Second, there are some interesting effects in the partially-sticky-price model that don't arise in either of the extreme cases.

To keep things relatively simple, we are going to use the exogenous-investment assumption from Section 14.6, so that investment is just $\bar{K}$. It will be useful to use the market clearing condition $Y_1 + c_1 + K$ to compute how much consumption is associated with a given level of employment:

$$c_1(L) = F_1(L) - \bar{K} \tag{14.7.1}$$

$c_1(L)$ defines a positive relationship between consumption and employment.

Suppose a fraction μ of producers have sticky prices and a fraction $1-\mu$ have flexible prices. The exact timing is as follows:

1. First, the sticky-price producers set their price. Once they set their price, the cannot change it. Let's call this price p_1^S.

2. Second, all the macroeconomic shocks are realized and any policies that the government will enact are put in place.

3. Finally, the flexible-price producers set their price, knowing everything that happened before. Let's call this price p_1^F.

The average price of a good will be simply a price index:[3]

$$p_1 = \mu p_1^S + (1-\mu)\, p_1^F \tag{14.7.2}$$

As we saw in Section 14.2, a flexible-price producer will want to set its price at a markup over marginal cost. Using (14.2.3) and (14.7.1): this means:

$$p_1^F = p_1 \underbrace{\frac{v'(1-L)}{u'(c_1(L))\, F_1'(L)}}_{\text{Real Marginal Cost}} \underbrace{\frac{\eta}{\eta-1}}_{\text{Markup}} \tag{14.7.3}$$

Replacing the pricing rule (14.7.3) into the price index (14.7.2) and solving for the price index p_1 we get:

$$p_1 = \frac{\mu p_1^S}{1-(1-\mu)\frac{v'(1-L)}{u'(c_1(L))F_1'(L)}\frac{\eta}{\eta-1}} \tag{14.7.4}$$

Equation (14.7.4) implies a positive relationship between p_1 and L. What's going on? Producers with flexible prices adjust their price in response to everything that happens in the economy. If economic shocks take place that lead to higher employment, then their marginal costs of production rise. In response to this, they set higher prices.

Now imagine that in some previous period 0 the price level was p_0. Inflation between periods 0 and 1 is

[3]This is a fixed-weight price index. Technically, one should allow for the fact that as prices change consumer substitute between flexible-price and sticky-price producers. We are going to assume this away. For small shocks, this doesn't make much difference.

given by $\frac{p_1-p_0}{p_0}$, so higher p_1 means higher inflation.[4] Therefore equation (14.7.4) gives us a positive relation between inflation and employment. In Chapter 12 we called this relationship a Phillips Curve. So now we have a theoretical justification of what gives rise to a Phillips Curve: in a partially-sticky-price model, shocks that make employment go up will also make prices go up.

We can also write down the Phillips Curve as a relationship between prices and *output* instead of employment. Invert the production function to write $L(Y_1) = F_1^{-1}(Y_1)$ and replace $L(Y_1)$ into (14.7.4) to obtain:

$$p_1 = \frac{\mu p_1^S}{1-(1-\mu)\frac{v'(1-L(Y_1))}{u'(Y_1-\bar{K})F_1'(L(Y_1))}\frac{\eta}{\eta-1}} \tag{14.7.5}$$

Equation (14.7.5) says that there is a positive relationship between prices and output, an equivalent way of expressing the Phillips Curve relationship.

Adding a Phillips Curve to an IS-LM Model

With fully sticky prices, the IS-LM model boiled down to a system of two equations in two unknowns: Y_1 and i. With partially sticky prices, we also have to solve for the period-1 price level p_1 (or equivalently, period-1 inflation). The Phillips Curve relation, which as we saw is derived from flexible-price-producers' optimal pricing, gives us an additional equation relating Y_1 and p_1, so we now have a system of three equations in three unknowns. The IS equation and the Phillips Curve are easy to see graphically because they only involve two of the three endogenous variables: the IS equation relates Y_1 to i and the Phillips Curve relates Y_1 to p_1.

The LM equation is a little bit trickier because it has all three endogenous variables in it. One way to see the LM equation graphically is to write down a modified version of it. Start from the LM equation (14.4.11) and replace the term p_1 with the expression that comes from the Phillips Curve (14.7.5) to obtain:

$$M^S = m^D(Y_1, i)\frac{\mu p_1^S}{1-(1-\mu)\frac{v'(1-L(Y_1))}{u'(Y_1-\bar{K})F_1'(L(Y_1))}\frac{\eta}{\eta-1}} \tag{14.7.6}$$

As with the basic LM equation (14.4.11), the right hand side of (14.7.6) is increasing in Y_1. Recall that the right hand side of the LM equation represents the demand for nominal money balances. There are two effects going in the same direction. First we have the basic effect we already know about: higher GDP means households want to carry out more transactions so, other things being equal, they demand more money. In addition, we have an effect coming from the Phillips Curve: since higher GDP is associated with higher prices, the nominal amount of money required to carry out a given level of transactions also goes up. On the other hand, the right hand side of (14.7.6) is decreasing in i for the usual reason that the demand for money is decreasing in the interest rates. Overall, this implies that the modified LM equation (14.7.6) describes a positive relationship between GDP and interest rates, just like the basic LM equation (14.4.11).

Figure 14.7.1 shows how the IS-LM model fits in with the Phillips Curve. The top graph shows the simplified IS equation (14.6.2) and the modified LM equation (14.7.6); the bottom graph shows the Phillips Curve (14.7.5). Together, the IS and LM curves determine the level of Y_1 and i, and then the Phillips Curve

[4] Once we are thinking of three periods: 0 (the past), 1 (the present) and 2 (the future), there are two inflation levels to keep in mind: π_1 (inflation between 0 and 1) and π_2 (inflation between 1 and 2). Here we are referring to π_1.

tells us what level of p_1 (and therefore inflation π_1) goes with this level of Y_1. We are holding constant $\frac{M}{p}$ (the real money supply), p_1^S (sticky prices) and π_2 (expected inflation between periods 1 and 2).

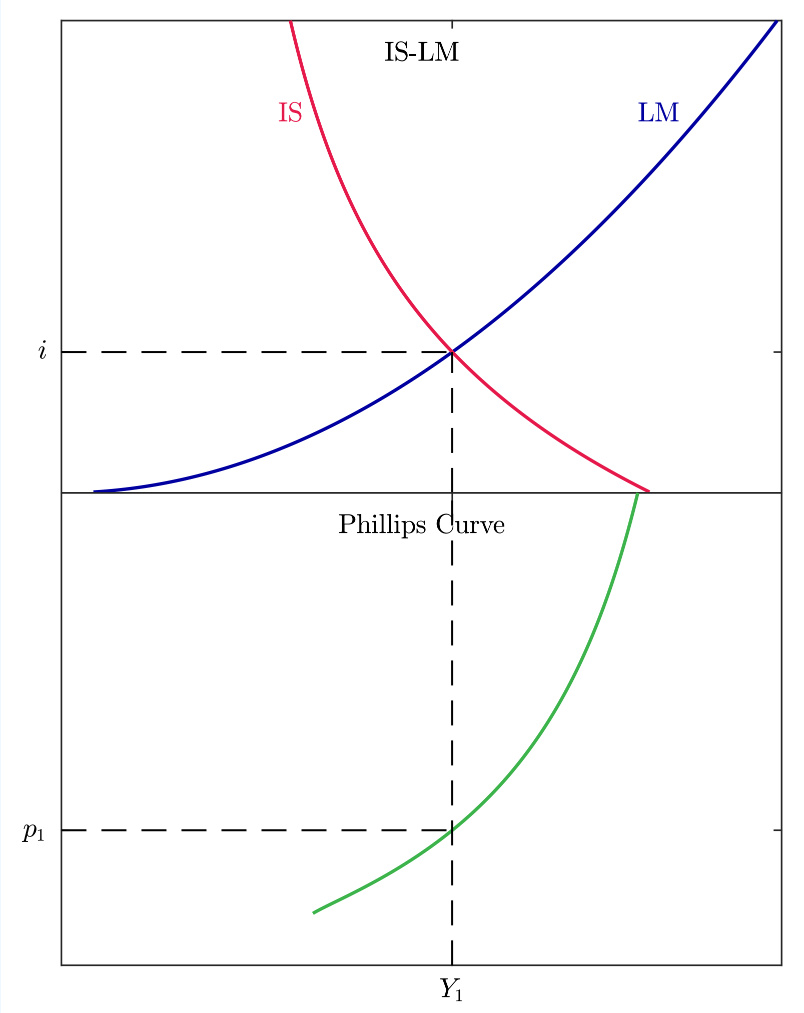


Fig. 14.7.1: *How the IS-LM curves and the Phillips Curve fit together.*

Let's re-do some of the exercises we did with fully sticky prices to see whether the conclusions change once we have partially sticky prices. Figure 14.7.2 shows the effects of a change in impatience, as captured by the discount factor β. Inspecting equations (14.6.2), (14.7.6) and (14.7.5), the only place where β shows up is in the IS equation, which shifts up and to the right. This is the same effect shown in Figure 14.5.1. Just like with purely sticky prices, the result is an increase in GDP, as impatient households demand more output. With partially sticky prices, we also have an increase in the price level, as flexible-price producers raise their prices to maintain their markups.

Figure 14.7.3 shows the effects of a productivity shock, with the production function increasing to $Y_1 = AF_1(L)$ with $A > 1$. As we saw before, productivity does not enter the IS and LM equations.[5] However, it

[5]Technically, productivity does enter the modified LM equation (14.7.6). The Phillips Curve shifts and (14.7.6) builds the

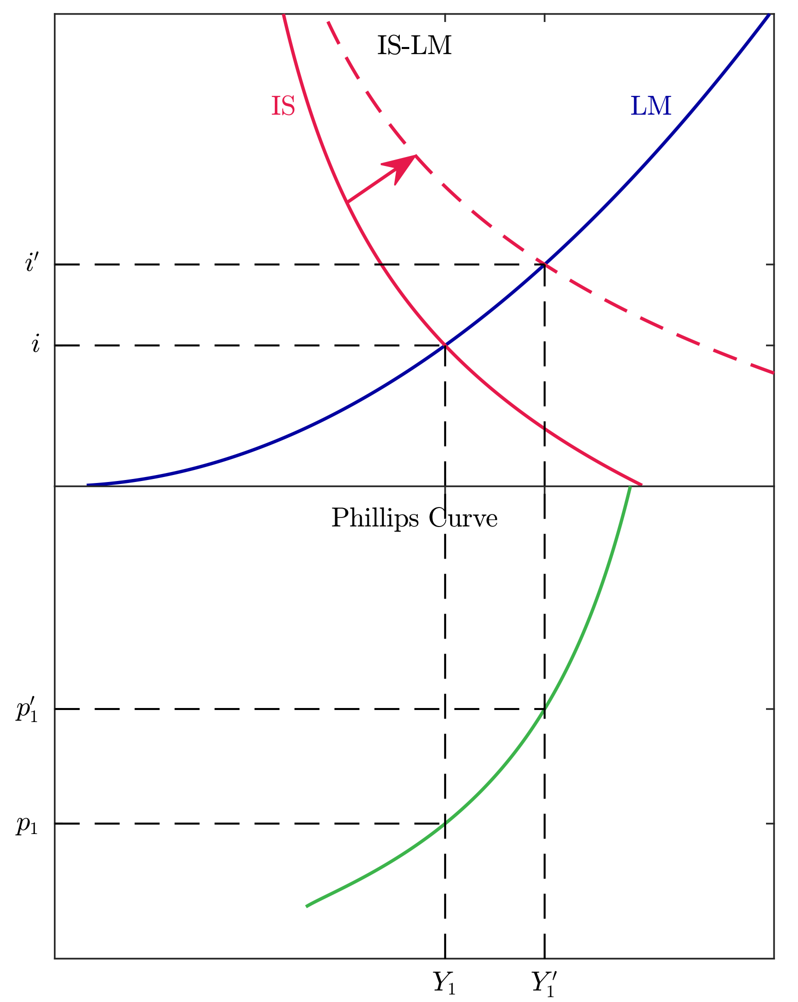


Fig. 14.7.2: *An increase in impatience.*

does enter the Phillips Curve:

$$p_1 = \frac{\mu p_1^S}{1-(1-\mu)\frac{v'\left(1-L\left(\frac{Y_1}{A}\right)\right)}{u'(Y_1-\bar{K})AF_1'\left(L\left(\frac{Y}{A}\right)\right)}\frac{\eta}{\eta-1}}$$

Higher productivity lowers marginal costs through three different channels. First, given a level of employment L, the marginal product of labor $AF_1'(L)$ rises. This means less extra labor is needed to produce an extra unit of output. Second, the amount of labor needed to produce a given level of output $L\left(\frac{Y}{A}\right)$ falls. Since there is a diminishing marginal product of labor, lower employment means a higher marginal product of labor, which also implies lower marginal costs. Third, lower employment means more leisure and therefore a lower marginal utility of leisure $v'\left(1-L\left(\frac{Y_1}{A}\right)\right)$. Since the opportunity cost of producing output is giving up

Phillips Curve relationship into the LM equation, so it also shifts. This is not terribly important so in order to be able to visualize things in a two-dimensional graph, we are going to disregard it.

leisure, this means lower marginal costs. Therefore, the Phillips Curve shifts down: other things being equal, flexible price producers lower their prices to maintain their desired markups. As in the fully-sticky-price case, productivity shocks have no effect on output and interest rates and they lower employment. With partially sticky prices, we also have a fall in the price level.

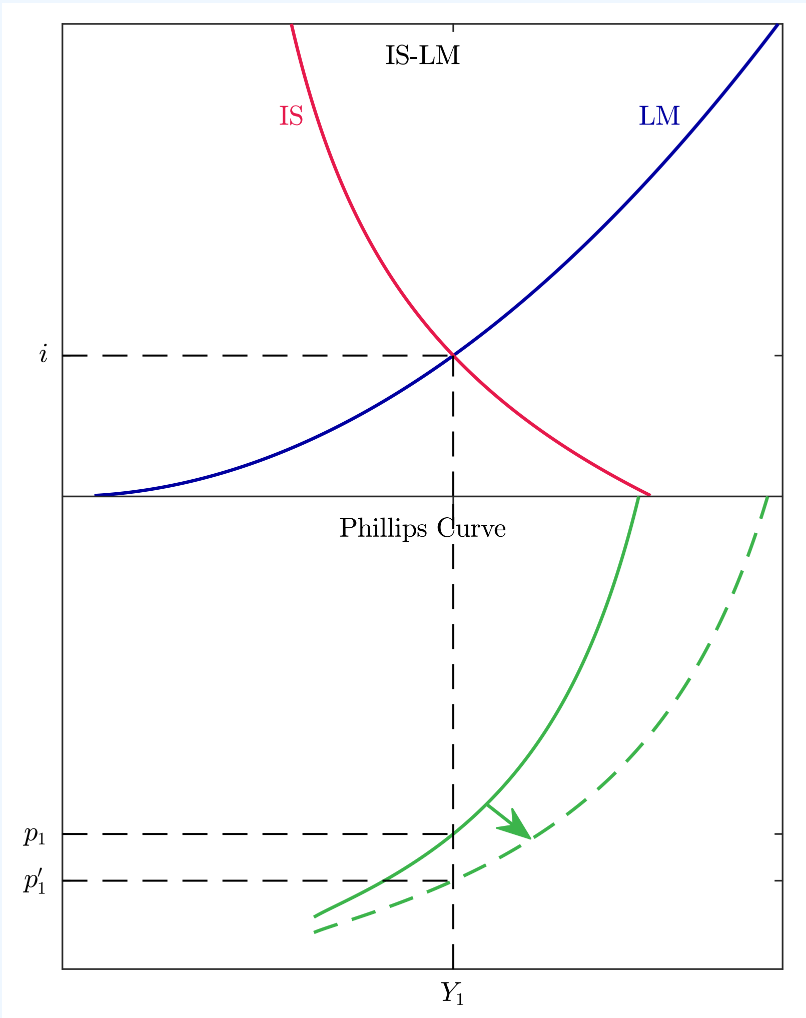


Fig. 14.7.3: *An increase in productivity.*

Exercises

14.1 **Too Many Customers**

Jackson & Jackson is a shampoo manufacturer. Its engineering department estimates that if it decides to produce q units of shampoo, the total production cost would be $c(q) = aq + \frac{b}{2}q^2$ dollars. Its marketing department estimates that if it sets a price of p dollars per unit of shampoo, it will sell $q(p) = \alpha - \beta p$

units, where $\frac{\alpha}{\beta} > a$.

(a) What price should Jackson & Jackson set in order to maximize profits?

(b) Suppose Jackson & Jackson has chosen the profit-maximizing price and advertised it to all its clients, so that at this point it is impossible to change it. Soon after, it learns that the marketing department turned out to be too pessimistic, and actual demand for shampoo is $q(p) = \alpha' - \beta p$, where $\alpha' > \alpha$. If Jackson & Jackson decides to satisfy every purchase order it receives, how many units will it end up selling? How does this number depend on α and α'? Explain.

(c) For what values of α' would it be advantageous for Jackson & Jackson to turn away some customers?

(d) Now suppose that the marketing department had the correct prediction but the engineering department underestimated production costs, which turn out to be $c(q) = a'q + \frac{b}{2}q^2$, with $a' > a$. For what values of a' would Jackson & Jackson want to turn away some customers?

14.2 **Stickiness and the Phillips Curve**

How does the slope of the Phillips Curve (14.7.4) depend on μ? What does this mean?

14.3 **Reserve Requirements**

Suppose the government decides to lower banks' reserve requirements.

(a) What will happen to the (M1) money multiplier?

(b) What will happen to the (M1) money supply?

(c) If prices are partially sticky, what will happen to GDP, interest rates and the price level?

14.4 **Gold**

For a long time, gold was used as money. This meant that the quantity of money depended on the quantity of gold that had been dug up from mines. Suppose a new mine is discovered.

(a) What should we expect to happen to GDP, nominal interest rates and the price level?

(b) How does the answer depend on how flexible prices are?

14.5 **Future Inflation and Present Inflation**

Suppose the economy is well described by a New Keynesian model with partially sticky prices. Let π_1 refer to inflation between period 0 and period 1 and π_2 refer to inflation between period 1 and period 2, which we have taken as exogenous. Suppose there is an exogenous increase in π_2, for instance because the government announces that it will increase the money supply in the future. What will happen to π_1? What are all the steps that lead to this conclusion?

14.6 **The Effect of Investment**

Consider the simplified New Keynesian model with exogenous investment from Section 14.6, with a period-2 production function given by (14.6.1). Suppose the utility function is $u(c) = \log(c)$, the household's discount factor is $\beta = 1$ and inflation is $\pi = 0$.

(a) Derive an IS equation for this special case. This should be a relationship between Y_1 and i for given values of the exogenous parameters A and $\bar{K}$.

(b) Holding the interest rate constant, by how much does output increase if $\bar{K}$ increases by one unit? Explain in words why the answer depends on A.

Now consider a simplified version of the Old Keynesian model with exogenous investment. The consumption function is $c(Y_1) = a + bY_1$ and the period-2 production function is given by (14.6.1).

(c) Use the market-clearing condition to solve for Y_1 as a function of the exogenous parameters A, $\bar{K}$, a and b. Explain in words why it does not depend on i.

(d) By how much does output increase if $\bar{K}$ increases by one unit? Explain in words why the answer depends on b but not on A. Explain why the answer is different from the answer to part (b).

14.7 **Blowing Stuff Up**

Imagine that the economy is well described by the New Keynesian model with perfectly sticky prices, with one twist. There is already some capital in place, which can be used to produce output in period 2. Denote this capital by K_0 and, for simplicity, assume that it does not depreciate. The total capital stock that's used in production in period 2 is therefore:

$$K = K_0 + I$$

where I is investment.

(a) Let $I(K_0, r)$ be the level of investment as a function of K_0 and r. How does it depend on K_0?

(b) Write down the market-clearing condition for period-1 goods.

(c) Derive an IS equation for this economy.

(d) Suppose someone blows up part of the original capital stock, so that we start with $K_0 - X$ instead of K_0. How does the IS curve shift in response to X? What happens to GDP and interest rates? Explain.

14.8 **A Recession**

Consider the following data:

	2015	**2016**
Capital Stock	10,000	10,000
Employment	1	0.81
GDP	4,000	3,600
Labor Income	2,000	1,800
Capital Income	2,000	1,800
Investment	800	600
Consumption	3,200	3,000
Price level	100	90
Nominal Interest Rate	5%	10%

Suppose that 2015 was a totally normal year (almost identical to 2012, 2013 and 2014). This question asks you to think about what are the possible causes of recessions and how to use the data to determine what may have caused the 2016 recession.

(a) Name two possible causes of recessions that *don't* seem to fit the data. For each of them, explain:

   i. Why, according to some model, it's possible for this to cause a recession. You can make the argument using equations, graphs or words, as long as you are precise.

   ii. What features of the data indicate that this is *not* what happened.

(b) Name one possible cause of recessions that *does* fit the data. Explain:

   i. Why, according to some model, it's possible for this to cause a recession. You can make the argument using equations, graphs or words, as long as you are precise.

   ii. What features of the data indicate that this is what may be going on?

CHAPTER 15

Monetary and Fiscal Policy

Most policymakers around the world, especially in central banks, have some version of the New Keynesian model in mind when they are setting policy. In this chapter we'll look at what the model tells us about the effects of monetary and fiscal policy.

15.1 Fiscal Policy

So far we have modeled an economy where there is no government spending. Imagine now that there is a government, which chooses some level of spending G. How does this change the model? What is the effect of an increase in G?

For now let's imagine that there is only government spending in period 1 (Exercise 15.6 will ask you to look at the effect of government spending in period 2). We are going to imagine that government spending is financed entirely by lump-sum taxes and that Ricardian equivalence holds, so we don't need to specify when the government collects taxes. Furthermore, we are going to assume that government spending does not enter the household's utility function, i.e. households do not value public goods (Exercise 15.7 will ask you to look at what happens under different assumptions in this regard).

The period 1 goods market clearing condition is now:

$$Y_1 = c_1 + K + G$$

Solving for c_1 and replacing this in the Euler equation (14.4.4) leads to the following modified version of the IS relationship:

$$u'\left(Y_1 - G - K\left(i - \pi\right)\right) = \beta\left(1 + i - \pi\right)u'\left(F_2\left(K\left(i - \pi\right)\right)\right) \tag{15.1.1}$$

An increase in government spending leads to a horizontal rightward shift of the IS curve, as show in Figure 15.1.1. Output and the interest rise.

What's going on? The government is directly demanding goods and producers are responding by producing more goods. This is, at least in part, the logic behind the "fiscal stimulus" plans that are sometimes carried out. The objective is to make production and employment rise by directly demanding goods and services, with the understanding that producers will expand output to meet demand.

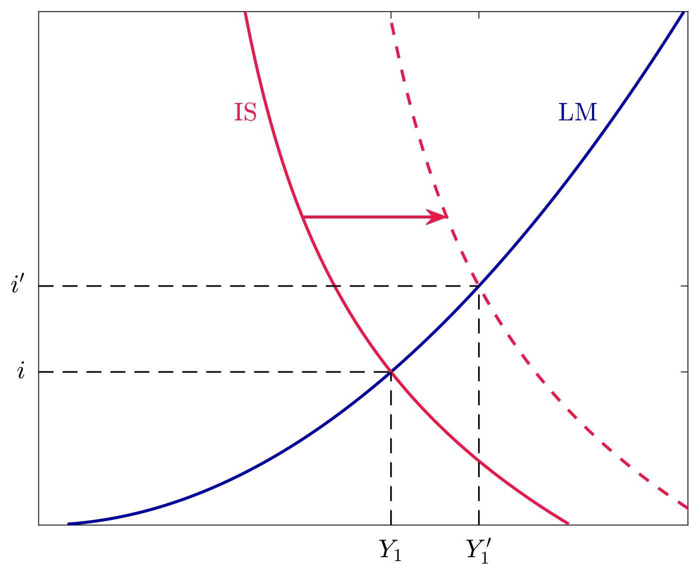


Fig. 15.1.1: *The effects of an increase in government spending.*

The Government Spending Multiplier

An increase in G leads to a rightward shift in the IS curve, resulting in higher GDP and higher interest rates. We'll now look at two further questions about this. First, we'll ask *how much* does the IS curve shift. We'll see that the New Keynesian model and the Old Keynesian model that we looked at in Section 14.6 give different answers to this question. Next, we'll ask how much of the effect will be on GDP and how much on interest rates.

Suppose that there is an increase in G. Holding everything else constant (in particular, holding i constant), how much does the IS curve shift to the right? We'll ask this first with the New Keynesian IS curve and then with the Old Keynesian IS curve. Mathematically, what we'll be trying to compute is:

$$\frac{\partial Y_1}{\partial G} = -\frac{\frac{\partial \Delta}{\partial G}}{\frac{\partial \Delta}{\partial Y_1}}$$

(while holding i constant). For the New Keynesian IS curve (15.1.1) we have:

$$\begin{aligned} \frac{\partial Y_1}{\partial G} &= -\frac{\frac{\partial \Delta}{\partial G}}{\frac{\partial \Delta}{\partial Y_1}} \\ &= -\frac{-u'\left(Y_1 - G - K\left(i-\pi\right)\right)}{u'\left(Y_1 - G - K\left(i-\pi\right)\right)} = 1 \end{aligned} \tag{15.1.2}$$

so the IS curve shifts to the right by the exact amount of G.

Instead, by adding government spending into the Old Keynesian IS (14.6.4) we get:

$$Y_1 = c(Y_1) + K(i - \pi) + G \tag{15.1.3}$$

Therefore, using the implicit function theorem, we have:

$$\begin{aligned} \frac{\partial Y_1}{\partial G} &= -\frac{\frac{\partial \Delta}{\partial G}}{\frac{\partial \Delta}{\partial Y_1}} \\ &= \frac{1}{1 - c'(Y_1)} > 1 \end{aligned} \tag{15.1.4}$$

so the IS curve shifts more than one-for-one with G. For instance, if the marginal propensity to consume is 0.75, the IS curve will shift to the right by 4 dollars for each dollar of government spending.

What explains the difference between (15.1.2) and (15.1.4)? Why does the Old Keynesian model predict that the IS curve shifts more than one-for-one with changes in G, while the New Keynesian model does not? The answer has to do with Ricardian equivalence.

Let's take the New Keynesian model first. When the government increases spending on public goods, there is a direct effect: producers increase output and earn additional income from selling these goods to the government. However, they understand that, either now or in the future, the government will increase taxes to pay for this spending. That's why we don't need to be specific about when the government collects taxes: the present value of these taxes will be exactly equal to the additional income generated by selling goods to the government. Therefore the present value of after-tax income has not changed, so consumption does not change. As a result, the IS curve shifts exactly by the amount of the increase in G.

In the Old Keynesian model, this works differently. Households pay no attention to the fact that future taxes will increase, so Ricardian equivalence does not hold. Therefore we do need to be specific about when the government collects taxes. Assuming households base consumption decisions on current *after-tax* income, then they will react differently to an increase in spending paid for by current taxes (which households pay attention to) or by future taxes (which households ignore). Exercise 15.3 asks you to compute how different the reaction will be. Let's assume that the increase in government spending is paid for entirely with future taxes. (Equation (15.1.3) implicitly assumes this; otherwise consumption would be $c(Y_1 - \tau_1)$). Households just see that they are earning extra income from selling goods to the government, so they go out and consume more. How much more? That depends on the marginal propensity to consume: they will consume an extra $c'(Y_1)$ per dollar of extra income. But this is not the whole story. This additional consumption will lead producers to increase production further, and make them earn extra income, which in turn leads to extra consumption, and so on. Mathematically, what results is a geometric series. For one dollar of additional government spending, we get:

$$\begin{aligned} \frac{\partial Y_1}{\partial G} &= 1 + c'(Y_1) + (c'(Y_1))^2 + (c'(Y_1))^3 + \dots \\ &= \frac{1}{1 - c'(Y_1)} \end{aligned}$$

which is exactly what (15.1.4) is saying.

What we have asked so far is how far the IS curve moves to the right. The overall effect of the change in G will depend on how the IS curve interacts with the LM curve. If the LM curve is very steep, then the interest rate will rise a lot and GDP will increase little. Conversely, if the LM curve is relatively flat, then GDP will increase a lot with little increase in interest rates. Figure 15.1.2 illustrates these different cases. Exercise 15.2 asks you to think about what underlying models of money demand would produce steep or flat LM curves.

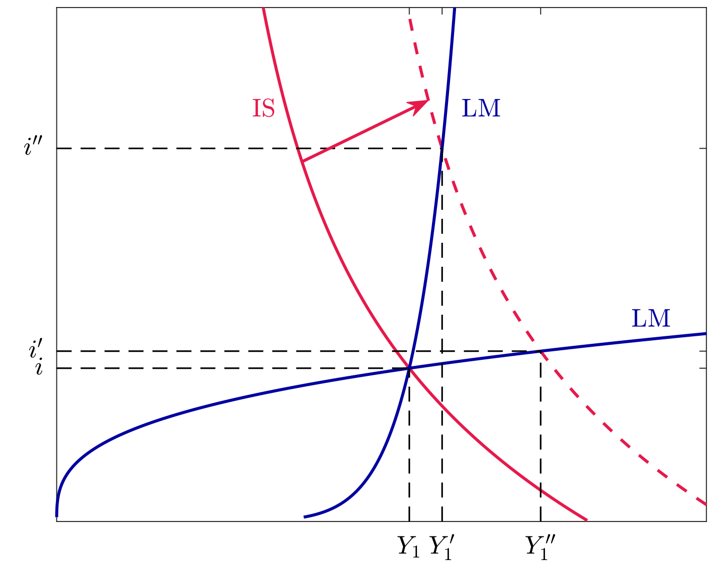


Fig. 15.1.2: *Crowding out.*

The term "crowding out" is used to refer to situations where an increase in G results in higher interest rates but not higher GDP, as shown in Figure 15.1.2 for the case of a steep LM curve. The term comes from the idea that G "crowds out" other types of spending. In particular, it is sometimes said that higher government spending may "crowd out" investment if it leads to a rise in interest rates.

The quantity $\frac{dY_1}{dG}$ is sometimes known as the "government spending multiplier". It is the answer to the question: "if government spending rises by one dollar, by how much does GDP rise?" It's called a "multiplier" because GDP changes by some multiple of the change in G. The size of the multiplier depends on:

1. The size of the shift in the IS curve. As we saw, in our version of the New Keynesian model, this will be equal to 1, whereas the Old Keynesian model says this will be greater than 1.

2. The slope of the LM curve, since this determines the degree of crowding out.

The second point is a little bit subtle. The idea of the multiplier is to ask how GDP changes with an increase in government spending *holding monetary policy constant*. However, holding monetary policy constant could mean more than one thing. Figure 15.1.2 implicitly assumes that by "holding monetary policy constant" we mean "holding the money supply constant". However, monetary policy is often described simply in terms of choosing an interest rate (instead of choosing a money supply that will then result in an interest rate). If by "holding monetary policy constant" we meant "adjusting to money supply so that the interest rate remains constant" then that changes the multiplier. Exercise 15.1 asks you to think about this.

There is a lot of disagreement about the size of the multiplier. In particular, there is disagreement about whether it's greater than 1, as would be implied by an "Old Keynesian" IS curve (combined with either a relatively flat LM curve or a monetary policy response that keeps nominal rates unchanged). Ideally, we would like to have many experiments where G is changed randomly and measure how the economy reacts to these. Since we don't run these sorts of macroeconomic experiments, we need to figure out the right way to interpret the data that we do have, and there is quite a bit of disagreement on how to do that. Ramey (2011) surveys some of the evidence on measuring the multiplier and finds that the most plausible values are between 0.8 and 1.5.

The size of the multiplier is important because, if the multiplier is large, then fiscal policy can be very powerful: a relatively small change in the level of government spending can have a large effect on GDP.

15.2 Monetary Policy

Suppose that there is some shock that, other things being equal, would lead to a recession. For instance, suppose that households become pessimistic about future productivity, so the IS curve shifts to the left. Other things being equal, this would lead to lower GDP and a lower interest rate.

Suppose the government wants to prevent GDP from falling. One possible way to respond is to increase the money supply. This will lead to a rightward shift in the LM curve, further lowering interest rates and counteracting the fall in output, as shown in Figure 15.2.1.

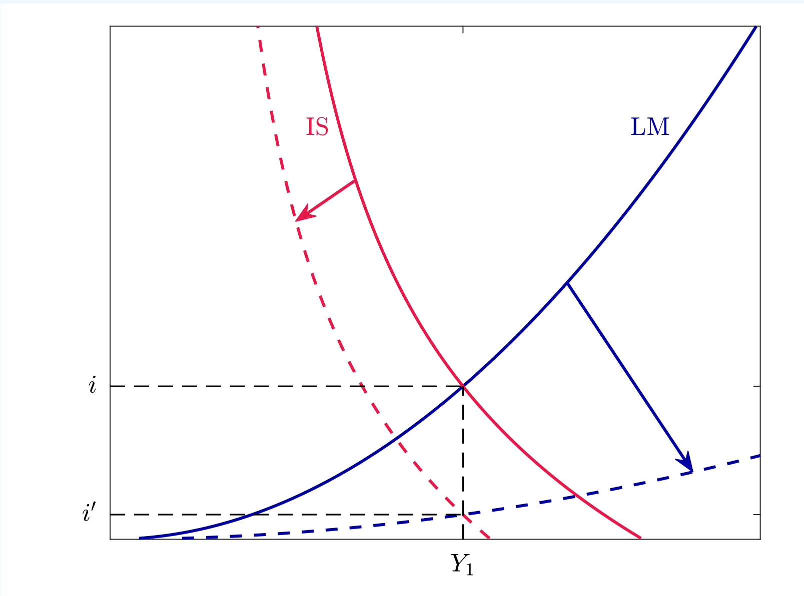


Fig. 15.2.1: *Using monetary policy to counteract a negative shock.*

What's going on? Pessimism about the future, other things being equal, leads households to reduce consumption and investment, which leads to lower output. In order to persuade households not to reduce their spending, the government tries to engineer a fall in the interest rate, to generate movement *along* the new IS curve. A lower interest rate means that more investment projects are worth doing, so investment rises, and present goods are cheaper relative to future goods, so consumption rises. If the government gets the size of the policy reaction exactly right, then it can offset the fall in GDP exactly, as in the example in Figure 15.2.1.

Often this type of policy is simply described as "lowering the interest rate". Ultimately, by controlling the money supply, the Central Bank can control the interest rate, so it is sometimes useful to think of the Central Bank as just picking what interest rate it wants. Indeed, in practice that's how most central banks operate these days. They decide on a target level for the interest rate and then conduct open market operations to adjust the money supply however much it takes for the target they chose to actually be the equilibrium rate.

Note that the objective of policy need not be to stabilize GDP. As we saw in Chapter 13, a completely efficient economy where the First Welfare Theorem holds may still have fluctuations in GDP in response to shocks of various kinds. One possible objective for monetary policy is to get the economy to behave the way it would if prices weren't sticky, which would not entail complete stabilization. One challenge in implementing this objective is that it's hard to know in real time what kinds of shocks are taking place, which makes it hard to decide when monetary policy should refrain from attempting to stabilize GDP.

Expectations-Augmented Phillips Curve

For a long time it was believed that the Phillips Curve implied a fundamental tradeoff: higher output (and employment) was thought to necessarily go together with higher inflation. Policymakers, the argument went, faced a choice: would they rather increase employment or lower inflation? If policymakers wanted to increase employment, then the Phillips Curve implied that they had to be willing to tolerate higher inflation. Conversely, if they wanted to combat inflation, then they had to be willing to tolerate lower employment.

No matter what choice the policymakers made, it was believed that macroeconomic policy offered the tools to pick any point on the Phillips Curve. Figure 15.2.2 illustrates how, according to the model, monetary policy can be used to increase output and employment, while inducing higher inflation. The figure shows an increase in the money supply. This shifts the LM curve to the right, leading to higher GDP. Since marginal costs have risen, flexible price producers raise prices, generating inflation. Conversely, in order to reduce inflation, the same policy can be used in reverse, which lowers inflation, employment and GDP.

The Phillips Curve defined by equation (14.7.4) takes as given the level of sticky prices p_1^S. Let's now go back and think about how sticky-price producers set their prices. Assume that these producers are smart. They understand that, once they have chosen a price, they will be stuck with it. So they will try to choose a price that is approximately right on average. Specifically, let's assume that they form an expectation of what the flexible price producers will do and then set their own price equal to that.

$$p_1^S = \mathbb{E}\left(p_1^F\right) \tag{15.2.1}$$

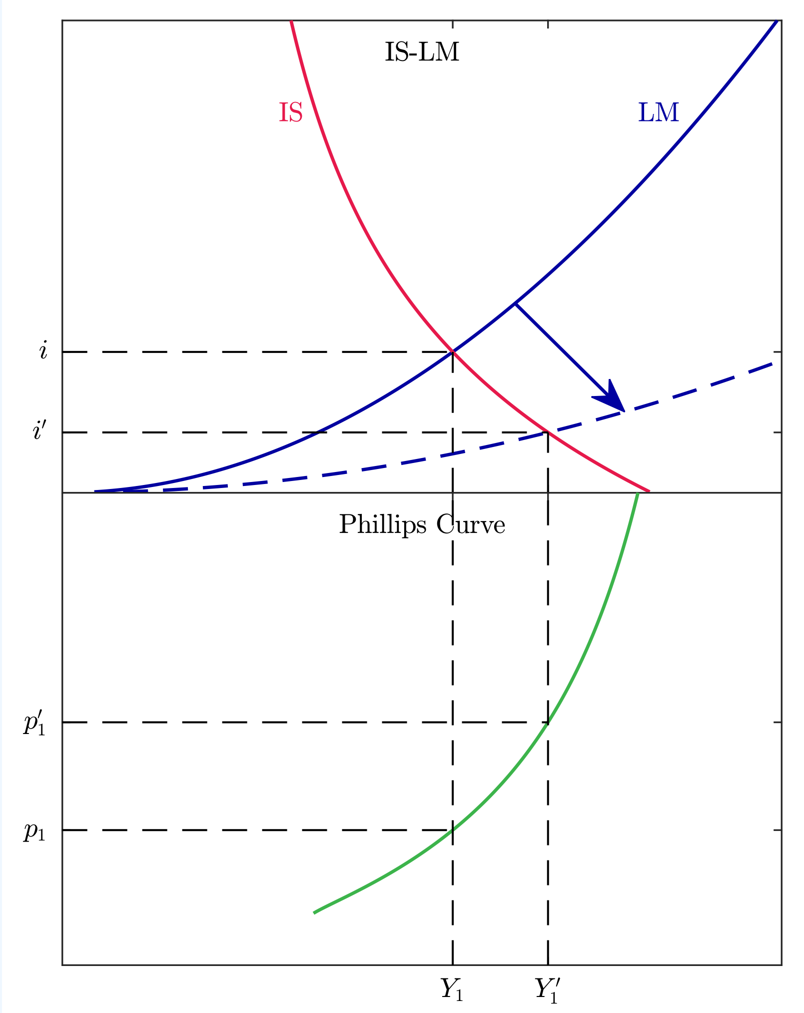


Fig. 15.2.2: *The effects of monetary expansion on GDP, inflation and interest rates.*

Now let's compute the price index. Using (14.7.2), we get:

$$p_1 = \mu \mathbb{E}\left(p_1^F\right) + (1-\mu)\, p_1^F \tag{15.2.2}$$

The average price is a weighted average of what sticky producers thought that flexible producers would do and what they ended up doing. The *expected* average price is then:

$$\begin{aligned}\mathbb{E}\left(p_1\right) &= \mathbb{E}\left[\mu \mathbb{E}\left(p_1^F\right) + (1-\mu)\, p_1^F\right] \\ &= \mathbb{E}\left(p_1^F\right)\end{aligned} \tag{15.2.3}$$

so it's also equal to the expectation of what the flexible-price producers will do.

Depending on what shocks and policies end up taking place, the actual average price might be different from what the sticky price producers expected. Define the price surprise ϵ as:

$$\epsilon \equiv \frac{p_1}{\mathbb{E}(p_1)} \tag{15.2.4}$$

ϵ measures the ratio between the *actual* average price and the *expected* average price. $\epsilon > 1$ means that the price level turned out higher than expected, so there was higher-than-expected inflation. Conversely, $\epsilon < 1$ means lower-than-expected inflation. Replacing (15.2.2) and (15.2.3) into (15.2.4):

$$\begin{aligned}\epsilon &= \frac{\mu \mathbb{E}\left(p_1^F\right) + (1-\mu) p_1^F}{\mathbb{E}\left(p_1^F\right)} \\ &= \mu + (1-\mu) \frac{p_1^F}{\mathbb{E}\left(p_1^F\right)}\end{aligned} \tag{15.2.5}$$

Since the sticky price producers are stuck, the price surprise depends only on how the flexible-price producers deviated from what was expected.

Now turn to the flexible price producers. They will set prices as a markup over marginal costs, according to equation (14.7.3). Using (15.2.1) to replace p_1^S, this reduces to:

$$\begin{aligned}p_1^F &= \mathbb{E}\left(p_1^F\right) \frac{v'(1-L)}{u'(c_1(L)) F_1'(L)} \frac{\eta}{\eta-1} \\ \Rightarrow \frac{p_1^F}{\mathbb{E}\left(p_1^F\right)} &= \frac{v'(1-L)}{u'(c_1(L)) F_1'(L)} \frac{\eta}{\eta-1}\end{aligned} \tag{15.2.6}$$

Finally, replacing (15.2.6) into (15.2.5) and rearranging we get:

$$\frac{\epsilon-\mu}{1-\mu} = \frac{v'(1-L)}{u'(c_1(L)) F_1'(L)} \frac{\eta}{\eta-1} \tag{15.2.7}$$

What is equation (15.2.7) telling us? It says there is a positive relationship between ϵ and L, i.e. between price surprises and employment. Once we take into account how the sticky price producers set their prices, the model doesn't quite predict a relationship between inflation and employment. It only gives us a relationship between *higher-than-expected* inflation and employment. Equation (14.7.4) gives us a Phillips Curve only because we were holding p_1^S constant.

Rational Expectations, the "Natural Rate" and the Value of Commitment

Suppose it turns out that $\epsilon = 1$, which means that the sticky price producers got their expectation of p_1^F exactly right. Then (15.2.7) reduces to:

$$\frac{v'(1-L)}{u'(c_1(L))} = \frac{\eta-1}{\eta} F_1'(L) \tag{15.2.8}$$

which is exactly condition (14.2.4) for the flexible-price case. This implies that if there were never any price surprises, the New Keynesian model would behave exactly like the RBC model.[1] Equation (15.2.8) is sometimes described as a vertical Phillips Curve. It is vertical because inflation does not show up anywhere, so if you were to plot it in terms of inflation against L you would find a vertical line, with the same level of employment for any possible level of inflation.

The level of economic activity that comes out of equation (15.2.8) is known as the “natural rate” (of employment, unemployment, GDP, etc.). The natural rate is the answer to the question: “if prices were completely flexible, what is the level of this variable that would prevail?” One way of restating the meaning of equation (15.2.7) is that output and employment can only deviate from their natural rate if there is higher-that-expected or lower-than-expected inflation.

How exactly do people form their expectations? One popular hypothesis is that they form these expectations rationally.[2] What does this mean? This means that expectations are derived correctly from knowledge of how the economy works, including how the government usually behaves, and are updated on the basis of all available information. Does this mean that expectations are always correct, so that (15.2.8) always holds? No. Even with rational expectations, people understand that random factors will cause inflation to differ from what they predicted. But the predictions themselves are not systematically biased, so they are correct on average.

If the rational expectations hypothesis is correct, higher *average* inflation cannot lead to higher *average* employment. Rational expectations imply that producers expect the high average inflation, so it doesn't come as a surprise and does not lead to higher output. Therefore, on average, output and employment are equal to their natural rate, no matter what the level of inflation is. In other words, if expectations are rational, then over longer periods there is *no* tradeoff between inflation and employment, and the long-run Phillips Curve is indeed vertical.

Suppose we accept the hypothesis of rational expectations. How much inflation should producers expect? This depends on a lot of factors. One in particular is very important: how producers expect the government to behave. Let's think a little bit about what the government might want to do and how that gets built into expectations. Formally, what are going to do is describe a game, first analyzed by Barro and Gordon (1983), between the government and sticky-price producers. The government tries to pursue beneficial policies taking as given the decisions that sticky-price producers have taken and the sticky-price producers, when setting prices, try to anticipate what the government will do.

Let's put ourselves in the position of a government that has to decide on macroeconomic policy. Sticky-price producers have already set their prices at $p_1^S = \mathbb{E}(p_1)$. The government knows that, by changing monetary policy, it can change the level of GDP, and understands the relationship between GDP and inflation implied by the expectations-augmented Phillips Curve (15.2.7). Assume the government is benevolent. What does the government want to do?

Let Y_1^* be the ideal level of GDP that the government would like to attain. What do we know about Y_1^*? First: Y_1^* is not infinity. In order to increase output, producers need to work more. At some point, the

[1] More precisely, like an RBC model where there is either a tax or monopoly power so that the term $\frac{\eta-1}{\eta}$ is present.

[2] Another hypothesis is that they just extrapolate from recent experience: if inflation was 3% last year, they expect 3% again this year. This is known as “adaptive” expectations.

marginal utility of leisure will be greater than the marginal utility of the consumption goods that you can obtain by working harder. A benevolent government wants to increase GDP up to the point where equation (13.1.3) from the RBC model holds, but no more. (Remember, in the RBC model, the First Welfare Theorem holds)

Second: Y_1^* is greater than the "natural" rate of output Y_1^N, which is defined by $Y_1^N = F(L)$ with the level of L that satisfies (15.2.8). If we compare equations (13.1.3) and (15.2.8), we can see that the difference comes from the term $\frac{\eta-1}{\eta}$, which measures monopoly power. If $\eta \to \infty$, then we are back to perfect competition and the natural rate of output coincides with what a benevolent government wants to attain. Away from that limit, the government would like to undo the effects of monopoly power on the economy. The government reasons: "All these monopolist producers are producing a bit less than they should, reducing quantity in order to keep their (relative) price high. Collectively, their efforts are self-defeating: they cannot all raise their price relative to each other, and the only effect is to decrease total output. If, by raising demand, I can get them all to increase output, this will make everyone better off".

Let's imagine that the government tries to balance two objectives: getting Y_1 as close as possible to Y_1^* and keeping inflation close to zero. The government's objective is to maximize:

$$W = -\left(Y_1 - Y_1^*\right)^2 - \phi \pi_1^2 \tag{15.2.9}$$

Equation (15.2.9) says there are two things the government tries to avoid: GDP away from Y_1^* and inflation away from zero. The parameter ϕ measures how much the government cares about each of the objectives. A high value of ϕ means the government really dislikes inflation. The quadratic terms in the objective imply that large deviations from the target are disproportionately more painful than small deviations.[3] In other words, the marginal cost of deviating from target is increasing in the size of the deviation.

The government cannot just choose any values of Y_1 and π_1 it wants. If it could, the solution would be simple: $Y_1 = Y_1^*$ and $\pi_1 = 0$. Unfortunately, the government is constrained by the Phillips Curve: policies that raise GDP will also raise inflation. Rather than work with the full-blown Phillips Curve (15.2.7), let's propose a simplified version that captures the same idea:

$$\pi_1 - \mathbb{E}\left(\pi_1\right) = a\left(Y_1 - Y_1^N\right) \tag{15.2.10}$$

The economic content of equation (15.2.10) is just like that of equation (15.2.7): output will be above its natural rate if and only if inflation is higher than expected. The only difference is that we have written a simplified, linear version, instead of the original (15.2.7). a is just a parameter which governs the slope of the Phillips Curve; higher a means a steeper Phillips Curve. Note that the Central Bank takes $\mathbb{E}(\pi_1)$ as given: by the time the Central Bank chooses policies, sticky-price producers have already made their decisions based on $\mathbb{E}(\pi_1)$ and cannot change them.

[3] We'll just take the fact that the function is quadratic as an assumption, although it can be justified as a second-order Taylor approximation to the representative household's utility.

The Central Bank solves:

$$\max_{Y_1,\pi_1} -\left(Y_1 - Y_1^*\right)^2 - \phi\pi_1^2$$
$$s.t. \tag{15.2.11}$$
$$\pi_1 - \mathbb{E}\left(\pi_1\right) = a\left(Y_1 - Y_1^N\right)$$

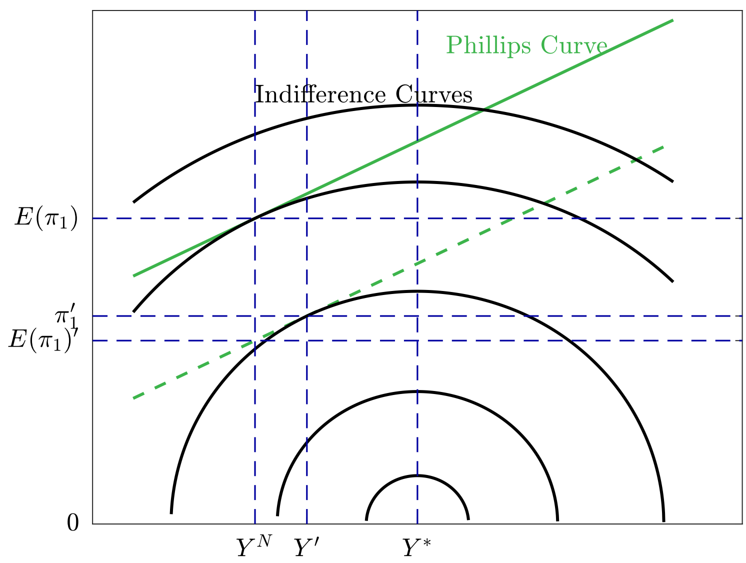


Fig. 15.2.3: *The Central Bank's inflation-output trade-off given a Phillips Curve.*

Figure 15.2.3 shows the Central Bank's problem graphically. Once expectations have been set at $\mathbb{E}(\pi_1)$, the Phillips Curve (15.2.10) is a constraint which limits the combinations of output and inflation that are attainable. The slope of the Phillips Curve is given by the parameter a and it goes through the point $\left(Y_1^N, \mathbb{E}(\pi_1)\right)$: if inflation is equal to expectations, output will be at its natural level. The objective function (15.2.9) implies that the Central Bank's indifference curves are ellipses centered on the ideal outcome $Y_1 = Y_1^*$, $\pi = 0$: the further away from this outcome, the worse it is for the Central Bank. The Central Bank optimizes by choosing a point that is on the best indifference curve that is consistent with the Phillips Curve.

The Lagrangian for problem (15.2.11) is:

$$L(Y_1, \pi_1, \lambda) = \left(Y_1 - Y_1^*\right)^2 - \phi\pi_1^2 - \lambda\left[\pi_1 - \mathbb{E}\left(\pi_1\right) - a\left(Y_1 - Y_1^N\right)\right]$$

with first order conditions::

$$-2\left(Y_1 - Y_1^*\right) + \lambda a = 0$$
$$-2\phi\pi_1 - \lambda = 0$$

and therefore:

$$\begin{aligned}
-2\left(Y_1 - Y_1^*\right) - 2\phi\pi_1 a = 0 && \text{(replacing } \lambda = -2\phi\pi_1\text{)} \\
-2\left(\frac{\pi_1 - \mathbb{E}\left(\pi_1\right)}{a} + Y_1^N - Y_1^*\right) - 2\phi\pi_1 a = 0 && \text{(using (15.2.10) to replace } Y_1\text{)} \\
\Rightarrow \pi_1 = \frac{a}{1+\phi a^2}\left(Y_1^* - Y_1^N\right) + \frac{1}{1+\phi a^2}\mathbb{E}\left(\pi_1\right) && (15.2.12)
\end{aligned}$$

Equation (15.2.12) tells us what inflation level the government will choose as a function of:

- The distance between the target level of output and the natural level $Y_1^* - Y_1^N$. If this distance is large, the government's desire to raise output is strong, so the government will be willing to bring about higher inflation in order to increase output.
- Expected inflation $\mathbb{E}\left(\pi_1\right)$. If inflation expectations are high, achieving low *actual* inflation means creating a negative inflation surprise, which is costly in terms of output. Therefore the government will respond to higher expected inflation with higher actual inflation. Notice that

  $$\frac{\partial \pi_1}{\partial \mathbb{E}\left(\pi_1\right)} = \frac{1}{1+\phi a^2} < 1$$

  so the government responds less than one-for-one to expected inflation. If the Phillips Curve is very steep (high a) or the government strongly dislikes inflation (high ϕ) then it will respond little to inflation expectations. Instead, if the Phillips Curve is flat or the government doesn't mind inflation very much it will be more responsive of inflation expectations.

Now let's impose the hypothesis of rational expectations. In this context, rational expectations means that sticky-price producers have figured out (15.2.12). They understand what the government is trying to do and how it trades off its different objectives. Therefore the rational way to set expectations is to set

$$\mathbb{E}\left(\pi_1\right) = \pi_1 \tag{15.2.13}$$

What are (15.2.12) and (15.2.13) telling us? The government sets inflation in response to inflation expectations, but under rational expectations, expected inflation rationally anticipates what the government will want to do. Using the rational expectations hypothesis, we can solve for what output and inflation will be. First, replace (15.2.13) into (15.2.12) and solve for π_1:

$$\begin{aligned}
\pi_1 &= \frac{a}{1+\phi a^2}\left(Y_1^* - Y_1^N\right) + \frac{1}{1+\phi a^2}\pi_1 \\
\Rightarrow \pi_1 &= \frac{1}{\phi a}\left(Y_1^* - Y_1^N\right) && (15.2.14)
\end{aligned}$$

Now replace (15.2.13) into (15.2.10) and solve for Y_1:

$$\begin{aligned}
0 &= a\left(Y_1 - Y_1^N\right) \\
\Rightarrow Y_1 &= Y_1^N && (15.2.15)
\end{aligned}$$

Equations (15.2.14) and (15.2.15) tell us what the level of output and inflation will end up being. Let's focus on (15.2.15) first. This says that output will be exactly at the natural rate, so the government will not have any success at all in raising it. Why is this? According to (15.2.10), the only way to raise output above the natural rate is with higher-than-expected inflation. But under rational expectations, inflation cannot be higher-than-expected.[4] Hence the government's attempts to raise output will be futile.

Figure 15.2.3 shows this result graphically. If expectations were set at some level like $\mathbb{E}(\pi_1)'$, the Phillips Curve would be the dotted line. The Central Bank's optimal decision would be to choose output Y' and inflation π_1'. But this inflation does not coincide with what the market expected, so this outcome is inconsistent with rational expectations. The rational-expectations outcome is for the market to expect $\mathbb{E}(\pi_1)$, which leads to the solid Phillips Curve. When faced with this Phillips Curve, the Central Bank chooses Y^N and $\pi_1 = \mathbb{E}(\pi_1)$, which confirms the market's expectation.

Let's now think about what (15.2.14) is telling us. This economy will experience positive inflation. Inflation will be higher if $Y_1^* - Y_1^N$ is high, ϕ is low or a is low. Why is this? Each of these factors means that, taking $\mathbb{E}(\pi_1)$ as given, it is very attractive for the government to choose higher inflation. High $Y_1^* - Y_1^N$ means that raising GDP above the natural rate is highly desirable; low ϕ means that inflation is not too unpleasant; low a (a not-too-steep Phillips Curve) means that the rise in inflation per unit of additional GDP is not large. Due to rational expectations, the factors that make choosing higher GDP and inflation desirable are fully anticipated, so they end up leading to inflation, but *not* to higher GDP.

The government in this problem faces what's known as a "time-inconsistency" problem. What does this mean? Imagine that the government could announce a level of inflation *before* sticky producers set their prices, and was then committed to sticking to the announcement. This would change the government's problem entirely. If the government is committed to an inflation level, there can never be surprise inflation, and hence GDP will be at its natural rate. Knowing that it will be committed, the government no longer has any reason to choose a level of inflation higher than zero, so we'd end up with $\pi_1 = 0$ and $Y_1 = Y_1^N$. This is a strictly better outcome than the problem with no commitment: it has the same level of GDP but lower inflation. Notice that it's important for the government to actually commit to this. If, once expectations have been set, the government could disregard its commitments, it would want to deviate from its announced plan and set the inflation given by (15.2.12) instead. "Time inconsistency" refers to the fact that the government would like to commit to an outcome, but then has incentives to undo this commitment.

This type of argument has been extremely influential in the design on macroeconomic policy institutions. In the last couple of decades many countries have introduced reforms to make their central banks more independent of elected governments. The idea behind this is to try to isolate monetary policy from the forces that push policymakers (for entirely benevolent reasons!) to renege on their commitment to low inflation. Rogoff (1985) argued that one way in which society could deal with the time-inconsistency problem is by appointing a "conservative central banker", i.e. Central Bank authorities who dislike inflation more than the average person. The logic is that if the central banker's preferences have very high ϕ, then equation (15.2.14) implies that inflation will be lower, and output will end up at its natural rate anyway.

[4]In this model there are no shocks. In a model with shocks, rational expectations would mean that inflation cannot be higher than expected on average; there could be shocks that lead to higher-than-expected inflation as long as there are other shocks that lead to lower-than-expected inflation.

15.3 Monetary Policy Regimes

As we have seen, expectations in general, and expectations about monetary policy in particular, are important determinants of economic decisions such as setting prices. This forces us to think about monetary policy *regimes* and not just isolated monetary policy decisions. A monetary policy regime is a set of norms, objectives or procedures that govern how monetary policy is chosen. If expectations are rational, they will depend on the type of policy regime that is in place. One possible regime, which is implicit in the Barro and Gordon (1983) game, is for the Central Bank to have full discretion to pursue objective (15.2.9) at every point in time. As we saw, this tends to produce more inflation than is socially desirable even if (especially if!) the Central Bank is benevolent. Other monetary policy regimes attempt to overcome the inflation bias that comes with discretionary policy.

Money Growth Rules

One extreme form of non-discretionary policy is for the Central Bank to commit to keeping the money supply growing at a constant rate, for instance 2% per year, or even not growing at all. If this commitment is firm (a big if!), then the Central Bank is not be able to *choose* a point in the output/inflation tradeoff and thus cannot yield to the temptation to increase GDP and inflation with monetary expansion: monetary policy is chosen automatically by whatever the money-growth rule specifies. If expectations are rational (another big if!), they anticipate this, so a sufficiently tight money-growth rule can be effective in keeping inflation low.

One problem with a money-growth rule is that it is quite inflexible. (That's the point, it's designed to be inflexible). This means that it prevents monetary policy from acting to stabilize the economy in response to shocks. For instance, the actions illustrated in Figure 15.2.1, where the money supply is increased to bring down interest rates and stabilize GDP, would be precluded by a strict money-growth rule. The inflexibility of a money-growth rule may lead the economy to be more volatile than it otherwise could be. Partly as a result of this, money growth rules have fallen out of favor in recent years. Exercise 15.8 asks you to go over an example where changes in transactions technology create volatility under a money-growth rule.

The Gold Standard and Fixed Exchange Rates

Another extreme form of a non-discretionary policy, which was important historically, is the so-called Gold Standard. Under the simplest version of this system, the Central Bank commits to exchange currency or Central Bank reserves for gold at a fixed exchange rate. Modern versions, which are conceptually quite similar, use a foreign currency like the US dollar or the Euro instead of gold. If the Central Bank is committed to a fixed exchange with respect to either gold or a foreign currency (again, a big if), this removes any discretion it might have in setting monetary policy: it must adjust the monetary base in response to any requests to exchange gold for currency or vice versa. This prevents the Central Bank from actively choosing a point in the inflation/output tradeoff. If the regime is credible, this can be effective in keeping inflation low.

It has the disadvantage that it fixates on the price of one particular good. In the gold standard, it's the price of gold. In a fixed exchange rate, it's the price of the US dollar (or more broadly, the imported goods that can be bought for US dollars at relatively stable nominal prices). If the relative price of this good with respect to other goods changes, then that will have side-effects on the economy. For instance, suppose that jewelry becomes very fashionable, so gold becomes expensive relative to other goods. The Central Bank has committed to keeping the nominal price of gold in terms of currency fixed. This means that the nominal price of other goods has to fall (otherwise, everyone would rush to sell their currency to the Central Bank to get gold). The Central Bank will be forced to contract the money supply, creating deflation and a fall in GDP. These types of side effects led Keynes to call the gold standard a "barbarous relic".

Inflation Targeting

Inflation targeting has become popular in recent decades. Under this regime, the Central Bank is given a formal mandate to keep inflation as close as possible to a pre-specified target (2% per year is a typical figure). In the purest version, the Central Bank is supposed to ignore any other objective it might have, such as keeping employment high, and just focus on keeping inflation as close as possible to the target. In principle, this approach can allow the Central Bank to respond to shocks without giving it the full discretion in setting policy that can lead to inflation bias.

Figure 15.3.1 shows how this would work, continuing the example from Figure 15.2.1. In the example, worsening expectations about the future make the IS curve shift to the left. If the Central Bank did not react to this (for instance, because it was following a zero-money-growth rule), then output would fall from Y_1 to Y_1' and the interest rate would fall to i'. This would induce movement along the Phillips Curve, so that the price level would be p_1' instead of p_1. A Central Bank that followed an inflation target *would* react to this because it doesn't want inflation to fall below its target. In order to keep the price level at p_1 it will increase the money supply, bringing the interest rate all the way down to i''. This would have the double effect of keeping the inflation rate on target and restoring output to its original level. Therefore even though the Central Bank is instructed to only worry about inflation, by doing so it also stabilizes output and employment. Blanchard and Gali (2007) refer to this as a "divine coincidence".

Figure 15.3.2 shows how inflation targeting would be applied to the example from Figure 14.7.3. In the example, a temporary increase in productivity leads to a shift in the Phillips Curve. If there is no reaction from monetary policy, output does not change and employment falls. However, staying at the original output Y_1 while the Phillips Curve has shifted down would mean that the price level is p_1' instead of p_1, so the Central Bank would miss its inflation target. (In the background, what's going on is that flexible-price firms set lower prices because higher productivity means lower marginal costs). In order to hit its inflation target, the Central Bank increases the money supply, lowering the interest rate until the increase in consumption and investment means that output reaches Y_1', which restores the intended price level. In this example, inflation targeting does not stabilize output. Instead, it makes output react to the shock in a way that it wouldn't under a fixed-money-growth rule. Note, however, that in the Pareto efficient RBC economy, output *would* react to a productivity shock: if productivity is temporarily high it is desirable to produce more output.

Taylor Rules

A Taylor Rule, named after Taylor (1993), is similar to an inflation target but with a built-in procedure for correcting mistakes. The Central Bank first sets a target π^* for the level of inflation it wants to attain. Then,

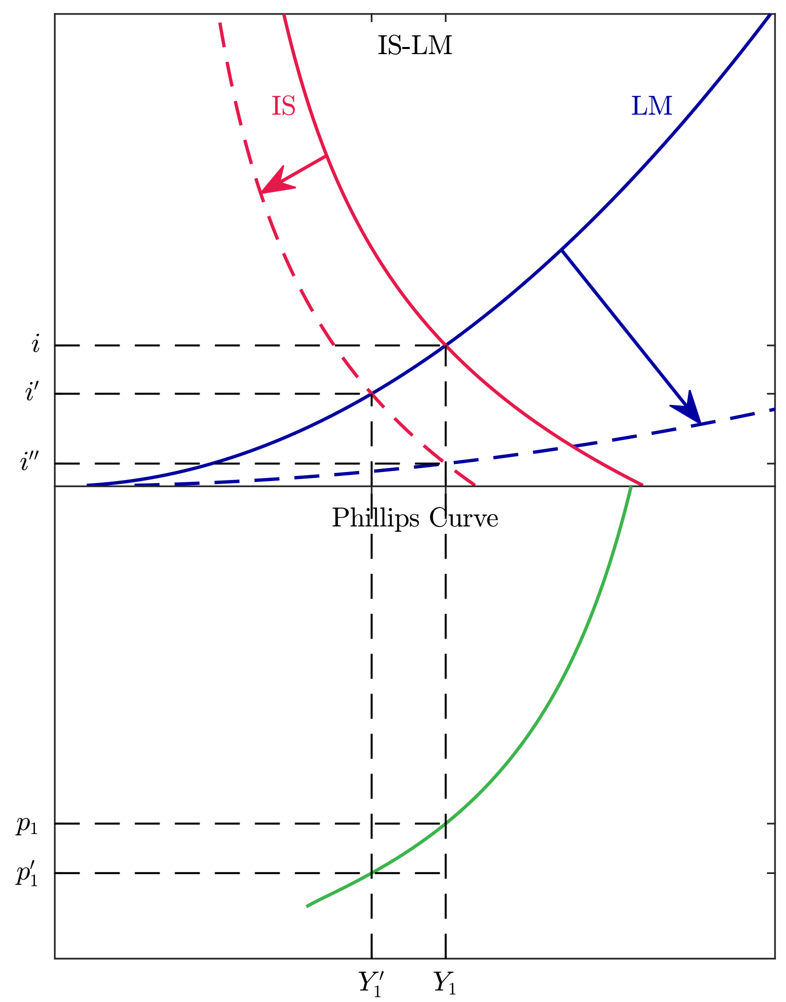


Fig. 15.3.1: *Monetary policy under inflation targeting.*

at any point in time it sets (i.e. targets by adjusting the money supply) the nominal interest rate i according to a formula like:

$$i = r^N + \pi + a(\pi - \pi^*) + b(Y - Y^N) \tag{15.3.1}$$

where:

- Y^N is the natural level of real GDP; Y is its current level.
- π^* is the target level of inflation; π is its current level.
- r^N is the natural level of the real interest rate.
- a and b are parameters; a typical level is $a = b = 0.5$.

What's the idea behind a Taylor Rule? Monetary policy attempts to keep the economy at the natural level

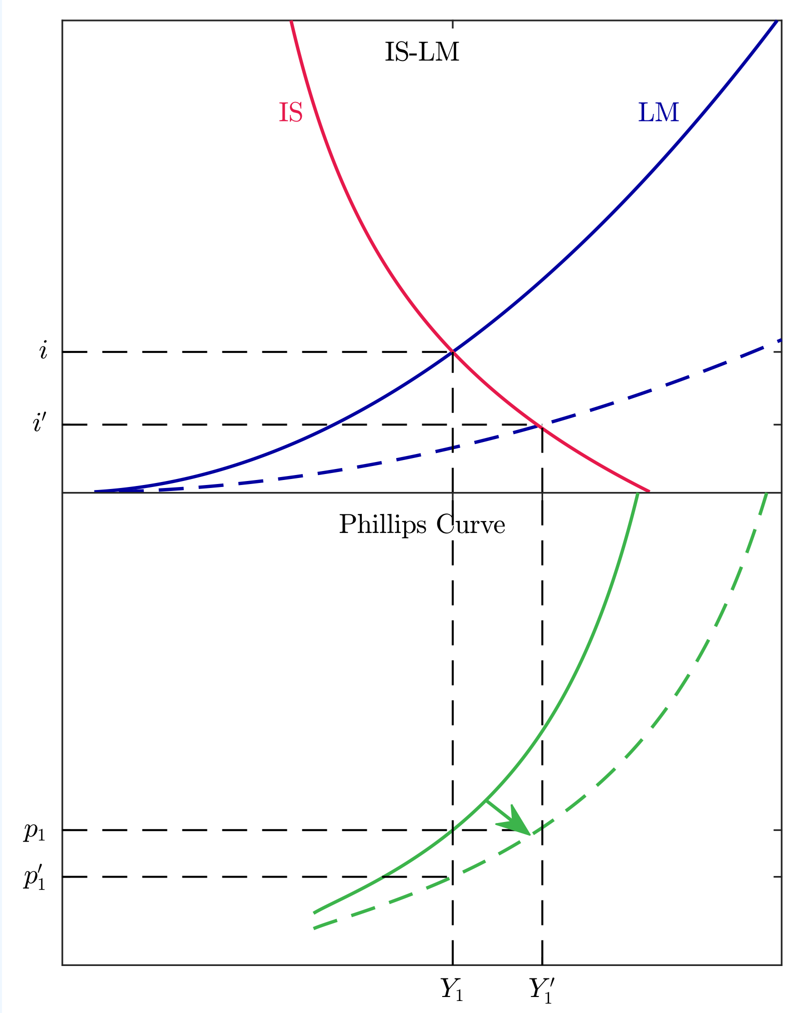


Fig. 15.3.2: *Inflation targeting response to a productivity shock.*

of GDP Y^N and the inflation target π^*. Note that the target is the natural level of GDP Y^N and *not* the first-best level Y^*, which would necessitate higher-than-expected inflation. If the Central Bank is succeeding in keeping both GDP and inflation on target, then it just sets the nominal interest rate at $i = r^N + \pi^*$: the natural real interest plus target inflation. (One practical challenge is to accurately measure Y^N and r^N in real time). If inflation starts to deviate above the target, the Central Bank adjusts the interest rate upwards. Crucially, since $a > 0$:

$$\frac{\partial i}{\partial \pi} = 1 + a > 1$$

so the nominal interest rate adjusts more than one-for-one to deviations of inflation from its target. This is sometimes known as the "Taylor principle". It means that if inflation starts to rise above its target, the Central Bank increases the *real* interest rate. This implies a upward/leftward movement along the IS curve, lowering GDP, and a downward/leftward movement along the Phillips Curve, lowering inflation. In other words, if inflation starts to get out of hand, the Central Bank is willing to generate a recession in order to bring it back in line. Conversely, if the economy starts to experience a recession, with Y below its natural level Y^N, then the Central Bank will lower the interest rate, generating inflation and a rise in GDP. The Central Bank is willing to tolerate some inflation in order to prevent output from falling below its target.

One useful metaphor is to think of the Taylor Rule as a thermostat. A thermostat adjusts the intensity of the boiler to keep the temperature close to a target. Under a Taylor Rule, the Central Bank adjust the interest rate to keep both inflation and GDP close to their targets. The parameters a and b indicate how strongly the Central Bank responds to deviations in inflation and GDP respectively. Pure inflation targeting would correspond to $a \to \infty$ and $b = 0$, where the Central Bank does whatever it takes to keep inflation at its target and does not look at GDP.

Interpreting the Historical Data

We saw in Chapter 12 that the Phillips Curve relationship sometimes seems to hold but not always. Let's see if we can make sense of that by considering different possible policy regimes and different shocks that the economy might experience.

Suppose first that monetary policy does not respond to macroeconomic shocks (for instance because we are under a fixed-money-growth regime) and the main shocks have to do with shifts in either the IS or the LM curve: changes in expectations of future productivity as in Figure 14.5.2; in the money supply as in Figure 14.5.3; in government spending, as in Figure 15.1.1, etc. In this case expectations of inflation will be approximately constant, so whenever inflation deviates from its usual level it comes as a surprise. Then any shocks will lead to movements *along* a fixed Phillips Curve. Expansionary shock will lead to higher output, employment and inflation, and vice versa. If we look at data generated by an economy like this, we will see a clear Phillips Curve. Arguably, this was a plausible description of the US economy until the mid-1960s.

Suppose instead that monetary policy follows inflation targeting (and follows it perfectly) and shocks consists of some mixture of shocks that move the IS curve and productivity shocks that, as we saw in Figure 15.3.2, do not. Then we will see that inflation is almost constant no matter what the shocks are. For shocks that move the IS curve, monetary policy will respond so that neither inflation nor GDP react at all. For productivity shocks, monetary policy will respond as in Figure 15.3.2. Inflation will not react but GDP and employment will. If we look at data generated by an economy like this, the Phillips Curve will look almost perfectly flat. Arguably, this has been a plausible description of the US economy since the late 1980s.[5]

Now suppose that the monetary policy regime is not entirely clear, so the main thing that happens is the people keep changing their inflation expectations in a way that does not exactly correspond to rational expectations. Suppose further that the government reacts to changing inflation expectations the way equation (15.2.12) and Figure 15.2.3 say it will: if expected inflation rises, then the government pursues higher inflation, but less than one-for-one, and vice versa. This means that when expected inflation $\mathbb{E}(\pi)$ rises, actual inflation π rises but *unexpected* inflation ϵ falls. If we go back to the expectation-augmented Phillips Curve (15.2.7) or its simplified version (15.2.10), this means that output and employment will be *lower*. If we look at data generated by this economy, it will look like there's a Phillips Curve in the opposite direction than usual!

[5]See Fitzgerald and Nicolini (2014) and McLeay and Tenreyro (2019) for a discussion of this point.

What's going on? If they expect high inflation, sticky-price producers raise their prices. The Central Bank then faces an unpleasant tradeoff. If it wants to maintain low overall inflation, it must persuade flexible price producers to *lower* their prices to balance out the increases from sticky price producers. The only way to do so is to engineer a recession in order to lower marginal costs. Instead, if the Central Bank wants to avoid a recession it must tolerate higher inflation. According to (15.2.12), the Central Bank chooses to compromise and tolerates both higher inflation and lower output and employment. This outcome is sometimes known as "stagflation" (for "stagnation" plus "inflation"). Conversely, if inflation expectations fall, the Central Bank will take advantage of this to obtain both higher output and lower inflation. Arguably, this pattern is a plausible description of what went on from the 1970s until the mid-1980s.

15.4 The Liquidity Trap

In Section 15.2 we looked at how monetary policy can be used to offset negative shocks to the economy. Lowering the interest rate (or, more precisely, increasing the money supply so that the LM curve shifts, leading to a lower interest rate) produces a movement *along* the IS curve, which can offset the effect of negative shocks on GDP. Now we'll see that this type of policy has some limits.

The Zero Lower Bound on Nominal Interest Rates

Figure 15.4.1 shows how the LM curve shifts as the money supply increases. As we know, a higher money supply shifts the LM curve down and to the right: if there is more money around, people will only hold it if either they need to carry out more transactions or the opportunity cost falls. However, the LM curve never goes below $i = 0$. Why not?

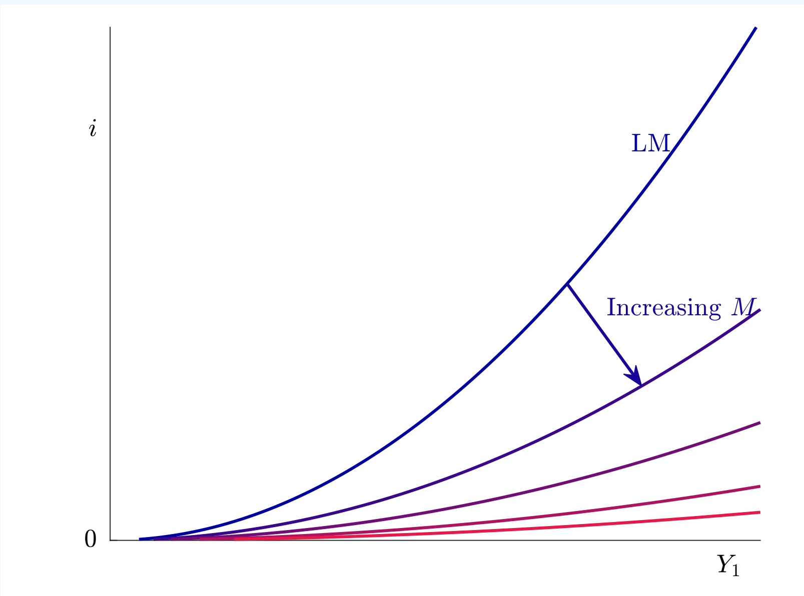


Fig. 15.4.1: *The LM curve and the zero-lower-bound on interest rates.*

Remember, the LM curve is just the representation of the money-market equilibrium condition. If $i = 0$, there is no opportunity cost of holding money: other assets like bonds also pay zero interest. If interest rates reach this point, then further increases in the money supply cannot lower the interest rate any further: people are perfectly willing to hold more and more money instead of other assets. In other words, since money always pays zero interest, it cannot be the case that other assets pay negative interest rates because people would just hold money instead.[6]

Limits on Monetary Policy

Suppose now that an economy suffers a large negative shock, as shown in Figure 15.4.2. This is like the shock we looked at in Section 15.2, just larger. In fact, the negative shock is so large that even bringing the interest rate all the way down to zero with very expansionary monetary policy is not enough to restore output to its previous level. This situation is known as a "liquidity trap".

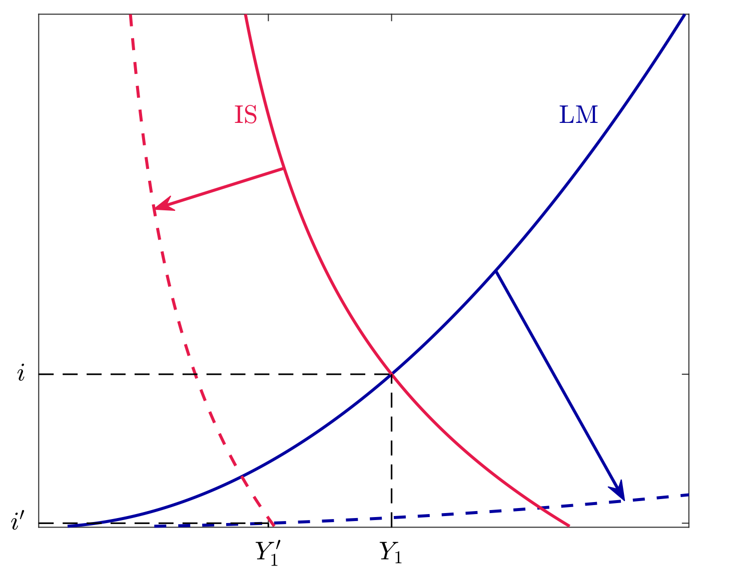


Fig. 15.4.2: *The liquidity trap.*

It's a trap in the sense that conventional monetary policy has no power to help the economy escape. It is sometimes said that expanding the money supply in a liquidity trap is like "pushing on a string".[7]

[6]This argument has been tested recently. Some countries like Switzerland and Sweden have had negative nominal interest rates. It turns out that the theoretical argument that once the interest rate becomes negative people would hold all their wealth as physical cash in a safe deposit box in order to earn zero interest is not exactly right. Storing physical currency has its own disadvantages: it can get stolen or lost, safe deposit boxes are costly and, unlike bank deposits, physical cash cannot be used to make online payments. Still, it is believed to be unlikely that interest rates could be very negative for very long.

[7]This metaphor is often attributed to Keynes, but it's unclear whether he is the original source.

Fiscal Policy in a Liquidity Trap

Many economists have argued that since monetary policy is ineffective, when an economy is in a liquidity trap it would be a good idea to use fiscal policy instead. As we saw above, an increase in G leads to a shift in the IS curve, so in principle this can be used to offset the effects of a negative shock.

In fact, as long as the economy is in a liquidity trap, there would be no "crowding out" effect from higher government spending, because the LM curve is flat at zero, so the shift in the IS curve would translate one-for-one into higher output. Part of the argument in favor of the American Recovery and Reinvestment Act of 2009 was precisely this. The economy was in a deep recession and the interest rate was already very close to zero, so there was little scope to restore usual levels of employment and GDP using monetary policy alone. Therefore, it was argued, an increase in government spending was the main tool of macroeconomic policy available. Exercise 15.6 asks you to look at the importance of timing in this type of fiscal policy.

Forward Guidance

What else can be done when an economy is in a liquidity trap? Krugman (1998) famously argued that it would be useful if the Central Bank could "credibly promise to be irresponsible". What does this mean? A Central Bank is often described as "responsible" if it is committed to pursuing low inflation without falling into the temptation to try to push output above its natural level. However, if the economy finds itself in a liquidity trap, it may actually be useful for the Central Bank to persuade the public that it will pursue *high* inflation. Let's see why that may be.

Let's go back to the beginning our our analysis. One of the things we are holding constant is expected inflation between periods 1 and 2, which we denoted π. Expected inflation matters because it determines how nominal interest rates are converted to real interest rates. We haven't really said very much about where this expected inflation comes from. Given what we saw in Chapter 11 about the relationship between money and inflation, it seems reasonable to assume that it depends, at least in part, on expectations about *future* monetary policy. Suppose that the Central Bank could make a credible announcement of what monetary policy is going to do in the future, and thereby the Central Bank could affect π. What would the Central Bank want to do?

As we saw in Figure 14.5.4, higher expected inflation results in an upward shift of the IS curve, because it lowers *real* interest rates for any given level of *nominal* interest rates. The combination of promising higher inflation with keeping nominal interest rates at zero makes the real interest rate negative. According to the model, this can succeed in raising output when the economy is in a liquidity trap. Therefore, in a liquidity trap it could be useful for the Central Bank to convince the public that it will *not* keep inflation low.

More broadly, it is increasingly recognized that communication about *future* policy is a very important aspect of how central banks do their job. Central banks are increasingly choosing to provide "forward guidance", i.e. indications of what they plan to do in the future, as a way to exert influence on the economy by changing expectations.

Exercises

15.1 **Interaction Between Fiscal and Monetary Policy**

The US government has decided to go to war in order to conquer Canada an incorporate it as the 51st US state. The Canadian government has politely agreed to carry out the war according to the following rules:

- Each country will build a large amount of tanks and set them on fire.
- The country whose tanks make the most noise will be declared the winner of the war.
- If the US wins the war, Canada will become a US state; if Canada wins the war we will leave them alone.
- Either way, no policies will change in any of the two countries, nothing besides the tanks will be destroyed and no one will be hurt.

In preparation for the "war", the US government orders a large amount of new tanks from its suppliers of military equipment.

Suppose throughout that the economy is well described by a New Keynesian model with partially sticky prices.

(a) Suppose the Federal Reserve follows a constant-money-supply policy.
   i. What will happen to GDP in the US?
   ii. What will happen to nominal interest rates?
   iii. What will happen to the price level?

(b) Suppose now that the Federal Reserve decides to adjust the money supply to keep nominal interest rates constant.
   i. What does the Federal Reserve need to do to the money supply?
   ii. How does the reaction of GDP compare to part (a)?
   iii. How does the reaction of the price level compare to part (a)?

(c) Suppose now that the Federal Reserve follows a strict inflation-targeting policy.
   i. What does the Federal Reserve need to do to the money supply?
   ii. How does the reaction of GDP compare to part (a)?
   iii. How does the reaction of nominal interest rates compare to part (a)?

15.2 **Money Demand and Fiscal Policy**

Let's look at some special cases of money-demand functions:

- Case 1:

$$m^D = a \cdot Y_0$$

- Case 2:

$$m^D = b - x \cdot i_1$$

where a, b and x are positive constants.

In case 1 the money demand depends on GDP but is completely insensitive to interest rates. Case 2 is the opposite: money demand depends on the nominal interest rate but not on GDP.

(a) Find an expression for money velocity in each of the two cases. Does the quantity theory of money hold in each case?

(b) Draw the LM curve that results from each of these two assumptions. How does the LM curve shift in response to an increase in the money supply?

(c) Suppose the government decides to undertake a fiscal expansion, i.e. increases G.

i. How does that shift the IS curve?

ii. What is the effect on interest rates and on GDP in case 1 and case 2 respectively?

iii. Suppose the fiscal expansion was undertaken with the objective of increasing GDP. You are trying to judge whether the policy is working, but you don't have the latest GDP figures yet. You do, however, have very good data on interest rates. How could you use the IS-LM model together with data on interest rates to get a sense of the effectiveness of the policy?

15.3 **Taxes and Spending in the Old Keynesian Model**

Suppose household consumption behavior is well described by the Old Keynesian model, with a consumption function:

$$c_1 = a + b(Y_1 - \tau_1)$$

where a and b are constants, $b < 1$ and τ_1 is the level of taxes collected by the government, so that $Y_1 - \tau_1$ is the household's after-tax income. The level of government spending is G.

(a) Use equation (15.1.3) and the implicit function theorem to compute how much the IS curve shifts to the right in response to an increase in government spending G and in response to a decrease in taxes τ_1, i.e. compute $\frac{\partial Y_1}{\partial G}$ and $\frac{\partial Y_1}{\partial \tau_1}$.

(b) If this model is correct, will period-1 GDP increase more in response to an increase in government spending of 100 million dollars or a tax cut of 100 million dollars? Explain.

15.4 **Patriotic Consumption**

The President has just tweeted:

> We need to help our economy. I'm going to ask that each of you go out and buy stuff. If you all do that, that'll get the economy going.

Suppose the typical household reacts to the announcement as follows:

> I wasn't planning to do this, but the President seems like such a wise leader that I'm going to do my patriotic duty and buy myself a lawnmower.

(a) Describe one model according to which the overall effect of the appeal to patriotic consumption is detrimental to welfare. Describe what happens and why it's detrimental.

(b) Describe one model according to which the overall effect of the appeal to patriotic consumption has desirable consequences. Describe what happens and why its desirable.

You can make the argument using equations, graphs or words, as long as you are precise.

15.5 **Taxes and Inflation**

Suppose the economy is well described by a New Keynesian model with partially sticky prices. The government wants to bring down inflation. The Central Bank, for some reason, is unable or unwilling to change monetary policy so the government decides to try to use fiscal policy. The first idea it considers is to lower the level of government spending but it decides not to do it because this would cut into public services that are considered too important. Two other proposals are considered:

- Proposal 1: An immediate, temporary, increase in the level of taxes, done in a lump-sum way: everyone must pay an extra Δ in taxes this year.
- Proposal 2: An immediate, temporary, increase in consumption taxes: everyone must pay extra taxes in proportion to this year's consumption

Suppose that the size of the tax increase is such that the government will raise the same revenue from both plans.

(a) If the present and future level of government spending is unchanged, what should households expect about future taxes?

(b) Will either or the two policies be effective in lowering inflation?

i. If yes, what is the mechanism?

ii. If no, why not?

(c) Suppose a lot of households are borrowing-constrained, how does that affect the answer?

15.6 **The Timing of Fiscal Policy**

Equation (15.1.1) describes how to modify the New Keynesian IS equation when there is government spending in period 1. Now imagine that there is government spending in both periods: G_1 and G_2.

(a) Write down the market-clearing condition for goods in period 2.

(b) Derive an IS equation that includes both G_1 and G_2.

(c) How does a change in G_2 move the IS curve? Interpret what this means.

(d) Suppose that the government wants to use increased spending (higher G_1) to raise GDP. However, it chooses some spending projects that will take some time to get started, so it ends up increasing G_2 instead, leaving G_1 unchanged. What will happen?

15.7 **Government Spending in the Utility Function**

Suppose that government spending enters the utility function (because households care about public goods) and consider two possible utility functions:

$$u(c_t, G_t, l_t) = u(c_t) + w(G_t) + v(l_t) \tag{15.4.1}$$

$$u(c_t, G_t, l_t) = u(c_t + G_t) + v(l_t) \tag{15.4.2}$$

(a) Explain in words what each of these two utility functions mean.

(b) What would the Euler equation for intertemporal choice look like under each of these utility functions?

(c) Derive an IS equation in each case.

(d) Assume prices are sticky. How effective are increases in G_t in increasing output in each of the two cases? Explain why.

15.8 **ATMs**

Assume that the economy is well described by the New Keynesian model with partially sticky prices. As we did in Exercise 10.4 suppose that one day, suddenly and unexpectedly, ATMs are invented, which make getting cash more convenient than before.

(a) If the Central Bank follows a policy of keeping the money supply constant, what will happen to GDP, interest rates and prices?

(b) Describe in detail what the Central Bank should do if it wanted to prevent GDP from changing in response to this change.

15.9 **The Bond Market**

Suppose that the Central Bank is known to use monetary policy to try to stabilize GDP. Suppose that a new employment report is released, which suggests that employment growth is lower than people thought it was going to be. What will happen to the price of government bonds? Explain.

15.10 **Reputation**

In this exercise we are going to think about how the government can build a reputation. The government and producers play the game we looked at in Section 15.2, with two differences:

- Instead of just playing the game once, they are going to play the same game twice.
- Producers don't actually know the value of ϕ (i.e. how much the government dislikes inflation). They are going to try to figure it out by watching what the government does the first time they play the game.

Let's start with the first time they play the game. Producers don't know the value of ϕ but they have a guess, which we denote by $\hat{\phi}$. Assume that producers believe that the government's preferences are given by $\hat{\phi}$ and that the government will play the first game just like the single-game case.

(a) What level of inflation do they expect to see? Call this level of inflation π^E.

(b) Now assume that (i) ϕ is not actually equal to $\hat{\phi}$ (it could be higher or lower) and (ii) the government does indeed behave in the first game just like in the single-game example. What level of inflation does the government actually choose?

(c) Now suppose that in preparation for the second game, producers try to figure out the true value of ϕ by looking at what the level of inflation turned out to be. Use your result from part (b) to derive an expression for how the new guess about about ϕ (let's call this ϕ^*) depends on the level of inflation. Explain in words how ϕ^* depends on π and why.

(d) Now let's consider the second game. If producers' guess about ϕ is ϕ^*, what level of inflation do they expect?

(e) Find an expression for W (the level of welfare the government attains in the second game) as a function of ϕ and ϕ^*. Does the government want producers to think that ϕ is high or low? Why?

(f) Now go back to thinking about the first game. Suppose that in the first game, the government is not just trying to maximize W (as we assumed in part (b)), but is also trying to affect how producers will form their beliefs ϕ^*.

i. Will they choose higher or lower inflation than what you found in part (b)? Why?

ii. Will GDP be higher or lower than what you found in part (b)?

Bibliography

Daron Acemoglu, Simon Johnson, and James A Robinson. The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401, 2001.

Alberto F. Alesina, Edward L. Glaeser, and Bruce Sacerdote. Work and leisure in the US and Europe: Why so different? In *NBER Macroeconomics Annual 2005, Volume 20*, pages 1–100. MIT Press, 2006.

Daniel Andrei, William Mann, and Nathalie Moyen. Why did the q theory of investment start working? *Journal of Financial Economics*, 133(2):251–272, 2019.

Irena Asmundson and Ceyda Oner. Back to basics - what is money? Without it, modern economies could not function. *Finance and Development-English Edition*, 49(3):52, 2012.

Gadi Barlevy. The cost of business cycles and the benefits of stabilization. *Economic Perspectives - Federal Reserve Bank of Chicago*, 29(1):32, 2005.

Regis Barnichon. Building a composite help-wanted index. *Economics Letters*, 109(3):175–178, 2010.

Robert J. Barro and David B. Gordon. Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1):101–121, 1983.

Robert J. Barro and Robert G. King. Time-separable preferences and intertemporal-substitution models of business cycles. *Quarterly Journal of Economics*, 99(4):817–839, 1984.

Robert J. Barro and Xavier Sala-i-Martin. Convergence across States and Regions. *Brookings Papers on Economic Activity*, 22(1):107–182, 1991.

Susanto Basu, John G. Fernald, and Miles S. Kimball. Are technology improvements contractionary? *American Economic Review*, 96(5):1418–1448, 2006.

William J. Baumol. The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics*, 66(4):545–556, 1952.

Alexander Bick, Nicola Fuchs-Schündeln, and David Lagakos. How do hours worked vary with income? cross-country evidence and implications. 2018.

Olivier Blanchard. The Economic Future of Europe. *Journal of Economic Perspectives*, 18(4):3–26, 2004.

Olivier Blanchard and Jordi Gali. Real wage rigidities and the New Keynesian model. *Journal of Money, Credit and Banking*, 39(s1):35–65, 2007.

Nicholas Bloom, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. Does management matter? Evidence from India. *Quarterly Journal of Economics*, 128(1):1 – 51, 2013.

Jutta Bolt, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden. Rebasing 'Maddison': New income comparisons and the shape of long-run economic development. *Maddison Project Database*, 2018.

Arthur F. Burns and Wesley C. Mitchell. *Measuring Business Cycles*. NBER, 1946.

Colin Camerer, Linda Babcok, George Loewenstein, and Richard Thaler. Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics*, 112(2):407–441, 1997.

James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. For better or for worse: Default effects and 401 (k) savings behavior. In *Perspectives on the Economics of Aging*, pages 81–126. University of Chicago Press, 2004.

Peter A. Diamond. Wage determination and efficiency in search equilibrium. *Review of Economic Studies*, 49 (2):217–227, 1982.

Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–82, 2015.

John Fernald. A quarterly, utilization-adjusted series on total factor productivity. 2014. FRBSF Working Paper 2012-19.

Terry J. Fitzgerald and Juan Pablo Nicolini. Is there a stable relationship between unemployment and future inflation?: Evidence from us cities. 2014. Federal Reserve Bank of Minneapolis Working Paper 713.

Milton Friedman. *The Permanent Income Hypothesis: A Theory of the Consumption Function*. Princeton University Press, 1957.

Milton Friedman and Anna J. Schwartz. *A Monetary History of the United States: 1867-1960*. Princeton University Press, 1963.

Irwin Friend and Marshall E. Blume. The demand for risky assets. *American Economic Review*, 65(5):900–922, 1975.

Jordi Gali. Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review*, 89(1):249–271, 1999.

Robert E. Hall and Charles I. Jones. Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1):83–116, 1999.

James D. Hamilton. Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics*, 100(5):831–843, 2018.

John R. Hicks. Mr. Keynes and the "classics"; a suggested interpretation. *Econometrica*, 5(2):147–159, 1937.

Robert J. Hodrick and Edward C. Prescott. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16, 1997.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

David S. Johnson, Jonathan A. Parker, and Nicholas S. Souleles. Household expenditure and the income tax rebates of 2001. *American Economic Review*, 96(5):1589–1610, 2006.

Louis Johnston and Samuel H. Williamson. *What Was the U.S. GDP Then?* MeasuringWorth, 2019.

Charles I. Jones and Peter J. Klenow. Beyond GDP? Welfare across countries and time. *American Economic Review*, 106(9):2426–57, 2016.

Nicholas Kaldor. A model of economic growth. *The Economic Journal*, pages 591–624, 1957.

John Maynard Keynes. *The General Theory of Interest, Employment and Money.* London: Macmillan, 1936.

Paul R. Krugman. It's Baaack: Japan's Slump and the Return of the Liquidity Trap. *Brookings Papers on Economic Activity*, 29(2):137–206, 1998.

Pablo Kurlat. Deposit spreads and the welfare cost of inflation. *Journal of Monetary Economics*, 106:78–93, 2019.

Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6): 1345–1370, 1982.

Robert E. Lucas, Jr. *Models of Business Cycles.* Basil Blackwell, Oxford, 1987.

Robert E. Lucas, Jr. On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1): 3–42, 1988.

Robert E. Lucas, Jr. Why doesn't capital flow from rich to poor countries? *American Economic Review*, 80 (2):92–96, 1990.

Robert E. Lucas, Jr. and Juan Pablo Nicolini. On the stability of money demand. *Journal of Monetary Economics*, 73:48–65, 2015.

Angus Maddison. *The World Economy: A Millenial Perspective.* OECD, 2001.

Michael McLeay and Silvana Tenreyro. Optimal inflation and the identification of the phillips curve. In *NBER Macroeconomics Annual 2019, volume 34.* University of Chicago Press, 2019.

Dale T. Mortensen. The matching process as a noncooperative bargaining game. In *The Economics of Information and Uncertainty*, pages 233–258. University of Chicago Press, 1982.

Jonathan A. Parker, Nicholas S. Souleles, David S. Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, 103(6):2530–53, 2013.

A. W. Phillips. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861 - 1957. *Economica*, 25(100):283–299, 1958.

Christopher A. Pissarides. Short-run equilibrium dynamics of unemployment vacancies, and real wages. *American Economic Review*, 75(4):676–690, 1985.

Raymond P. Powell. Economic growth in the USSR. *Scientific American*, 219:17–23, 1968.

Edward C. Prescott. Theory ahead of business cycle measurement. *Federal Reserve Bank of Minneapolis Quarterly Review*, (4):9–22, 1986.

Edward C. Prescott. Why do americans work so much more than europeans? *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1):2–13, 2004.

Valerie A. Ramey. Can government purchases stimulate the economy? *Journal of Economic Literature*, 49 (3):673–685, 2011.

Valerie A. Ramey and Neville Francis. A Century of Work and Leisure. *American Economic Journal: Macroeconomics*, 1(2):189–224, 2009.

John Rawls. *A Theory of Justice*. Harvard University Press, 1971.

Kenneth Rogoff. The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics*, 100(4):1169–1189, 1985.

Jeffrey D. Sachs. Tropical underdevelopment. 2001. NBER Working Paper 8119.

Thomas J Sargent. The ends of four big inflations. In *Inflation: Causes and effects*, pages 41–98. University of Chicago Press, 1982.

Thomas J. Sargent and François R. Velde. *The Big Problem of Small Change*. Princeton University Press, 2014.

Robert Shimer. The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49, 2005.

Gene Smiley. Recent unemployment rate estimates for the 1920s and 1930s. *Journal of Economic History*, 43 (2):487–493, 1983.

Robert M. Solow. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1): 65–94, 1956.

Lawrence H. Summers. Some Skeptical Observations on Real Business Cycle Theory. *Federal Reserve Bank of Minneapolis Quarterly Review*, 10(4):23–27, 1986.

George G. Szpiro. Measuring risk aversion: an alternative approach. *Review of Economics and Statistics*, 68 (1):156–159, 1986.

John B. Taylor. Discretion versus policy rules in practice. In *Carnegie-Rochester Conference Series on Public Policy*, volume 39, pages 195–214, 1993.

Richard H. Thaler and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Yale University Press, 2008.

James Tobin. The interest-elasticity of transactions demand for cash. *Review of Economics and Statistics*, 38 (3):241–247, 1956.

World Bank. *Doing Business 2012.* 2011.

Kurlat has filled an essential void by producing a text, accessible to undergraduates with a moderate amount of mathematical training, that introduces students to the frontier of modern macroeconomics. I will be teaching from it this year and recommend it to instructors at the intermediate or masters level and to mathematically-inclined students who want to learn what macroeconomics is all about.

Gabriel Chodorow-Reich, Harvard University

This outstanding book covers modern macroeconomic ideas with extreme rigor but without heavy math and keeping the focus on real-world applications and policy implications. Readers will find a very accessible coverage of microeconomic foundations and a thoughtful treatment of long-run and short-run macroeconomic models. Every instructor who teaches undergraduate macroeconomics at an intermediate or advanced level should consider using this book.

Alp Simsek, Massachusetts Institute of Technology

A fantastic introduction to macroeconomics for advanced undergraduates. I use it as background reading for my masters level course at the LSE on growth as well. What makes the book stand out is a very accessible, indeed, a sparkling conversational style combined with analytical rigour and a masterly overview of the literature. It is rare to find a textbook that is so lucidly written and yet analytically so solid.

Maitreesh Ghatak, London School of Economics

Pablo Kurlat is Associate Professor of Economics at the University of Southern California. He got his Ph.D. in Economics from MIT in 2010.

His research focuses on information economics, monetary policy and financial markets.

This book grew out of his teaching notes for the Intermediate Macroeconomics course he used to teach at Stanford University.