An Introduction to the Dismal Science of Economics for Those Concerned with more Noble Pursuits

Miles Corak

November 2010

Graduate School of Public and International Affairs

University of Ottawa

55 Laurier Avenue East

Ottawa ON K1N 6N5 Canada

mcorak@uottawa.ca

Contents

Preface and objectives

STUDENTS will read this book with different objectives, but the most common has to do with the need to develop a reasonably good familiarity with the basic ideas of economics that can be used to engage in discussions of public policy, or simply to intelligently read a newspaper. The basic principles of economics are discussed, illustrated and explained in a way that will appeal to students who have no previous experience in the subject, who recognize a personal need to learn at least its the basic vocabulary and grammar, and who have an interest in public policy issues both domestic and international.

The content falls out of several years of teaching to graduate students of public and international affairs who for a variety of reasons never completed an economics course during their undergraduate education, or if they did it was some time ago, and an experience—for whatever reason—not to be repeated. It is hard to generalize on the reasons why motivated, engaged, and interesting people of the sort that I have taught have essentially no formal training in a subject that so directly touches the lives of all citizens, in all countries; rich or poor, north or south. Part of this may have to do with the type of learning and teaching styles that have come to dominate the subject, an analytical treatment, possibly orientated to mathematics and without reliance on history, politics, philosophy and biography that offer alternative ways of entering into some of the basic principles and tools of economic analysis.

No single method is used here, but the basic lessons of economics—both microeconomics and macroeconomics—are approached from a variety of angles in a manner that will hopefully be appealing.

The material is organized into three blocks. The first outlines the scope and method of economic analysis. It certainly falls short of offering a definition of the subject, something that has eluded even the greatest thinkers and is likely at any rate to change through time, but seeks through example to clarify some basic tools and methods: the value of theoretical models and reasoning; the important questions that have motivated the great thinkers; and the distinction between positive and normative analysis. It also introduces some of the basic and core concepts—scarcity, opportunity costs, marginal analysis—and uses them to highlight public policy issues like population growth and free trade. Adam Smith, David Ricardo and Thomas Malthus are all discussed in the context of their times, the important public policy issues they tried to illu-

minate and understand, and why what they had to say some 200 years ago still resonates in public policy discussions today.

The second block treats microeconomic theory, focusing on competitive markets and the appropriate tools to analyse them. Here Alfred Marshall reigns, and his famous contribution to the theory of value—demand curves, supply curves, and their interaction through the idea of equilibrium—is developed and applied to agricultural commodities, oil, specific taxes, and the labour market. This falls short of many important aspects of the theory of value, but the door is left open through a discussion of consumer and producer surpluses and market inefficiencies.

The final block addresses macroeconomic theory, beginning with John Maynard Keynes's fundamental contribution to our understanding of economic depressions and high unemployment rates. It describes the national accounting system in Canada from a historical perspective, outlining its development through an interaction between the theoretical advances of Keynes, the pressing needs of public policy, and developments in the technology of measurement. Inflation and public policy are addressed through the writings of Milton Friedman, and the public policy pronouncements of one of Canada's most successful central bankers.

The importance of recognizing that economics is a growing and developing— indeed unfinished—science is stressed throughout. This growth comes from the interaction of pressing public policy concerns, theoretical understanding, and the capacity to view the world through numbers. It is always supported by the thoughts of the great thinkers who came before, and whose contribution is lasting.

[Add and discuss the following quote from Hayek.

A certain humility to approaching economics and its use in public policy is cerainly in order among many members of the discipline and is in general an important trait in the social sciences regardless of the discipline. Hayek is speaking to economists when he says:

> Nobody can be a great economist who is only an economist—and I am even tempted to add that the economists who is only an economist is likely to become a nuisance if not a postitive danger [Studies in Philophophy, Politics, and Economics, 1967]

But philosopher, political scientist, statistician could all just as equally be substituted into this phrase.]

Chapter 1

Introduction

I work for a rather large university, a university with thousands of students, in the capital city of Canada, in fact within sight of the heart of political life. I work in a rather large building, some 12 stories housing tens and tens of classrooms, offices, lecture halls and other rooms necessary for the life of students, administrators, professors, and support staff. I have an office on the eleventh floor overlooking the eastern part of the city and in view of busy streets that intersect the campus and structure the steady flow of students between the various buildings, both university and private, that they need to enter and leave.

And like them, and like many others, at various points in the day I need a coffee. A cappuccino; that is what I prefer. On the ground floor of my building there is indeed a coffee shop, a Starbucks. But I have other choices. I can leave the building and walk down a busy thoroughfare a little more than a block to another coffee shop called Second Cup. Or I can walk slightly further still outside the formal confines of the university buildings to a third alternative, Timothy's World Coffee.

You can see the layout on the map labeled as Figure 1.1. My office building is represented by the large circle, and the three coffee shops by the smaller circles. The distances are not really that far, though admittedly they seem— and perhaps in a certain sense are—a lot further in the heart of winter when the temperatures in Ottawa fall to points that contribute to its status as one of the world's coldest capital cities, and certainly they seem further when it rains and I don't have my umbrella.

The cappuccino forms an important part of my day, every day. I usually have one at about 10:30, though this can vary: sometimes as late as 11:00, but never after lunch. Only once a day, but every day. The coffee shops in the area of the university do a good trade. They are often quite busy when I am there, and usually through the entire day and into the evenings, particularly the evenings closer to examination time. There is money to be made on Laurier Avenue East selling coffee: not just because there are a good number of people continually moving through the area, and not just because the type of people that are in want or need are likely more prevalent among the rushed and hurried

Figure 1.1: Distances from my office to the nearby coffee shops
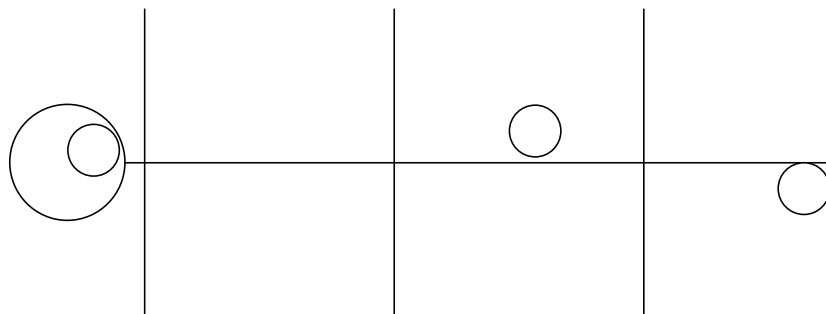
population of the university, but also because these coffee houses have learned how to price their goods to maximum effect.

I am interested only in cappuccini, and only in small cappuccini, ignoring all the other tempting items on the menus hanging from the cafe walls. So you would think that after at least a couple of years of steady visits to these various stores—and I don't find myself going to any particular establishment more than the others, not being concerned with the colour of my cup, the appeal of the logo, the ambiance, the music or any other aspects of the decor by which the various establishments distinguish themselves—so you would think that I would know the price of a cappuccino. And you would think I would know it exactly because I am economist. Wouldn't you?

It is striking that I don't. But in a sense I am probably not that different, in spite of my professional training, from many of the others who regularly visit one of these three places for their morning pick-me-up.

One day near the beginning of a new school year, after having been away for the summer and before returning to my regular routine, I decided to take a survey, a price survey of my favourite drink. After all it was easy to do. All of the three places sell essentially the same drink, of the same quality, the same size, and delivered with roughly the same efficiency. All the prices are posted and readily viewed by all. You can see the results of my survey in what I have labeled as Table 1.1.

The prices are almost the same, but surprisingly not exactly the same. After all we are talking about essentially the same thing, and with a price difference of as much as 10 cents it is clear that some of the establishments over the course of a year must be taking in a good deal more than others from all the cappuccini they sell. Why is that? Or more generally what determines the price in the first place? Why all around $3.00?

I suspect that most people can offer reasonable guesses, there is after all a good deal of mileage to be had from common sense, from the stuff we just know

Table 1.1: The price of a small cappuccino near my office

Establishment	Price in dollars	Distance from my office
Starbucks	$3.00	in the same building
Second Cup	$2.95	1 1/2 blocks away
Timothy's World Coffee	$2.90	2 1/2 blocks away

though not certain where or how it was acquired. "Location, location, location," as the old adage goes, probably has something important to say. Most people probably have a good deal of economic common sense, and probably don't appreciate how well they know the subject called "economic theory." And a question of this sort is certainly a fundamental part of economic theory: what determines the price of something?

My fixation on a cappuccino might give the appearance of trivializing this question. But upon reflection there is a sense in which it is very important, determining how we lead our lives, the satisfaction we get from the things and people around us, and indeed reflecting the type of society we live in and pass on to the next generation. What determines how much we earn for the hours of work we put in, or the gains we will receive from the hours of study we devote, or the value we place on the air, water and natural environment around us. And if these questions are about prices then it is not far to other questions like: why do some in the world have so much, and others so little?

Or for that matter why in the rich countries are the disparities in incomes and capabilities so wide. Right next door to the Second Cup where I regularly enjoy a 3 ounce drink of black and nutritionally useless liquid for the before tax price of $2.95 stands a Roman Catholic church, in front of which are regularly lined up dozens of men waiting, waiting, waiting to be admitted to a small kitchen for a free meal. Jobless out of choice? Or jobless for lack of opportunity? They may be unemployed, or they may be unemployable, but they are clearly hungry.

As relevant as these questions are, and hopefully we will be able to offer intelligent answers to them all by the end of this book, I would like to start our discussions about economic theory and public policy with another question: how many babies should we have?

Chapter 2

The method and scope of economics

1. Positive versus Normative statements. 2. Economic methods and models. 3. The two great questions: how much surplus society creates, how the surplus is distributed. 4. A definition of the subject.

ON JULY 25th, 2008 the British newspaper *The Guardian* published an article called: "Doctors' advice to Britons: have fewer children and help save the planet."[1] The article summarized the major messages from an editorial in the *British Medical Journal*, a journal that is owned by the British Medical Association, and has as its vision "to be the world's most influential and widely read medical journal."

Its web site goes on to explain that the mission of the journal is to "lead the debate on health, and to engage, inform, and stimulate doctors, researchers and other health professionals in ways that will improve outcomes for patients."[2] The web site also explains that the circulation for its print edition is close to 200,000 copies every four weeks. This is clearly an influential publication, which seeks to be actively engaged in and to significantly influence public policy debate.

The article in *The Guardian* summarized an editorial by John Guillebaud a professor of family planning at University College London and his co-author Pip Hayes, a General Practitioner. In an interview with the newspaper Dr. Guillebaud stated that "a really green thing to do is to have one child less." He goes on to explain that "an opportunity is missed when a doctor in this country [the United Kingdom] is talking to a young couple .... We think doctors might join in a gentle campaign of suggesting people have one child less."[3]

---

[1] Ian Sample (2008). "Doctors' advice to Britons: have fewer children and help save the plant." *The Guardian.* Friday, July 25th, page 3.

[2] See http://resources.bmj.com/bmj/about-bmj, accessed on September 12, 2008.

[3] See the audio clip of an interview at http://www.guardian.co.uk/world/audio/2008/

We should not have any doubt that this is a serious public policy proposal being put forward by a serious observer. The proposals are meant to change public policy: they are explicitly intended to inform and influence the medical profession, to get doctors to treat their patients differently, and to change the way people behave. The stated goal of all of this, as the headline makes clear, is to save the planet from environmental degradation.

2.1   Positive versus Normative statements

How can we respond to the suggestions that couples should be having fewer children? One letter writer to the *The Guardian* expressed his views this way:

> In one of the greatest lines ever written for a great show, Homer Simpson comes to realize the magnitude of his despair as a father. "I have three children and no money," he wails. "Why can't I have no children and three money?" I watched this in the company of my three children. We all laughed like drains. Those who take a professional interest in the welfare of children may have noted that I was perhaps laughing a little harder than my children, or that it is a line I tend to repeat to them frequently, especially when we wait in line for cinema tickets and they ask if they can have popcorn at sickeningly marked-up prices.
>
> Anyway, an editorial in the British Medical Journal tells us that to save the planet, we should limit ourselves to at most two children. God, am I tired of doctors telling us what to do. First they tell us not to smoke. Then they they tell us not to drink a bottle of wine a night. And now they're telling us how many times men can impregnate their partners.
>
> Actually, I am trying to write this while my children cavort around me. One of them is playing the ukelele and asking me, every 12 seconds, how much longer I am going to be. The second child is making helpful suggestions as to which of our three children should be culled in order to oblige the BMJ. And the third – the one chosen by the second for euthanasia on the grounds he is "most useless" – is going crazy with a football. There are times when you do see the cold-hearted logic behind the BMJ's pronouncement.
>
> Well, not really. As long as you're not trying to write, having three children is great. I used to be a committed no-children man, but circumstances changed, and I find myself a convert to the larger family. Not too large. Four children or more are too many. But two are not enough: all pairs of children do is bicker. Three argue in many more exciting and unusual combinations. And as for their

jul/25/john.guillebaud.fewer.children , accessed on September 12, 2008. The reference to the original article is: John Guillebaud and Pip Hayes (2008). "Population growth and climate change." *British Medical Journal.* Volume 337, No. a576. August 2nd, pages 247-48.

> ruining the planet − pah ... it's not population that's the problem,
> it's how the population lives.
>
> But doctors need to be sat on. Their job is to try and cure
> us when we get ill − nothing else. If we do not exercise ceaseless
> vigilance, it will not be long before they end up recommending that
> we eat each other to save resources. And, unlike Swift's Modest
> Proposal, or Soylent Green, they won't be being satirical.[4]

In effect: who do doctors think they are telling us how many children to have?
That is a legitimate way to proceed.

Doctors have a technical and scientific expertise in health and disease, and
therefore the letter writer wonders on what basis they have the authority to
pronounce on a basic value and desire such as the number of children to have,
and on its relationship to climate change. The "right" number of children is
something couples should decide on their own, reflecting their values, beliefs
and circumstances. So-called experts sometimes rely on their technical or pro-
fessional credentials to promote their own values or political agenda. We should
be wary of this pattern: when the sailors, who have the technical knowledge to
sail a ship and keep it afloat, feel their knowledge gives them the authority to
set the direction.

In fact, the *British Medical Journal* makes a point of having the authors
of its editorials, which are distinct from the other scientific and peer reviewed
articles in the journal, disclose any potential conflicts of interest. Aside from
this editorial writers seem to be given license to express their views. In this
particular case, Dr. Guillebaud "has received fees and expenses from manu-
factures of contraceptives for educational presentations, research projects, and
short term consultancies, and is a patron of the Optimum Population Trust.
[Dr. Hayes] is a Trustee of the Optimum Population Trust." So just maybe
these authors are pursuing a pre-conceived agenda, and using the popular con-
cern about environmental degradation to give their beliefs more traction in the
public arena.

At this level it is hard to resolve the argument. Someone believes that a
family of three is a good thing; someone else believes it is not. Maybe the
doctors are in a conflict of interest; maybe they are not.

But there is another line of attack in the writer's argument. If we take the
doctors on their word and assume they are disinterested and objective observers,
is their logic correct? Does it make sense that lower fertility rates in the United
Kingdom will save the planet? This is a very different tack, and is a direction
that at least in principle can lead to agreement and consensus.

The editorial page writers of *The Guardian* also voiced an opinion on the
Guillebaud-Hayes articles. In addition to questioning the objectivity of the
authors they focus on their logic:

> The preoccupations of an age are often given away by its choice
> of prophet. In the 90s, Karl Marx came back into vogue, not as

---

[4] Nicholas Lezard (2008). "Now doctors are telling families to have only two kids. As a
father of three, I know what's good for us." *The Guardian*. Saturday, July 26th, page 34.

the John the Baptist of the class struggle, but as a reliable guide to globalisation and its discontents. Old Whiskers was even the subject of a long New Yorker essay, which argued that Wall Street types had nothing to lose by reading him. Over the last couple of years, it has been Thomas Malthus's turn in the spotlight. The spectre of "Pop" Malthus, as students referred to his work on population growth, has hovered over the recent arguments about record food and fuel prices. His warning about how growing populations would outstrip food supply are often echoed by greens and on blogs. And today the British Medical Journal weighs in, with an online opinion piece that is essentially Malthus-lite.

The problem that the BMJ authors and others highlight is real; the solution they give, however, is plain wrong. True, a shortage of food is only heightened by a rising population; the same goes for tackling climate change. The statistic one often hears from the population control lobby is that the world will have 9.2bn people by 2050. Someone born in 1950 who lives to be 100 will see the global population grow well over threefold. That may sound scary, but it does not prove a direct causal link between commodity supply or greenhouse gases, and population growth. The proper link is between consumption and commodities, and between emissions and climate change. Anything else is a side issue.

The UK has around 60 million people; but the average British citizen creates nearly 10 times more carbon dioxide emissions than the average Indian, and 166 times more than the typical Ethiopian. So the best way to deal with climate change is not for Ethiopia to curb its (runaway) population growth, but for the British and others in the west to reduce their greenhouse gas emissions. Since most of the population in 2050 will be in the poorest countries with the smallest carbon footprints, birth control will do little on global warming. For the British to limit themselves to two children per family, as the BMJ piece argues, is besides the point: birth rates in the UK are just 1.7 children per couple. The authors of the BMJ piece are associates of the Optimum Population Trust. That group believes the UK should only have 17 million people; which 17 million it does not say.

Population control has a terrible reputation: India's forced sterilisation programme was among the blackest points in its recent history. Just as there is a reason why prophets come back into fashion, so there is normally a reason why history turned its back on them. In Malthus' case, he was simply wrong.[5]

It is not population, but consumption patterns that matter: not the number of people but the amount of greenhouse gas produced per person. The author of

---

[5] Editorial page (2008). "Population control: An old misconception." *The Guardian*. Friday, July 25th, page 38.

the editorial argues that the population of the UK is already below replacement rates, and the concern of public policy should not be on fertility rates but on "greening" consumption patterns.

On this basis we can have an argument about the proposal that potentially will garner a consensus and give appropriate guidance for the conduct of public policy. Is the logic right? What predictions does it generate? Can these be tested, and do the actual numbers bring them into question? And even if the logic stands up, can the proposed policy be implemented in a way that actually improves things, rather than making them worse or entailing other unintended and costly consequences that might outweigh any possible benefits?

The difference between these two approaches—the emotional father versus the calculating newspaper journalist—illustrates an important distinction: that between a "normative" argument and a "positive" argument. Economic analysis is used to understand societies, and to develop policies to influence and change how they function. Positive economics refers to the scientific part of this: the science—the body of systematized knowledge—concerning what is; normative economics refers to the prescriptive part: the recommendations concerning what ought to be. We can criticize the proposal to have doctors encourage people to have fewer babies because it offends our values, or we can criticize it because it is not logical, cannot be effectively implemented, and even if implemented will not address the issue at hand.

Positive economics describes economic events and situations, and develops theories to explain them. It strives to follow scientific principles: developing models to represent reality, generating hypotheses or predictions from these models, and testing these predictions with actual observations to determine if they are refuted or not. It is, or attempts to be, objective.

Normative economics reflects a value judgment on what should be done. While a normative statement might be built upon some objective view of reality, it requires a further leap based upon something other than the facts. It is, explicitly or implicitly, ethical.

Both are a necessary part of public policy discourse. Policy makers need to understand the world as it is, how it works, and the effectiveness of the instruments available for changing it. But policy makers also need to have a vision of the ideal, of the social goals they value and wish to achieve.

Even so it is easy to confuse the two. After all economics is concerned with issues that directly touch the lives of everyone: whether or not someone will have a job, how much they will earn, how their children will succeed in life, and—as we have just seen—how many children they have in the first place. These issues are part of the everyday media, and a good part of what keeps governments busy. It is not surprising that most everyone has, and indeed should have, an opinion on matters of this sort.

It is also not surprising that there is no shortage of "experts" who are only too willing to express their opinions and advice. Self-proclaimed experts will be much more numerous on matters of economics than they are on matters of cataracts or sub-atomic particles, because the subject matter of economics is so directly relevant to how we manage ourselves individually and collectively. It is

much more likely that both citizens in the street and highly trained professionals will shape their objective analysis of a problem or issue to suit their ethical vision of the world.

Someone who has, for example, spent his whole career studying family planning, teaching it, and making educational presentations on the effectiveness of contraceptives to the point of being an acknowledged expert and even being paid by companies for his opinions and analyses, may be much more inclined to view solutions to society's problems, real problems that everyone agrees need addressing, through the lens of population control.

Even for observers who profess to be scientific there is a risk of going beyond the logic of an objective analysis, of applying its reasoning out of context, or of weighing the alternative public policies and the trade-offs they imply in a particular fashion, one that reflects pre-conceived views of the way the world "should" be. And how much easier it must also be for those with no real expertise but strong pre-conceived views to shape objective conclusions to fit those views. In fact, it may well be that the values of *The Guardian* editorial writer have also informed his or her way of addressing the issue: a writer for a centre-left newspaper, quick to point out the terrible extremes of a policy gone wrong a generation ago in India without regard to other more measured and respectful implementations.

So it is not surprising that "normative" and "positive" statements are often confused. After all a normative statement, a policy recommendation, must rest on some sort of analysis of the problem; it cannot be independent of positive economics. The flip side of this is, at least in principle, not the case. Milton Friedman, for example, has written:

> Positive economics is in principle independent of any particular ethical position or normative judgments. As Keynes says, it deals with "what is," not with "what ought to be." Its task is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances. Its performance is to be judged by the precision, scope, and conformity with experience of the predictions it yields. In short, positive economics is, or can be, an "objective" science. Of course, the fact that economics deals with the interrelations of human beings, and that the investigator is himself part of the subject matter being investigated in a more intimate sense than in the physical sciences, raises special difficulties in achieving objectivity....[6]

---

[6]Milton Friedman (1953). "The Methodology of Positive Economics." In *Essays in Positive Economics*. Chicago: University of Chicago Press. page 4. In this passage the author is referring to a statement by John Neville Keynes (not John Maynard Keynes) in which he distinguishes between "... a *positive science* ... [as] a body of systematized knowledge concerning what is; a *normative ... science* as a body of systematized knowledge relating to criteria of what ought to be, and concerned therefore with the ideal as distinguished from the actual." John Neville Keynes (1955). *The Scope and Method of Political Economy*. Fourth edition. New York: Kelley and MacMillan, pages 34-5.

This distinction between positive and normative flows out of a philosophical heritage based in logical positivism. No doubt there are many who believe that it is impossible to make such a distinction.

"Post-modernists," for example, believe that ....

Friedman, in fact, acknowledges this in the later part of this quotation, as have others. There are after all no economic facts without theories. "Facts do not organize themselves," writes Gunnar Myrdal, "into concepts and theories just by being looked at; indeed, except within the framework of concepts and theories, there are no scientific facts but only chaos. There is an inescapable *a priori* element in all scientific work."[7] This issue certainly needs to be acknowledged, but without accepting that it stymies all attempts at objective inquiry. A science progresses through the interplay between developments in theory, observations designed to test the hypotheses or predictions it implies, and changes in both our power to conduct measurement and the public policy relevance of the undertaking.

Our first lesson in economics is to understand the distinction between positive and normative statements, to appreciate that this distinction can at times be challenging to make, and to develop a familiarity with the scientific approach that disciplines our analysis of public policy. An important aspect of this approach is the development and testing of "models."

2.2 Economic methods and models

So how many babies should we have? The editorial writer of *The Guardian* suggests that the claim that more babies will lead to environmental degradation, and the policy recommendation that people in the UK should all have one child less, rests on the outdated and incorrect theory put forward by Thomas Malthus. This is in reference to an economist who authored a book published in 1803 called, in all its glory: *An Essay on the Principle of Population; or, A View of its Past and Present Effects on Human Happiness; with an inquiry into our prospects respecting the future removal or mitigation of the evils which it occasions.*[8]

As far as I can tell the difference in views between the editorial writer and the authors of the BMJ article can be summarized by equation (2.1).

$$\text{greenhouse gas emissions} = \text{number of people} \times \frac{\text{greenhouse gas emissions}}{\text{person}}$$

$$(2.1)$$

---

[7] Gunnar Myrdal (1954). *The Political Element in the Development of Economic Theory.* Cambridge, Massachusetts: Harvard University Press, page vii.

[8] T. R. Malthus (1803). *An Essay on the Principle of Population.* Selected and introduced by Donald Winch. Cambridge: Cambridge University Press. This edition uses the text of the 1803 edition but shows additions and corrections made in the 1806, 1807, 1817, and 1826 editions. The book was popular in its day, engendering wide ranging discussions in the Britain of the early 1800s, but its ideas–as our discussion makes clear in a small way–still resonate today. In fact, the book was first published in 1798, but the 1803 edition is definitive.

The total amount of greenhouse gas emissions produced over a generation is the product of the number of people and the average amount of emissions per person: more people for fixed emissions per person means more greenhouse gases; more emissions per person for a fixed number of people also means more greenhouse gases.

In a way this helps us to clarify the nature of the arguments put forward: Doctors Guillebaud and Hayes assume that greenhouse gas emissions per person will stay constant for the next generation, and even if the UK population is below replacement it should fall further in order to reduce total emissions to make enough of a difference. The response by the writer of *The Guardian* editorial points out that the real issue is with the emissions per person. This is what has to be reduced, rather than assumed constant, to have an impact on total emissions.

This taxonomy helps to clarify the issues. A taxonomy is simply a listing, a categorization: a mental filing cabinet. It helps in organizing our thoughts and the sometimes many issues that may be at play. But it doesn't seem to prove or disprove either case. Possibly an appeal to the facts might help: by how much would the population have to fall in order for it to be conceivable that the reduction in UK emissions would have a global impact? In a similar vein by how much would emissions per person have to fall? Would either of these changes be feasible? Interestingly neither party has much to say in this regard.

A taxonomy, valuable as it may be, is not all that makes up a theory. *The Guardian* writer must have more in mind if he or she is correct in asserting that Malthus was "simply wrong." Equation (2.1) is a mathematical relationship, a tautology—and as such is true by definition. Taxonomy is part of theory, but it is not all. A theory is a set of assumptions about technical and social relationships that generates predictions about the issues of interest not yet observed and that, at least in principle, can be refuted by data.

Before exploring what Malthus actually said it is interesting to ask in what sense he could be wrong. Or in other words how do we evaluate, as positive economists, whether a theory is valid or not? We could judge a theory in two ways: in terms of the logic, or in terms of the facts. A theory could be logically inconsistent. It could be wrong if it contradicts itself, or if the conclusions do not follow from the assumptions. It is important therefore to check the logical consistency and the completeness of the theory. Formal logic and clear abstract thinking are essential to this.

But even if the model has a clear logic it can be wrong because the predictions that follow may not be verified when they are confronted with the facts. Understanding the hypotheses that are generated by the theory and gathering the appropriate facts to test them is challenging, particularly for a science that cannot—like many of the physical sciences—as easily rely on controlled experiments. But this is the ultimate and most important aspect of the scientific method. Malthus could be wrong because the predictions from his theory don't find support in the facts, or to be strictly correct are disproved by the facts. This seems to be the premise of the editorial in *The Guardian.*

Tests of this sort have to be constructed with care because the model must

be applied to situations for which it is intended. How to do this is not simple, clear, or a matter of abstract reasoning, but rather is more of an art. Since models are not universal, there is a good deal of intuition or judgment needed in applying them. The world is always changing, and a model appropriate for a particular time and place may not fit in another context. "It is at this point," Milton Friedman wrote, "that the 'amateur' is separated from the 'professional' in all sciences and that the thin line is drawn which distinguishes the 'crackpot' from the scientist."[9] In other words, Malthus may only appear to be wrong because we are applying his model inappropriately to situations for which it was not intended.

In addition there is a particular and important way in which we should *not* evaluate a theoretical model: whether the assumptions upon which it is based are realistic. In fact, Friedman asserts, "the more significant the theory, the more unrealistic the assumptions" (Friedman 1953, page 14). He is not saying that unrealistic assumptions guarantee a significant theory. But rather that a theory is significant if it explains a lot in a simple way. By abstracting from the details, and by focusing just on the factors relevant for determining the matter of interest a theory must by necessity be false in its assumptions. The issue is not whether the assumptions are realistic, but whether they are good enough for the job at hand.

The map I offered in Figure 1.1 to aid in the explanation of coffee prices according to the distance from my office is obviously not a very realistic description of the geography in the area of my office. In Figure 2.1 I offer another alternative, one intended by the university, and put on its web site, to serve as a multi-purpose tool for almost anyone interested in finding their way from any university building to any other building. As such it is certainly more realistic than the map I drew, every building and street are listed, and so are parking spaces and bus stops, as well as all kinds of other information.

As good as this map is for the purpose of getting from any building to any other building on the university campus, and while it can also serve the purpose I intended—to show the distance from my building to the coffee shops—you might agree that it is too busy, too detailed for my purposes. The map offered in Figure 1.1 is better. Even if it is a patently unrealistic description of reality, and the nature of the walk between my office and the three coffee shops from which I can buy my coffee, it much more clearly illustrates the distances that are central to the determination of cappuccino prices, and in this simplicity it is a better model.

We should judge, in other words, a theoretical model not by the realism of its assumptions, but by the validity of its predictions. My model is based on the idea that the price of coffee declines with the distance from my office building, and in particular at a rate of about 5 cents for every block to block and a half. It predicts, for example, that if another coffee shop were to open between the Starbucks and the Second Cup a small cappuccino would sell for somewhere

---

[9]Milton Friedman (1953). "The Methodology of Positive Economics." In *Essays in Positive Economics*. Chicago: University of Chicago Press, page 25.

Figure 2.1: Campus of the University of Ottawa

between $3.00 and $2.95.

The ideal test of this model would be to run a controlled experiment by introducing a coffee shop exactly between the Starbucks and the Second Cup, hold everything else the same, and to observe the price the owners post for a cappuccino. Though this method is a standard in some of the physical sciences and in medicine—for example to test the efficacy and safety of newly proposed drugs—it is less often, though sometimes, used in the social sciences.[10] So we are left to two alternatives: wait and see what happens if at some point in the future a new coffee shop happens to open up; or go back to the data to uncover the possible existence of coffee shops of which we were not aware. Neither of these will be ideal, and our data rarely offers a perfect test of our hypotheses. But this is the nature of an imperfect science.

In this particular case, however, it does happen that there is a fourth coffee shop where I could have bought my cappuccini. It is called Cafe Nostalgica, a student run cafe located not far from the Second Cup, but closer to the heart of the campus. Its location is presented as the extra small circle I have added to the updated version of Figure 1.1 called Figure 2.2. As the crow flies it is only one block away from my office, and arguably between the Second Cup and the Starbucks.

This may not be a perfect test of the prediction of my theory: the new cafe is located off the main street closer to other population centres in the campus, and since it is student run it may have different pricing policies than a profit maximizing firm or different degrees of willingness to pay from its more loyal customer base. But when I asked the barista for the price of a small cappuccino

---

[10] A "controlled experiment" refers to the method used in many of the sciences to determine if an intervention has a causal impact on an outcome. For example, in the medical sciences there is a need to determine if a new drug cures a disease, the proposed drug being the "intervention," the outcome being the state of health of those who take the drug, and the "causal" impact being the extent to which the drug leads to improved health. Those proposing a cure have the challenge of determining the causal impact of the drug. They might administer the drug to a group of individuals and observe what happens to their health over time, but this would not be terribly convincing. If health improved it might be due to the drug, or it might be due to other things in the environment that had also changed. Because all of the other possible influences on health cannot be observed and taken to account there is a need to also observe the health of a "control" group, a group of individuals who are similar to those receiving the drug in all other respects, the only difference being that they have not been administered the drug.

A "controlled experiment" involves the random assignment of the intervention to those enrolled in the experiment to create two groups of individuals: a "treated" group, and a "control" group. The difference in the outcome between these two groups represents the causal impact of the intervention. This is because all other changes in the outcome brought about by other influences would have the same impact on both groups, and would therefore be differenced away. The classic example of this method is the experiment run on hundreds of thousands of school-aged children in the United States during the 1950s to test the effectiveness of the Salk vaccine as a cure for polio. This experiment, along with the challenges of convincingly implementing controlled experiments, is described by David Freedman, Robert Pisani, Roger Purves (2007), *Statistics*, Fourth Edit ion, New York: W.W. Norton, pages 3-10. In their description they make reference to Francis Thomas Jr. (1955). "An evaluation of the 1954 poliomyelitis vaccine trials: summary report." *American Journal of Public Health*. Vol. 65, pages 1-63. In fact, economics is increasingly making use of such methods to test the predictions of theory.

Figure 2.2: Updated locations of nearby coffee shops
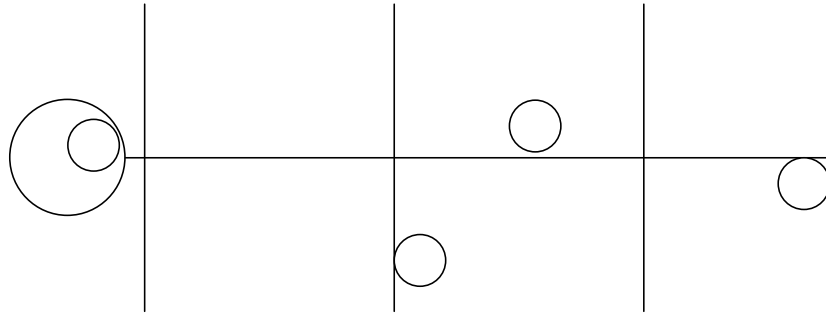
she said, I was pleasantly surprised to hear, $2.98. This does not prove my theory for the reasons discussed above—there may be other things going on here that are not being taken into account—but proving a theory is too high a standard. What is important is that it does not disprove the theory.

2.3 The two great questions: how much surplus society creates, how the surplus is distributed

Well, just how many babies should we have? What did Malthus say? What predictions did his model imply? Was he wrong?

In developing his theory Malthus, the positive economist, makes two assumptions: one behavioural and the other technical.

Based on observations from a number of historical experiences in different countries and calculations made by other theorists, he finds that human populations double in as few as 10 years to as many as 25 years or so. Malthus is beginning with the facts that are available to him. He takes the most conservative estimate as his starting point. The behavioural assumption is that the human population, like other species in the plant and animal kingdoms, has a natural tendency to increase geometrically—as the numbers 1, 2, 4, 8, 16, 32, 64, 128, 256 as he puts it—when its environment provides for an above subsistence amount food. This assumption, therefore, adopts a rather long period of time as the unit of analysis, a generation, a period measured not in months or years, but in decades.

The technical assumption concerns the productivity of land, and again this is based upon the empirical evidence available to him. "If it be allowed, that by the best possible policy, and great encouragements to agriculture, the average produce of [Europe] could be doubled in the first twenty-five years, it will be allowing probably a greater increase than could with reason be expected. In

the next twenty-five years, it is impossible to suppose that the produce could be quadrupled. It would be contrary to all our knowledge of the properties of land." (Malthus, page 18.) He is assuming that technology does not change at a very fast rate, if at all.

The third element of the model is an equilibrium condition. The geometric increase in human populations cannot possibly be matched by increases in production of the food and resources to sustain it, and as a result human populations are propelled to levels that oscillate around a steady state of misery. The population eventually outstrips the carrying capacity of the environment, and is kept in check by these physical limits. The equilibrating mechanism, that keeps the population in check means, as Malthus—whose views led many to call economics the "dismal science"—puts it

> ... the human race will be constantly endeavouring to increase beyond the means of subsistence. But as by that law of our nature which makes food necessary to the life of man, populations can never actually increase beyond the lowest nourishment capable of supporting it; a strong check on population, from the difficulty of acquiring food, must be constantly in operation. This difficulty must fall somewhere, and must necessarily be severely felt in some or other of the various forms of misery, or the fear of misery, by a large portion of mankind. (page 15)

> The ... checks to population are extremely various, and include every cause, whether arising from vice or misery, which in any degree contributes to shorten the natural duration of human life. Under this head therefore may be enumerated, all unwholesome occupations, severe labour and exposure to the seasons, extreme poverty, bad nursing of children, great towns, excesses of all kinds, the whole train of common diseases and epidemics, wars, pestilence, plague, and famine. (page 23)

Well, you get the picture.

This model offers an explanation for the misery that Malthus saw around him in the Britain of the late 1700s and early 1800s. It predicts that the vast majority of the population will be living on the margins of life, and it implies that this is a "supply side" problem—driven by demographics. This is all positive economics, and Malthus certainly goes on to discuss normative economics, the policy proposals derived from the model that will improve the well-being of the significant mass of the population. But as good scientists we can keep these discussions separate for now. It is interesting to examine the actual workings of the model in more detail. It is still a model that animates public policy discussion not only in the UK and other rich countries, but particularly in poorer countries.[11]

---

[11]See for example: "Africa's population: The baby bonanza." *The Economist.* August 29, 2009, Vol. 392, Issue 8646. Available at http://www.economist.com/displaystory.cfm?story_id=14302837. Also see

We can give it more precision and clarity—as well as introduce some fundamental principles of economic analysis—by coupling the model with some ideas from Malthus's close friend and intellectual rival David Ricardo, who also lived and wrote during the same period. From Ricardo we take two important ideas: the concept of scarcity, and the method of marginal analysis. With these ideas we can supplement Malthus's thinking in a way that answers the two great questions of economics that are implicit in his analysis: how much of a surplus does a society generate? and how is that surplus divided among its population? And just as importantly we will also be able to answer the equally important question posed at the beginning of this chapter: why do I pay so much for my cappuccino? and who is making all the money?

If a resource is said to be "scarce" all we mean is that the quantity of it available is not sufficient to satisfy all the demands put on it. At some population level land becomes scarce in Malthus's model. Ricardo offers the classic statement:

> On the first settling of a country, in which there is an abundance of rich and fertile land, a very small proportion of which is required to be cultivated for the support of the actual population, or indeed can be cultivated with the capital which the population can command, there will be no rent; for no one would pay for the use of the land, when there was an abundant quantity not yet appropriated, and, therefore, at the disposal of whosoever might choose to cultivate it. On the common principles of supply and demand, no rent could be paid for such land, for the reason stated why nothing is given for the use of air and water, or for any other of the gifts of nature which exist in boundless quantity.[12]

When population is relatively small land has no economic value, no price, it is free like the air we breath. It is, in other words, not scarce. But at some point our demands upon the natural resources around us increase to the point that they exceed the available supply, and at that point—and only at that point—we have an economic issue. Scarcity implies that we face trade-offs in allocating what is available to all the uses it could be put. Indeed, many people have defined the subject of economics as the study of how to allocate scarce resources among the competing demands placed upon them.

When we reach the point of scarcity our actions entail a type of cost because of the trade-offs that most be made, the use of the resource for one purpose implies that it cannot be used for other purposes. In making a particular use we forgo other possibilities. For this reason we refer to these costs as "opportunity costs." They are the value we attach to the next best alternative use of the

http://economist.com/debate/overview/151 for an online debate organized by the *The Economist* magazine on whether the world would be better off with fewer people. Accessed September 14, 2009.

[12]David Ricardo (1817). *On the Principles of Political Economy and Taxation.* Volume I of The Works and Correspondence of David Ricardo. Piero Sraffa (editor). Cambridge: Cambridge University Press. Page 69.

resource, the value of the opportunity foregone. And in this context we have an economic problem: how to allocate resources in a way that maximizes the social good. Or another way of putting it: how to maximize the surplus in society, and how to allocate that surplus across the members of society. The first is a question of production, or growth, and the second is a question of distribution.

Microeconomic reasoning concerns questions of allocation, and it leads to the realization that the maximization of something—the total production less the total costs of production (which is what we mean by "surplus")—requires marginal reasoning; maximization occurs when the marginal benefits of an action just equal the marginal costs.

The most obvious example of a resource that has an opportunity cost is our time. There are a fixed number of hours in the day, and we would like to do more things than there is time available. Time is scarce. Each minute spent doing something means that something else cannot be done. The fact that my time has an opportunity cost is ultimately what is driving the pattern of cappuccino prices in the neighbourhood of my office. In making my decision as to where to buy my cappuccino I make a calculation comparing the marginal benefit to the marginal cost and stopping when they are the same: what do I give up by taking the extra minutes to walk a block and a half? is it worth the five cents I save? given I have come this far should I take another few minutes to go another block and a half? are those extra steps worth the additional five cents I would save. I will walk up to the point that the extra benefit—the 5 cents in savings—just equals the value of the next most productive thing that I could be doing with that few minutes foregone.

If economics is defined by the substantive objective of maximizing a surplus—often in reference to material goods—then it is also defined by the method of marginal analysis. It is at the margin that the size of a society's population is determined in Malthus's model, at the point where an additional increase in the population just attains a subsistence level of existence. Anything less would leave a incremental gain that would lead to further population growth; anything more would imply an incremental deficit that would reduce it. At this point society has maximized its surplus, and ironically—or perhaps not so ironically if you appreciate that the workings of the model also imply a theory of distribution—the vast majority of the population are living in abject misery.

We can make this much clearer by borrowing from Ricardo to amplify and adapt a bit the technical assumption that Malthus made concerning the productivity of land. Let us assume that we live in a one good world. That way we don't have to worry about how to measure the surplus in a society. In a one good model the surplus is just the amount of the good that is left after paying all the costs associated with producing it. And also there is no need to worry about prices. Since there is only one good there is nothing to trade and therefore no need for a rate at which one good exchanges for another. Let's call this good "corn."

There are two resources in society: land and labour. And two classes of individuals each associated with one of the resources: landowners who earn a rent, and workers who earn a wage. Rents and wages are paid in corn. Ricardo

continues this way:

> If all land had the same properties, if it were unlimited in quantity, and uniform in quality, no charge could be made for its use, unless where it possessed peculiar advantages of situation. It is only, then, because land is not unlimited in quantity and uniform in quality, and because in the progress of population, land of an inferior quality, or less advantageously situated is called into cultivation, that rent is ever paid for the use of it. When in the progress of society, land of the second degree of fertility is taken into cultivation, rent immediately commences on that of the first quality, and the amount of that rent will depend on the difference in the quality of these two portions of land.

> When land of the third quality is taken into cultivation, rent immediately commences on the second, and it is regulated as before, by the difference in their productive powers. At the time same time, the rent of the first quality will rise, for that must always be above the rent of the second, by the difference between the produce which they yield with a given quantity of capital and labour. With every step in the progress of population, which shall oblige a country to have recourse to land of a worse quality, to enable it to raise its supply of food, rent, on all the more fertile land, will rise. (Page 70)

Tim Harford eloquently translates this into an English more attune to our ears in the first chapter of his book *The Undercover Economist*.[13] The story can be illustrated with the help of Figure 2.3, with some specific numbers attached for the sake of concreteness. The productivity of each of four different types of land is indicated by the height of the vertical bars. Type 1 land is the most productive yielding, for every labourer working this land, 100 units of corn per period. Type 2 is not as fertile yielding, 90 units of corn per worker. Type 3 is even less fertile yielding 80 units. And so on. We can measure the output of corn in whatever units are convenient, tonnes, kilograms, bushels. But it is a flow, an amount of corn per unit of time, say per year to represent the annual cycle of agriculture. The vertical heights of the bars represent the amount of corn that a single labourer would be able to produce over the course of a year on each of four types of land.

This is strictly a technical relationship, rooted in the fertility of the soil. We can go out and do tests on the land and determine its productivity. Think of a lush meadow, scrub land, and progressively more rocky terrain, each successively less suited to the growing of an agricultural crop. This technical relationship on its own does not tell us how much land will be cultivated, how much corn will

---

[13]Tim Harford (2005). *The Undercover Economist: Exposing why the rich are rich, why the poor are poor—and why you can never buy a decent car!* Anchor Canada. The analogy that I am drawing between Ricardo's theory of rents and the differences in cappuccino prices near my office is inspired by Harford's more eloquent discussion in the opening chapter of his book. In my case the Starbucks in my office building is located on land that possesses, in Ricardo's terms, "peculiar advantages of situation."

Figure 2.3: Ricardo - Malthus corn model
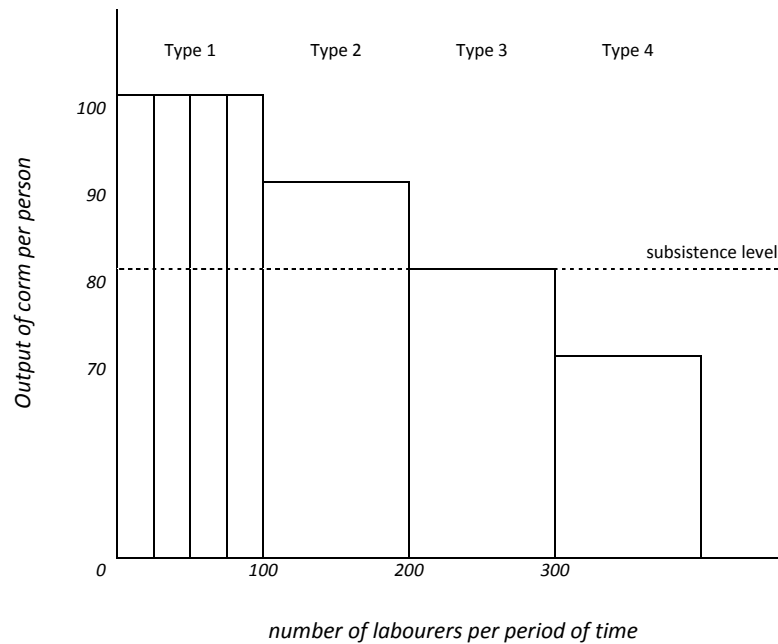
be produced, how large the population will be, nor how the corn is distributed among the population that it ends up supporting.

We need to introduce a behavioural assumption, and for this we borrow directly from Malthus. Assume that there is some subsistence level of corn needed by the workers, and for the sake of argument let's say this is 80 units of corn per period of time. This is shown in the figure as the horizontal dashed line. Anything below this amount will not be sufficient for life and the labourers die; anything above implies more than they need to live and is used to augment the population. With less than 80 units per head the population falls; with anything above it increases. Malthus assumes that any output above subsistence is used entirely to augment the population: when there is a surplus people marry sooner and have more children, and more of these children survive to adulthood to start their own families.

The third aspect of the model is the adoption of an appropriate time horizon, one that will drive this simple economy to an equilibrium of a steady state in which the pressures of population growth are just balanced by the limits of the environment. In equilibrium the model comes to rest, the amount of corn, the size of the population, and the distribution of corn among the two social classes all determined. In this model the time horizon must be long enough to permit

changes in population, something measured in decades as opposed to weeks, or years. The appropriate application of this model requires adopting this longer term frame of reference.

Equilibrium is not necessarily an optimum or ideal position. Rather the analogy in this model is to physics, it is a position of rest in which there is a balance of forces leading to no further movement in the variables determined by the system.

The horizontal axis of the figure measures the number of labourers that work the land each year. Again choose an appropriate unit: hundreds, thousands, millions. It does not matter, but again this is a flow, something measured per unit of time. We assume that all labourers are the same, offering the same effort and skill to the task of farming. Imagine that individuals live for two periods: childhood and adulthood. In adulthood each individual matches with another individual to form a household. The household raises four children into adulthood if it has access to an above subsistence amount of corn, and these children then go on in the next period to establish their own households and raise their own children. So the population increases in the way that Malthus imagined: first two then four, eight, and so on.

So over the course of decades and generations this is how the model works. Imagine the first household that settles the territory chooses to work type 1 land, since this is the most productive, by striking a deal with the landowner of the particular parcel of land upon which they have settled. This household, represented by the very first thin vertical bar in Figure 2.3 produces 100 units of corn per person per year. This implies a surplus of 20 units of corn per person after they each consume the 80 that is necessary for survival.

How much of this surplus does the household get to keep, and how much does the landlord receive as rent? We should be clear, firstly, that there is more to our little society than we have made explicit. The fact that we have "landlords" implies the idea of property rights. So in the background there is a social order being assumed, a way of enforcing and legitimizing the property rights. We need not be concerned about how this came about, but it is clear that the landowners have in some sense the legitimate right to their land. The labourers are there only after an agreement has been reached, they are not squatting on the land or doing anything unlawful. The terms of that agreement are set through a negotiation between the landowner and the worker, and determine the distribution of the surplus.

So we need to understand who has the bargaining power to set the terms of the contract in their favour, and in that way we will understand how the surplus is distributed. In economics strength comes from having options; or to put it in another way, strength comes from scarcity, from controlling the scarce factor of production. When our society is first settled land is not scarce. The demands being put on the land do not exceed its capacity. In fact, just the opposite: it is labour that is scarce. If labourers begin negotiations with a landowner both parties know that if the terms of the bargain are not to their liking the workers can just as easily negotiate with another landowner of type 1 land. Workers have options, and they are in a position to play one landowner

off against another. They pursue their self-interest to the limit by offering the landowner some minimal amount, and extracting the surplus for themselves.

As time goes on their children are more likely to survive through infancy, become adults and, since times are good and the future continues to look good, marry early raising their own children under equally favourable conditions. As the population expands more and more of type 1 land is brought into cultivation—represented by the succession of thin vertical bars in the figure. But as Ricardo describes in the previous quote, at some point the edge of type 1 land is reached and the next incremental increase in population forces the next generation onto type 2 land. This is an important turning point because the demands placed on type 1 land are exceeded by its availability, more workers would like to work on type 1 land than is available. It is now scarce, and with this scarcity the bargaining power shifts to the landowner. The landowners now have options and can extract some of the surplus. How much?

Households forced to work on type 2 land would gladly pay landowners of type 1 land a much higher rent than they have hereto received. In fact, it would be worth their while to pay anything up to the difference in the productivity of type 1 and 2 land as rent. Working type 1 land, producing 100 units of corn and paying a rent of up to 9 units of corn would still net the worker more than working type 2 land where a maximum of only 90 units can be produced. Information is perfect: landowners know this, all workers know this, and consequently the terms of the bargain between the two groups are eventually renegotiated. Landowners of type 1 land are able to extract a rent of 10 units of corn, and all workers—whether they work type 1 land or type 2 land—take home 90 units. The rent on type 1 land is the difference between its productivity and the productivity of the marginal land, in this case type 2.

Landowners of type 2 land do not, of course, receive any rent because that type of land is not scarce. But as you can imagine with time that too will also change since 90 units of corn per year is still above subsistence, and consequently population will continue to grow. At some point population growth pushes out the margin to the next type of land. Type 3 land will eventually become the rent-free, or marginal land. But it produces only 80 units of corn per year, just equal to the subsistence level. Now workers occupying this land compete with all other workers on all the previously cultivated land, types 1 and 2. Landowners on type 2 land have bargaining power because their land has become scarce. They demand some of the surplus, but landowners of type 1 have also benefited. The next best alternative of all labourers is 80 units of corn. All labourers regardless of the productivity of the land they actually work receive only a subsistence wage.

This is enough to survive and to continue to reproduce, but only enough. As population grows further the new margin upon which it can settle is even less productive, type 4 land yields a below subsistence amount of corn. The settlers that are forced onto this land will not survive, and population will decline. Over time it will oscillate and settle at the level supported by type 3 land. The economy eventually comes to an equilibrium with 300 units of labour working types 1, 2, and 3 land.

We are able to precisely answer the two great questions we have posed. The total production is given by the sum of the amount of corn produced by each individual on each type of land. We can get this by adding up the individual contributions, the thin vertical bars that we can imagine run through the entire picture. On type 1 land 100 labourers each produce 100 units of corn for a total of $100 \times 100 = 10,000$ units of corn per year. Similarly on type 2 land the total production is given by the area of the vertical bar. With each of 100 workers producing 90 units of corn this is equal to $100 \times 90 = 9,000$ units of corn per year. On type 3 land total production is $100 \times 80 = 8,000$. So each year the total production in the society is 27,000 units of corn. The total cost of producing this corn is the amount needed to keep the workers alive, with 300 workers all requiring 80 units of corn this amounts to 24,000 units of corn each year. So the surplus is $27,000 - 24,000 = 3,000$ units of corn per year. All of this surplus goes to the landlords. The economy supports a population of 300 labourers, and all of them live a life of misery receiving just enough to stay alive.

In this equilibrium society has maximized its surplus. If population was lower an incremental increase would add more to the total production than it would to total cost: if it was higher an incremental decrease would reduce the total costs more than it would reduce total production. The surplus is maximized when the marginal benefit just equals the marginal cost: when the contribution of an extra person to total production just equals the addition to total cost. Yet, the members of this society—more concretely the labourers—lead the miserable existence described by Malthus. This is the situation when the model comes to rest; it may be an equilibrium, it may represent the maximum surplus possible, but it is hardly a desirable outcome for the have-nots.

It is natural to want to change this outcome. We can as an exercise in positive economics evaluate the effectiveness of different policies intended to improve the well-being of those living in poverty. Malthus in fact did this in a careful and methodical way, evaluating a number of policies including: a redistribution of the surplus toward more equality; programs of income support; emigration; and technical changes that improve the quality of the land. His particular concerns were programs of income support, which in his day were known as the 'poor laws'. These laws required local parishes to support the poor, and Malthus sees them as causing rather than alleviating poverty. We might think of them as 'aid' of the sort rich countries give to the poor, or programs of income support for the poor living in the rich countries, what is sometimes called welfare. Malthus's logic in reaching this conclusion should be clear to us. Figure 2.4 amends Figure 2.3 by illustrating the consequences of giving those forced to live on Type 4 land sufficient aid to reach subsistence. This amounts to 100×10 units of corn—shown as the shaded area in the figure—which added to the 100×70 units of corn produced per person is just enough to give each of these 100 labourers the 80 units need to survive. These people would not have been able to survive solely on the fruits of their labour, but now do and the economy supports a population of 400 rather than 300.

What happens next is best left in Malthus's own words, who, speaking of the poor laws, says that their "first obvious tendency is to increase population

without increasing the food for its support. A poor man may marry with little or no prospect of being able to support a family without parish assistance. They may be said, therefore, to create the poor which they maintain ...".[14] The policy, in other words, induces a behavioural response: it causes the population to grow further, and creating another generation who are forced to live on even a crueler margin. We have not even said anything about where this aid comes from, but each attempt to help each successive generation of poor will require more aid. Ultimately this policy will not be sustainable, and end up doing no good, in fact just the opposite. A much greater population will be supported, on ever more marginal land, implying more environmental degradation without making any of the labourers better off.[15]

The only way to help the poor in this model is to put limits on population growth. Malthus would prefer that this happen voluntarily, through moral restraint and delayed marriage rather than through the vice and misery that would follow from the laws of nature. It is in this sense that the authors of the article in the *British Medical Journal* are in fact Malthus-lite, their policy recommendation being essentially the same. Malthus is right in the sense that his model has a logical consistency, but clearly the prediction of his model—that the only way to improve material well-being is to restrict population growth—is wrong. The vast majority of the population in the England of 2001 was astoundingly better off than in the England of 1801, yet the population increased from less than 10 million to almost 50 million over these 200 years.[16] The prediction of the model are refuted by the facts, and this should lead us to question whether the assumptions upon which it is built are appropriate for our times. This is what the editorial writers of *The Guardian* are trying to say.

This is a model that relies on two assumptions. The first is an assumption about human behaviour and the associated time horizon. The surplus in society is used only to support more people; it is consumed, never invested. This happens over decades and generations. The second is an assumption that the productivity of the land is subject to the law of diminishing marginal productivity, and that its productivity does not change over time. It is ironic that a model adopting a very long time horizon, one long enough for population to change, also assumes that technical know-how and productivity don't change. This may have been appropriate for Malthus's times, for someone who built his theory by

---

[14]Malthus (1803), page 100.

[15]In a statement that is surely among the first and most blunt claims that incentives matter, that policy makers must be aware that their actions may have unintended consequences, and that certainly labels Malthus as the first 'supply side' economist, he says: "Hard as it may appear in individual instances, dependent poverty ought to be held disgraceful. Such a stimulus seems to be absolutely necessary to promote the happiness of the greatest mass of mankind; and every general attempt to weaken this stimulus, however benevolent its apparent intention, will always defeat its own purpose. If men be induced to marry from the mere prospect of parish provision, they are not only unjustly tempted to bring unhappiness and dependence upon themselves and children, but they are tempted, without knowing it, to injure all in the same class as themselves." (page 101)

[16]For these population estimates see Figure 1.1 in Julie Jefferies (2005), "The UK population: past, present and future," in *Focus on People and Migration 2005*. United Kingdom: Office of National Statistics.

Figure 2.4: Evaluating a policy of aid using the Ricardo - Malthus corn model
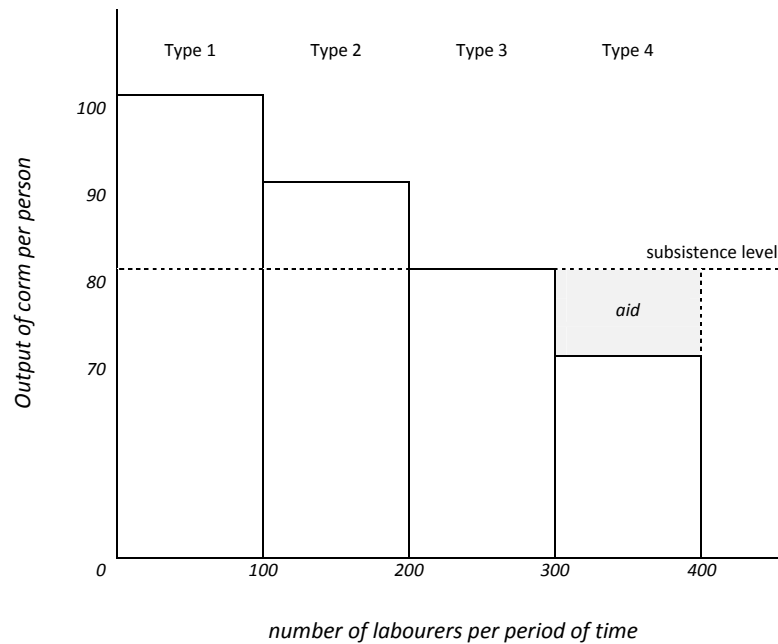
observing the facts of hundreds of years of feudal history in which technology in fact did not change fast enough to keep pace with population changes. But just barely appropriate. Malthus did not foresee the changes in technology that the industrial revolution was beginning to bring during his life, that would increase exponentially in the coming decades and continue into our times.

It is not simply that we need to assess once and for all if the model is "right" or "wrong." The real challenge is judging the appropriateness of a model to address the problem at hand. While we might legitimately question the use of this model by Drs. Guilleband and Hayes in the context of a growing and rich country to support what would appear to be pre-conceived normative views, this is not to say that it cannot be legitimately and correctly used in other contexts either to highlight issues or concerns that must be addressed, or outright to generate predictions and policy recommendations. For example, one of the strongest advocates for increased foreign aid to Africa, Jeffrey Sachs, has even said:

> I have been asked dozens of times if help for Africa would ultimately backfire in an even greater population explosion. Would greater child survival rates not translate directly into more adult hunger and suffering? Usually the questioner begins sheepishly, apologizing

> for what is to come. Then he or she explains that they do not want
> to sound callous, but that they really need to understand the issue.
> This is a very fair question. After all, Thomas Malthus asked almost
> the same one two hundred years ago.[17]

To make his case for increased aid Sachs needs to point out how the model does
not fit the current reality, or illustrate what modifications would be necessary
so that aid would have a long-lasting positive effect. In this sense the model is
useful, helping to sharpen and clarify arguments.

2.4   A definition of the subject

Our discussion of this model is meant as motivation: to appreciate the impor-
tance of the kinds of questions economics has historically addressed, to under-
stand the elements of economic methods, and to begin to understand some of
the underlying core concepts central to this way of thinking. As an example it
has offered a definition of economics, as an example it has demonstrated how
economics is done, and along the way it has illustrated the use of some of the
core principles.

But if you were to take a look at some of the more popular introductory
economics textbooks you might be hard pressed to suggest that there is a con-
sensus concerning the definition of the subject.[18] Certainly a quick scan of
a few textbooks written by particularly influential contemporary economists
does not suggest that these experts are using exactly the same definition when
speaking to those studying the subject for the first time. In a textbook he
has co-authored, Ben Bernanke, the current chair of the Federal Reserve Bank
in the United States, states that economics is "the study of how people make
choices under conditions of scarcity and of the results of those choices for so-
ciety."[19]A former Chairman of the President's Council of Economic Advisors,
Gregory Mankiw, is slightly more general in stating that "[e]conomics is the
study of how society manages its scarce resources."[20] While Paul Krugman, a
recent Nobel Prize laureate and *New York Times* columnist, and his co-authors
begin their book with a chapter called "Introduction: The Ordinary Business
of Life" in which they offer an even more general definition of economics as
"the study of economies, at the level both of individuals and of society as a
whole." They also distinguish two branches: "[m]icroeconomics is the branch
of economics that studies how people make decisions and how these decisions

---

[17]Jeffrey D. Sachs (2005). *The End of Poverty: Economic Possibilities for Our Times.*
London: Penguin Books, page 323.

[18]This and the following three paragraphs are informed by and draw from Roger E. Back-
house and Steven G. Medema (2009). "Retrospectives: On the Definition of Economics."
Journal of Economic Perspectives. Volume 23, Number 1 (Winter), pages 221-34.

[19]Robert H. Frank, Ben S. Bernanke, Lars Osberg, Melvin L. Cross, Brian K. MacLean
(2005). *Principles of Microeconomics.* Second Canadian Edition. Toronto: McGraw-Hill
Ryerson, page 3.

[20]N. Gregory Mankiw, Ronald D. Kneebone, Kenneth J. McKenzie, Nicholas Rowe (2008).
Principles of Microeconomics. Fourth Canadian Edition. Nelson Education, page 4.

interact;" while "[m]acroeconomics is the branch of economics that is concerned with overall ups and downs in the economy."[21]

The exposition of Malthus and Ricardo in the previous section is certainly intended to illustrate the usefulness of economic theory and models in discussing and evaluating important public policy issues, but it is also meant to highlight the crucial elements of what economics is, both in its subject matter and in its method. Classical economists, who were the contemporaries and precursors of Malthus and Ricardo, would have been comfortable defining the subject in terms of its substance, in terms of the issues addressed. To them economics deals with material well-being, and seeks to understand what determines a nation's wealth, what determines how national wealth grows, and how it is distributed. Indeed, Adam Smith entitled his book—which was published in 1776 and which many associate with the birth of economics—*The Wealth of Nations.*[22] I have tried to emphasize this heritage by focusing on the two great questions: how society creates a surplus? how the surplus is distributed among its members? Indeed, many non-economists today would also be comfortable with a definition of this sort, that is, economics as the study of material well-being.

But many people would also have the notion that economics has something to do with markets, with money, and at its root with exchanges of goods and services between individuals mediated by markets and money. In fact by the late 1800s and early 1900s the most current definition of economics was very much in line with a view that also incorporated individual behaviour as a core element of the subject. In one of the first and most influential textbooks on the subject the Cambridge UK economist Alfred Marshall wrote that economics

> is a study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of material requisites of wellbeing.
>
> Thus it is on the one side a study of wealth; and on the other, and more important side, a part of the study of man.[23]

While it is, in other words, the study of material well-being, it is also the study of human motivations and behaviour. As such Marshall's definition highlights a subject in transition. Defining economics solely in terms of the issues it addresses does not capture the essence of the subject as it is practiced today. In fact, Steven Levitt in writing to a broader, non-professional audience about the major findings of his applied economic research felt compelled to address this by entitling his book "Freakonomics," in large part because the subject matter—which ranged from the reasons for declines in rates of criminality to what makes

---

[21] Paul Krugman, Robin Wells, Martha L. Olney (2007). *Essentials of Economics.* New York: Worth Publishers, pages 3 and 4.

[22] Adam Smith (1776). *An Inquiry into the Nature and Causes of The Wealth of Nations.* New York: Modern Library Edition

[23] Alfred Marshall (1920). *Principles of Economics: An introductory volume.* Eighth Edition. London: MacMillan and Co., page 1.

a 'perfect' parent—was likely not in line with what most of his potential readers thought economics was about.[24]

All this is to say that since the time Marshall wrote economics has increasingly become defined not just in terms of the substantive issues it addresses, but in terms of the methods used. In fact, as the quotes from the Bernanke and Mankiw textbooks suggest, oftentimes it is not defined at all in terms of issues, but entirely in terms of methods. Our exposition of Malthus and Ricardo was also meant to highlight these methods and the core concepts upon which they are based. The concept of scarcity was central to all of this. In fact, an influential and often cited definition of economics, one that is at the root of contemporary textbook definitions, considers economics as "the science which studies human behavior as a relationship between ends and scarce means which have alternative uses."[25] The gradual adoption of this narrowing of the subject in terms of method or approach has had the effect of expanding the scope of acceptable subjects that are addressed by economists. The claim that "economics is what economists do" has more truth to it than when it was originally made some 70 or more years ago.[26]

Scarcity is at the centre of the economic approach to human behaviour. It is very important that we clearly understand this idea as well as the other core concepts that follow from it—opportunity costs and marginal reasoning—as well as the pitfalls that can arise when they are not fully appreciated, fully understood, or for that matter when they are misapplied. Some very important public policies affecting the lives of millions of people are, at their root, supported by these concepts. And it makes sense to examine them and their implications more explicitly.

---

[24]Steven D. Levitt and Stephen J. Dubner (2005). *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything.* New York: William Morrow.

[25]Lionel Robbins (1935 [1945]). *An Essay on the Nature and Significance of Economic Science.* Second Edition, Revised and Extended. London: MacMillan and Company. Page 16.

[26]This quote has long been ascribed to Jacob Viner, though apparently never appearing in his writings. See Gary S. Becker (1976), "The Economic Approach to Human Behavior," in Gary S. Becker, *The Economic Approach to Human Behavior*, Chicago: The University of Chicago Press, page 4, and Backhouse and Medema (2009, page 222). The latter also point out that the "move to economics as the analysis of choice had the effect of pushing to side questions of philosophy and ethics, history and institutions, broader conceptions of rationality, and various nonmathematical approaches to the subject" (page 230).

Questions to consider

1. Economics is sometimes defined not in terms of its subject matter, but rather in terms of its method: "Economics ... is not a body of concrete truth, but an engine for the discovery of concrete truth." If economics is a method or an approach, explain what that method is.

2. Define each of the following terms, explain the significance of the relationship of each pair:

   (a) Comparative advantage and Absolute advantage

   (b) Firm and Factor market

   (c) Microeconomics and Macroeconomics

   (d) Endogenous variable and Exogenous variable

   (e) Positive economics and Normative economics

   (f) Opportunity cost and Sunk cost

   (g) Scarcity and Rent

   (h) Marginal productivity and Average productivity

   (i) Inequality and Poverty

   (j) Efficiency and Equality

3. Answer the following questions by re-considering Figure 2.3.

   (a) Explain what each of the curves in the figure is meant to represent, and explain why they display the patterns they do. As a part of your discussion be certain to explain which variables are endogenous and which are exogenous.

   (b) Use economic reasoning to explain how the model reaches an equilibrium. What is the total production of corn in equilibrium? How is that total output distributed?

   (c) Complete the following table, and explain at what point the total surplus is maximized and why.

Number of Workers	Total Production	Total Cost	Surplus
100			
200			
300			
400			

   (d) Examine the implications (predictions) of the model in response to each of the following policies:

       i. A program of taxes and transfers in which some of the surplus is transferred from landlords to workers

    ii. A change in technology that increases the productivity of the land

   iii. Emigration

   iv. A medical intervention that permanently reduces infant mortality rates so that more children survive to adulthood

    v. A government imposed and fully enforced regulation that permits families to have only one child

      In all cases be certain to show how you would change the diagram, and to fully discuss the intuition of how the new equilibrium is reached.

4. Use the following template to draw the marginal benefit and marginal cost curves that determine how much time I spend walking for a cappuccino per day. Marginal benefits and marginal costs are measured along the vertical axis (or as it is often referred to, the Y axis) in cents per minute walked, and the minutes spent walking per day are measured on the horizontal (or X) axis. Do this using the information in Table 1.1, and your understanding of the opportunity costs of my time. Show how the curves—and the resulting equilibrium—-would change if it was raining.

cents per minute

0

number of minutes spent walking for a cappuccino per day
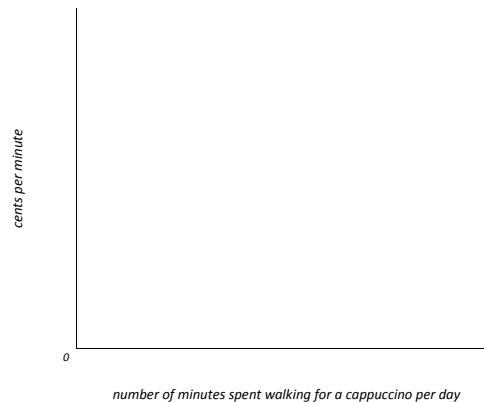
5. Is it rational for some people to drop out of high school, for others to end their formal education with high school degrees, and others still to obtain a university degree? Develop an economic model using marginal reasoning to answer these questions. In doing so explain how and why the marginal benefit and marginal costs vary with the number of years spent in school. In your answer be certain to account for both the direct costs of an extra year of education—for example like tuition fees— but also the opportunity costs. Also explain what parameters (that is, exogenous variables) are being held constant, and how they would change the curves. Derive some testable predictions from your model by hypothesizing a change in the underlying parameters.

6. Consider the document "Prosperity, inequality, and poverty" published by the Institute for Competitiveness and Prosperity, and listed in the further readings. In what ways does this document address the same questions that motivated the discussion in this chapter? How is surplus measured in a contemporary economy; how is inequality measured? According to this document who are the poor in contemporary Canadian society, and what are the policy recommendations put forward to improve their circumstances? In what way are these recommendations similar to those Malthus offered two hundred years ago? How are these similar or contrast with the discussion of inequality by Joseph Heath in chapters 11 and 12 of his book *Filthy Lucre*, which is also listed in the further readings?

Further readings

The intuition of scarcity and marginal reasoning, as well as an engaging and easy to read presentation of Ricardo's theory of rent, is offered in the first chapter of Harford (2005). The second chapter also discusses the idea of a surplus, but from the point of view of a consumer in a rich country facing the pricing policies of firms keen on extracting as much of it for themselves.

The Institute for Competitiveness and Prosperity (2007) offers an illustration of the two great questions but placed in a modern advanced country and using data and statistics. It also documents inequality and poverty in Canada, offering a portrait of those living on the margins. The policy recommendations to help them are based on supply side prescriptions.

One of these is to increase the education levels of the least advantaged. Question 4 asks about the optimal level of education for an individual, and offers you the opportunity to sharpen your skills at thinking marginally. Reading the indicated sections of the book *Human Capital* by Becker will help you.

Saez and Veall (2005), and the book by Heath (2009) document the extent of inequality in the United States and Canada, and offer an interpretation and evaluation of some public policy debates. They help answer the question of just how the surplus is distributed in North America, and what if anything should be done. On page 10 of the introduction to his book Heath says that the "first half [of the book] can be thought of as "favorite economic fallacies of the right"—arguments routinely peddled by conservatives not because those arguments make any sense, but because the conclusions are so agreeable to them. The second half consists of "favorite economic fallacies of the left." Chapters 11 and 12 deal with the left's perspective on poverty and inequality, but as you can imagine the entire book is a careful and accessibly written attempt to use positive analysis to uncover the misconceptions underlying less scientific thinking.

The definition of economics is discussed in Backhouse and Medema (2009), which is the principle source for the discussion in this chapter. They offer a nice, but brief, overview of the history of thought on this question, with ample references to support your answer to Question 1. Becker's famous 1976 essay—

which carries the same title as the book within which it appears—is a clear, if extreme statement, of the methods of economics. But the classic treatment is the book by Robbins (1935). In spite of being written in the 1930s it discusses the nature of economics and its significance in a way that would probably find a good deal of support among academic economists today.

1. Tim Harford (2005). *The Undercover Economist: Exposing why the rich are rich, why the poor are poor—and why you can never buy a decent car!* Toronto: Anchor Canada, Random House Canada. Chapters 1 and 2.

2. Institute for Competitiveness and Prosperity (2007). *Prosperity, inequality, and poverty.* Working paper 10. Available at http://www. competeprosper.ca/index.php/work/working_papers/ working_paper_prosperity_inequality_and_poverty/

3. Gary S. Becker (1993). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* Third Edition. Chicago: University of Chicago Press. Pages 108 to 130. Available at http://www.nber.org/books/beck94-1

4. Emmanuel Saez and Michael R. Veall (2005). "The evolution of high incomes in North America: Lessons from the Canadian evidence." *American Economic Review.* Vol. 95 number 3, pages 831-49.

5. Joseph Heath (2009). *Filthy Lucre: Economics for People Who Hate Capitalism.* Toronto: Harper Collins. Chapters 11 and 12.

6. Roger E. Backhouse and Steven G. Medema (2009). "On the Definition of Economics." *Journal of Economic Perspectives.* Volume 23, Number 1 (Winter), pages 221-33.

7. Gary S. Becker (1976), "The Economic Approach to Human Behavior," in Gary S. Becker, *The Economic Approach to Human Behavior*, Chicago: The University of Chicago Press. Chapter 1.

8. Lionel Robbins (1935 [1945]). *An Essay on the Nature and Significance of Economic Science.* Second Edition, Revised and Extended. London: MacMillan and Company. Chapters 1, 2, 3 and 6.

Chapter 3

Scarcity and its implications

1. Scarcity and choice. 2. The labour market and the production function. 3. A two-good model and the gains from trade. 4. The gains from trade and the distribution of the surplus 5. Comparative statics and public policy.

BILLY BRAGG is a British folk singer and writer with roots in the early punk music scene of the late 1970s. But, as the biography on his web site says, after "seeing how the Conservative government of Margaret Thatcher was changing the fabric of British society, particularly with the decimation of the mining communities, Bragg's songs became more overtly political. He became a fixture at political rallies and benefits, particularly during the 1984 Miners Strike."[1] He is firmly on the left of the UK political spectrum, and his more recent music deals with the meaning of national identity in economies that have embraced globalization. The song NPWA—No Power Without Accountability— speaks to these concerns in a way that will appeal to those who have lost jobs in declining resource and manufacturing industries, not just in the United Kingdom but in other rich countries. Indeed, it could well be an anthem for many critics of free trade and global governance.

> I grew up in a company town
> And I worked real hard 'til that company closed down
> They gave my job to another man
> On half my wages in some foreign land
> And when I asked how could this be
> Any good for our economy?
> I was told nobody cares
> So long as they make money when they sell their shares
>
> Can you hear us? Are you listening?
> No power without accountability!

---

[1] http://www.billybragg.co.uk/biography/index.html . Accessed October 2, 2009.

> I lost my job, my car and my house
> When ten thousand miles away some guy clicked on a mouse
> He didn't know me, we never spoke,
> He didn't ask my opinion or canvass for my vote
> I guess its true, nobody cares
> 'Til those petrol bombs come spinning through the air
> Gotta find a way to hold them to account
> Before they find a way to snuff our voices out
>
> Can you hear us? Are you listening?
> No power without accountability!
>
> The ballot box is no guarantee that we achieve democracy
> Our leaders claim their victory when only half the people have spoken
> We have no job security in this global economy,
> Our borders closed to refugees but our markets forced open
>
> The World Bank says to Mexico,
> We'll cut you off if you don't keep your taxes low
> But they have no right to wield that sword
> 'Cos they take their orders from the chairman of the board
>
> IMF, WTO,
> I hear these words just every place I go
> Who are these people? Who elected them?
> And how do I replace them with some of my friends?
>
> Can you hear us? Are you listening?
> No power without accountability![2]

These themes and concerns are also voiced by Michael Moore in his famous film "Roger and Me," which describes the impact of factory closures by the automobile manufacturer General Motors on the community of Flint, Michigan.[3] In part they are also the concerns that have made books by Canadian author Naomi Klein, and many other observers of global economic changes, best-sellers.[4] Klein looks at the issue from the perspective of workers in those "other lands," and documents appalling working conditions that in some cases even involve child labour. Clearly the changes and adjustments believed to be wrought by globalization are real, affecting the lives and livelihoods of many workers in resource and manufacturing industries all over the world. Nobody seems better off; except the big companies who somehow evade being held accountable. Billy Bragg's heart felt cry certainly strikes a chord. There is a sense in which he is right, terribly right.

---

[2]Billy Bragg (2002) "NPWA" from the album *England, Half English*. Lyrics available at http://www.billybragg.co.uk/releases/albums/england_half_english/npwa.html . Accessed October 3, 2009. I have taken the liberty of changing "Before they find a away..." as phrased on the web site to "find a way."

[3]See http://dogeatdog.michaelmoore.com/synopsis.html for a synopsis of the film. Accessed October 3, 2009.

[4]See for example Naomi Klein (19–). *No Logo*. Toronto: Random House.

Yet the strong majority of economists would suggest that he has it wrong, terribly wrong. In fact, "strong majority" may be understating the matter. The notion that free trade is a good thing is held almost unanimously by economists. One survey of professional economists in the United States found that 93% would agree with the claim that restrictions on free trade through tariffs and import quotas would reduce economic welfare.[5]

This belief rests in large part on the idea of scarcity, and the concepts that directly follow from it: opportunity costs, marginal reasoning and, as we will explore in more detail in this chapter, the gains from exchange. Though I have no concrete evidence to support this view my own impression is that most of my professional colleagues would feel that the concerns raised by Bragg, by Moore, by Klein, and by many other critics of free trade reflect the fact that they do not fully appreciate or understand these concepts and their long run implications. In short, that these critics are looking at the issue with a blind spot.

Yet, I'm certain those advocating for free trade are often accused of also having a blind spot. As we have already suggested there is a fine line between the crackpot and the scientist, and many policy advocates—both for and against free-trade—frequently claim the authority of science to support their pre-conceived views. But this aside, is there something in the economic method—which can legitimately lay claim to being scientific—that also blinds its practitioners to what others see so clearly?

3.1 Scarcity and choice

Scarcity is a simple, yet a deceptive idea. It certainly should not inform all public policy discussions. But when it should and doesn't, the decisions made are usually done from an overly short-term perspective, are not mutually consistent, generally have hidden or unintended consequences, and are not sustainable in the long-term.

Scarcity goes hand in hand with the necessity of making choices. Choice—the act of accepting a trade-off between alternatives—has an economic aspect when four conditions are met. First, we must have more than one desire or goal. Our "ends" have to be multiple. If we desired only one thing then there really isn't a choice to be made, we would devote all of our resources to that one objective and enjoy as much of it as we could. An individual with a severe drug addiction might in the extreme be characterized to behave in this way.

Second, our multiple goals must differ in their importance, and we must be able to consistently rank them in order of importance. If we desire A, B, and C, we must be able to say, for example, that we prefer A to B, and B to C. And if this is the case we must also be consistent and therefore also prefer A to C. If this was not the case—if in fact C was preferred to A—rational choice would not be

---

[5] Richard M. Alston, J. R. Kearl, and Michael B. Vaughan (1999). "Is there a consensus among economists in the 1990s?" *American Economic Review.* Vol. 82 No. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association (May), pp. 203-209.

possible. It is in this sense that we use the word "rational," to mean consistency in our ranking of the alternatives we face. It is the relationship between these alternatives that is important if we are to have a problem of choice. But if we prefer two objectives to exactly the same degree, or if our preferences are inconsistent across our multiple goals, then we cannot even state our goals in a way that makes logical choice possible.

The two remaining conditions concern the resources available to reach our goals: our time and resources must be limited; and finally, it must be possible to use these resources in alternative ways. When we say that our resources are "limited" we mean this in a comparative sense. If a resource is in comparative abundance it can be because there is simply so much of it, or that our wants are not all that great. But either way, since all of our wants can be fulfilled with the time and resources available to us there is no need to choose between them. Similarly if our resources do not have an alternative use—if they can be used only to fulfill one of our wants but not any of the others—they may be scarce in the sense that they are not sufficient to fully satisfy our desires, but there really isn't a choice to be made concerning how they should be allocated.

If we have multiple objectives, and if our desires for these goals exceed the time and resources that can be used to attain them, then given that these resources can be used in different ways it matters how we allocate them. It matters because our goals differ in their significance. In this context, our behaviour is characterized by choice. "Every act which involves time and scarce means for the achievement of one end involves the relinquishment of their use for the achievement of another. It has an economic aspect."[6] This is what we mean by an opportunity cost. In the context of scarcity every choice we make entails an opportunity foregone. This opportunity cost describes the relationships between our alternative goals. It is these relationships that are the subject of economics. To borrow once again from Robbins:

> The economist studies the disposal of scarce means. He is interested in the way different degrees of scarcity of different goods give rise to different ratios of valuation between them and he is interested in the way in which changes in ends or changes in means—from the demand side or the supply side—affect these ratios. Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative use.[7]

These "ratios of valuation" or opportunity costs are the prices we put on things. They express relative scarcities. Something is not valuable in and of itself—there is no absolute standard of value—but only in relation to what we have to

---

[6]Lionel Robbins (1935), *An Essay on the Nature and Significance of Economic Science.* Second Edition. London: MacMillan and Company, page 14.

[7]Robbins (1935), page 16. It is probably true that when Robbins' wrote the use of the masculine pronoun was not inaccurate. Most economists probably were men. But even so it is also probably true that the writing style reflects the prejudices of the times as not all of Robbins' professional colleagues were men, in particular he had a famous contemporary, Joan Robinson, who was affiliated with Cambridge University. Either way we shouldn't let this cultural baggage distract us from the significance of the statement.

give up to get it. Scarcity, in other words, makes the need to accept trade-offs inevitable.

We should be clear that the opportunity cost of a particular choice refers not to the value of all the possible alternative uses of our resources, but rather to the value of their use in the next best alternative: the one we would have otherwise chosen to pursue. Further, opportunity costs include both the direct and indirect costs associated with a choice. To spend an extra year in school has direct costs associated with the amount of extra money that must, for example, be spent on textbooks and tuition. That money could have been spent in other ways. But it also has an indirect cost reflected in the alternative ways in which the year could have been spent. The year could have been spent, for example, working in a paid job and therefore all the earnings from that job also represent an opportunity foregone.

If economics is a science, Malthus's analysis of material well-being led many to call it the dismal science, but if it were theatre then Robbins' definition of the method would surely imply economics would be a tragedy. We cannot have it all, and tragically have to choose. The idea that social and individual decisions involve trade-offs between competing ends is one of the first principles emphasized in many first-year economics textbooks.

We should be clear that as economists we are not making a judgment about the ends, but rather offering a way to think about the way they are achieved. Economics is significant when attaining one goal implies sacrificing another. It stresses the inevitability of trade-offs; it does not question the importance of the ends. More prosaically:

> ... when we are faced with a choice between ultimates, [Economics] enables us to choose with full awareness of the implications of what we are choosing. ... There is nothing in Economics which relieves *us* of the obligation to choose. There is nothing in any kind of science which can decide the ultimate problem of preference. But, to be completely rational, we must know what it is we prefer. We must be aware of the implications of the alternatives. For rationality in choice is nothing more and nothing less than choice with complete awareness of the alternatives rejected. And it is just here that Economics acquires its practical significance. It can make clear to us the implications of the different ends we choose. ... It makes it possible for us to select a system of ends which are mutually consistent with each other.[8]

There are certainly some public policy analysts who argue that our preferences are wasteful and demand too much from the resources around us. Others go further to say our preferences are influenced and manipulated by advertisers to promote over-consumption. Or that much of what we demand and consume is conspicuous consumption intended to signal status and our position in the hierarchy, a position that has to be reasserted with new fashion statements and

---

[8]Robbins (1935), page 152. Emphasis in the original.

more consumption once others have caught up.[9] The normative implication is that we should be scaling back on these socially wasteful wants and demand less of the resources around us. When society enters into a basic debate about where it wants to go, what should be the objectives of individual and social choice, economists do not really have a role to play, or at least not a role any different than other citizens. The significance of the subject, as the above quote stresses, is to make clear the consequences that follow from making a particular set of choices.

The most immediate and major message is a rule for making choices that in some sense allows us to get the most out of the trade-offs that we have to make, that maximizes the extent to which we can simultaneously pursue our competing goals, that puts us right on the cutting edge of those trade-offs and does not waste any opportunity to make the most out of the situation. That rule, as we have seen, involves equating a marginal benefit to a marginal cost. This is set in the context of an assumption that at some point the marginal cost in the use of our resources will exceed the marginal benefit. If the marginal cost was always greater than marginal benefit, then we would not choose any amount of this alternative; if it was always less we would choose an infinite amount and the resource would be exhausted.

Another major message is that the use of marginal reasoning immediately rules out the idea that bygones should matter in making choices on the margin. Another way of stating this is to say that sunk costs should not play a role in making choices under conditions of scarcity. A sunk cost is a cost that cannot be avoided whether we make a choice or not, and as such they are not a part of the opportunity costs we face. Robbins cites an example that he draws from Winston Churchill, who described the situation facing the government of his country immediately after the first World War.[10] At precisely 11 am on November 11th, 1918 the war ended. For years the economy had been geared to producing all sorts of weapons of destruction at great sacrifice to the material welfare of the population. At that precise moment, however, this productive capacity was of no use, yet many tanks, ships, and guns were still in the middle of production. There was no need for these weapons yet the government gave instructions to complete any projects that were more than 60% done, and for many weeks after the war ended the economy continued to produce military materials that had no use. The fact that they were partially completed was a sunk cost that could not be changed, yet it led even more resources to be used for a wasteful purpose. When decisions are made in this way they reflect regret or pride, and often can lead to significant waste.

This said, we should at the same time note that economic reasoning will not play a role in public policy if the issue is fundamentally not one of scarcity.

---

[9] Offer a footnote here to references dealing with conspicuous consumption, and positional goods–Velben, Hirsch, Frank–to illustrate that this can be the subject for positive analysis. But fundamentally we do not question preferences, and not only take them as given but also stable. Becker 1971 stressed the importance of stable preferences as part of the economic approach.

[10] Robbins (1935), pages 47-48.

When the four conditions outlined above do not in fact hold there is no problem of scarcity, and no role for economic analysis. Human rights are a case in point when they are defined as civil and political rights, as for example in parts of the *Universal Declaration on Human Rights* adopted in 1948 by the United Nations General Assembly. For example, Articles 3 to 21 of the *Declaration* deal with political and civil rights to which everyone is entitled. These range from rights to life, liberty, and security—prohibiting slavery, servitude, the slave trade, torture and cruel, inhuman, or degrading treatment or punishment—to rights to a fair trial, freedom from arbitrary arrest, and rights to freedom of movement, privacy, and freedom of thought and religion as well as rights to political participation. A right entails a duty, and therefore involves not just the right holder, but also an agent who is responsible and held accountable for upholding the right. The right is inherent to the individual, the duty is required of the state. The *International Covenant on Civil and Political Rights* accompanies the *Universal Declaration*, and is the legally binding instrument associated with these rights.

Rights of this sort are considered "universal and indivisible": they are absolute and non-conflicting. As such there is no preferential ranking among them: some rights are not held with more importance than others. It is not as if a government can permit arbitrary arrest and torture because, for example, security is more of a priority, and still claim it is respecting human rights. So in a sense we are not talking about multiple goals, but rather confront only a single objective.

Furthermore, the duty required by the state is often considered a "negative duty," as an obligation to refrain from an action. As such the resources required to undertake these duties are in comparative abundance. In the extreme no resources are required to refrain from undertaking an activity. Negative duties can be performed simultaneously without being limited by scarce resources.

In other words, none of the conditions motivating choice as an economic problem seem to hold. This is not an exercise in trade-offs. It does not entail opportunity costs. As a result, if a right is compromised it is a matter of adjudication, there is a need to make the duty-holder accountable through the law. A right, is a right, is a right; it holds for all people, in all places, at all times; it is upheld or it isn't; and the legal obligation to uphold it is immediate and absolute.

But all of this is not the case when we address economic and social rights, which are seen to entail "positive duties." The *International Covenant on Economic, Social, and Cultural Rights* and other covenants like, for example, the *Convention on the Rights of the Child* speak to rights of this sort. There is much debate in legal and other circles on their nature and significance compared to civil and political rights. The later are often seen as being legally binding; the former as without legal significance, being characterised as a set of policy goals rather than legal obligations.

Economic, social, and cultural rights entail the duty-holder to undertake actions, to offer assistance or aid. As such they entail an opportunity cost. For example, the *International Covenant on Economic, Social, and Cultural Rights* states that, among other things, governments will recognize the right of everyone

to health care, to education, and to an adequate standard of living. In some cases it is very specific about the actions that governments must undertake as, for example, in Article 13, which addresses the right to education:

> The States Parties to the present Covenant recognize that, with a view to achieving the full realization of this right:
>
> (a) Primary education shall be compulsory and available free to all;
>
> (b) Secondary education in its different forms, including technical and vocational secondary education, shall be made generally available and accessible to all by every appropriate means, and in particular by the progressive introduction of free education;
>
> (c) Higher education shall be made equally accessible to all, on the basis of capacity, by every appropriate means, and in particular by the progressive introduction of free education;
>
> (d) Fundamental education shall be encouraged or intensified as far as possible for those persons who have not received or completed the whole period of their primary education;
>
> (e) The development of a system of schools at all levels shall be actively pursued, an adequate fellowship system shall be established, and the material conditions of teaching staff shall be continuously improved.

So a government is faced with multiple goals, not just with respect to the well-being of children versus other groups in the population, not just with respect to the different rights individuals are deemed to hold, but even in the priorities it must set within one of those rights. All of these require resources. Does it allocate them to pensioners or to children; to the fight against poverty or to improvements in education; to primary education or to higher education?

In fact, the framers of the *Covenant* recognize that it may not be feasible to simultaneously fulfill all positive obligations, and that they may be subject to resource constraints. Article 2 establishes the nature of the legal obligation by stating that a signatory is committing to taking steps to fulfill its obligations "to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant by all appropriate means, including particularly the adoption of legislative measures."[11] This is describing an economic problem: "maximum of available resources" and "progressive realization" imply that the positive obligations should be a priority of government spending, but also that resource constraints may not imply immediate fulfilment of the obligation.

As economists we can accept the validity of these rights as an expression of a social goal, but not as a method to attain them. In particular, "the adoption of

---

[11] See the text of the *Convention* available at http://www2.ohchr.org/english/law/cescr.htm. This wording contrasts with the nature of the negative duties described in Article 2 of the *International Covenant on Civil and Political* Rights. It states that each "state party to the present *Covenant* undertakes to take the necessary steps...to adopt such legislative or other measures as may be necessary to give effect to the rights recognized in the present *Covenant*."

legislative measures" may be necessary or not, but it is certainly not sufficient. A right is a right, for all people, in all places, at all times. But in the face of multiple priorities and limited resources passing a law does not make it so immediately and everywhere. These goals still require the design of effective and sustainable programs.

Yet when it comes to acts of public policy informed by this objective it is sometimes the case that the lessons and implications of decision making under scarcity are not appreciated. There is a strong tendency to ignore opportunity costs, and look for the absoluteness of a single goal and the immediacy of unlimited resources. Consider this example involving investments in education that is drawn from a document published by the United Nations Development Program. The document is meant as a "practice note," a practical guide for field staff that is intended to inform and guide their policy making from a rights perspective. After outlining the principles underlying a human right-based approach to poverty reduction—one of which is universality and indivisibility—it offers the following practical example concerning human rights and support for education.

> Studies have traditionally shown that the rate of return on public investment in primary education is higher than on post-primary schooling. Government and donors have a limited amount of funds available to support the education sector. If new studies would to [*sic*] show that post-primary education yields higher returns, would they make a difference?
>
> The market logic would re-direct funds toward higher levels of education, as this provides the highest rate of return. The human rights-based logic, on the other hand, would be unaffected by these results. Government has the responsibility—the International Covenant on Economic, Social and Cultural Rights (article 13)—to respect, protect and fulfil the right to basic education. The article states, "*Primary education shall be compulsory and available free for all*". Hence, the results of the new studies would not really matter as to the priority objective. Primary education is a fundamental right and entitlement; it would keep the highest level of priority focused on basic education even if the rate of returns to such public investment would be lower than other options.[12]

This can only make any economic sense if society has one goal, or if resources are unlimited. But it makes no economic sense in the context defined by the authors: "limited amount of funds." To invest in primary education without regard to its marginal return, even to the point of driving it to zero, wastes resources that could be used to pursue other goals that have a higher marginal benefit. This

---

[12]United Nations Development Program (2003). *Poverty Reduction and Human Rights: A Practice Note.* (June). New York: United Nations Development Program. 19 Pages. Available at http://www.undp.org/governance/docs/HRPN_(poverty)En.pdf. Accessed October 4, 2009. The cited text appears on page 9 as Box 3 in a section of the document called "Integrating human rights in the programming process". Emphasis in the original.

higher benefit would increase society's surplus and in the future lead to more resources that could be allocated to other goals. That we have to choose between primary and higher education, between the rights of children and others, or that we do not have sufficient resources to meet these rights immediately is unfortunate and some would say even a tragedy, but to do otherwise and not make these choices in a rational way through marginal reasoning is not only a tragedy but also a waste.

3.2 The labour market and the production function

It is important to stress, once again, that the significance of economic analysis to issues of this sort is not with respect to the underlying goals, but rather to how those goals can be achieved. Investment in education is actually often seen as a central element in promoting individual well-being and economic growth. In particular, it is sometimes argued that it is one of the ways out of the Malthusian trap that limits the welfare of the individuals occupying the world described by Figure 2.3.

There are two related ways out of this trap. None of them involve the kind logic described above by the authors of the UNDP report, which over time will lead to a squandering of scarce resources. The first way is to invest the surplus, rather than consume it, in things that will increase productivity and reduce population growth. But this requires us to expand our model to include more than one good. And when we do that we also can begin to think about the second way to promote economic growth and well-being: trade. Scarcity, opportunity costs, and marginal reasoning inform the underlying logic, and most notably the logic of the case that economists make for free trade in international markets.

A stylized illustration of our original model depicted in Figure 2.3 is offered in Figure 3.1. In the top panel of this Figure I have taken the liberty of smoothing out the picture presented in Figure 2.3 of how marginal productivity diminishes. For the sake of discussion Ricardo assumed three types of land with productivity falling discretely—that is, in distinct steps—as more and more labour was applied. In Figure 3.1 I am assuming that there is a smooth and continuous decline in the productivity of labour, and represent this as the downward sloping line labeled "marginal product of labour." It is as if there were many, many different plots of land each slightly less productive, and in the limit an infinity of plots. Smooth and continuous relationships of this sort are assumed more often than not in much of contemporary economic theory.[13] I have labeled this curve the marginal product of labour. I have also added a marginal cost of labour

---

[13]To some significant degree this is because they permit the application of mathematics of calculus to the development and analysis of theoretical models. That said, Alfred Marshall did write, on the title page of his highly influential textbook, *"Natura non facit saltum,"* which translated from the Latin means "nature does not make jumps."

curve to represent the subsistence level of output needed to keep the workers alive.

This diagram depicts, in effect, the market for labour services. The labour market is a factor market, a market in which the resources used to produce the output, so-called "factors of production," are exchanged. In the labour market individuals—as workers—offer the services of their labour, they are the sellers; firms demand these services, they are the buyers. This is in contrast to a "goods market" in which the final output is exchanged for other goods that can be consumed: say corn for coffee. In a goods market the firms are the sellers, and individuals—as consumers—are the buyers.

Demand and supply curves are, respectively, marginal benefit and marginal cost curves. The demand curve slopes downward because the marginal productivity of labour falls as it is applied to the fixed factor of production, in this case land. Labour and land are imperfect substitutes in the production of corn. An additional labourer does not fully compensate for the fact that the amount of land of a particular quality is fixed. If this were not the case, if marginal returns were constant, we would be able to grow all the corn we wanted by simply using more and more labour on a given plot of land. In the context of diminishing marginal productivity a firm would only hire additional labour if what it had to pay was lower.

We have drawn the supply curve to be perfectly horizontal because of the very long time horizon assumed in the model, and because of the behavioural assumption that workers use their share of the output to support increased population levels. If the time horizon was very short, say a week, then the supply curve would be perfectly vertical rather than being perfectly horizontal.

In the diagram the quantity of labour is referred to as N, and the price of labour as W, representing the wage rate. The wage rate is measured in the amount of the good produced, so to continue with our analogy of the last chapter W is a given amount of corn. But we can generalize this slightly in a way that will prove helpful when we move to a situation in which more than one good is produced. W can represent a wage rate paid in money, so many dollars per hour of labour (or whatever other unit of time is appropriate: weeks, or years), and P can represent the price of the output, so many dollars per unit of output. If this is the case the "real" wage rate, that measured in terms of the good being produced, is W/P. This is measured as the number of units of output per unit of time since \$/hour divided by units of \$/unit of output is units of output/hour. By simply stating that W is the real wage rate we are assuming that P is equal to one, and this is as good a start as any for our purposes.

The bottom panel is added to represent the relationship between the total output of the economy and how much labour is used. It is called the "production function," the relationship between the amount of labour and the maximum amount of output that can be obtained. The shape of the production function reflects the assumption of a diminishing marginal product of labour. Each extra unit of labour adds less to output then the previous unit. The marginal product of labour in the top panel is the graph of the change in output per unit of labour as the amount of labour increases, it is the slope of the production function.

Figure 3.1: The labour market and the production function in a one good model
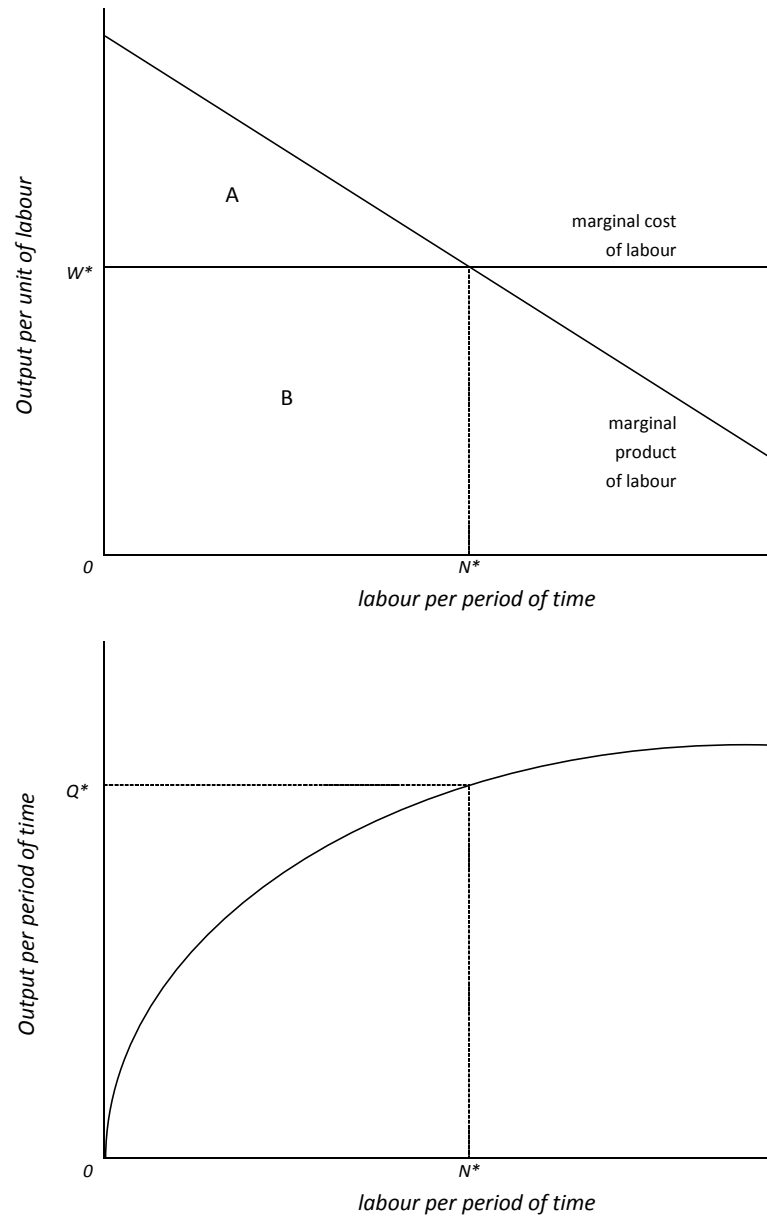
To be "on" the production function is to be "technically efficient." A firm, or for that matter a society, that does not use its resources efficiently is wasting them. If it was better organized and managed it could have more output for the same amount of labour. The production function is depicted for a given amount of the other factors of production, in this case land, but also for a given state of technology. These are the parameters, or exogenous variables, that are assumed constant during the period of our analysis. Given the amount of land and given technical know-how society cannot get more from its labour than that indicated by the production function. It gets less when it is not technically efficient. In our model we assume technical efficiency.

In equilibrium the model determines two unknowns: the quantity of labour, and the price of labour. It reaches equilibrium through the same bargaining process described in the last chapter, under the assumption that none of the underlying parameters change. When this process works itself out the economy is at rest with N^* labour services exchanged per period of time for W^* units of output per unit of labour. The total production of the economy in equilibrium is the area under the marginal product of labour curve up to the point N^*, which corresponds to the combined area of the triangle labeled A and the rectangle labeled B. This can be read directly from the functional relationship between labour and output in the bottom panel. When the level of employment is N^* the total production of the society is Q^*, and of this total landowners receive area A in rents and workers receive area B in wages. In this model the equilibrium level of employment is determined in the labour market. Once this is known the level of output is determined from the production function.

In his book *The End of Poverty* Jeffrey Sachs makes the case that the way out of this Malthusian trap is to take some of the surplus, or more specifically to offer some foreign aid, and invest it in another factor of production, capital. Foreign aid can help to kick start this process if it is invested in physical capital—like fertilizers, machines, infrastructure—that increases the productivity of the land, but also if the resulting surplus is re-invested rather than consumed in augmenting the population.[14] This is not a subtle change. There is a need for a third type of actor in our economy: an entrepreneur, or in other words a 'captialist'. Presumably the landlords need to change from living off rents, to be being future orientated and investing some of the rent in the productive capacity of the land. Alternatively, the workers themselves could begin to play this role if they too are able to become more future orientated in their consumption behaviour, something that admittedly is difficult to do when you are on the edge of survival. Presumably this is in part the role for foreign aid: to ensure

---

[14] Jeffrey D. Sachs (2005). *The End of Poverty: Economic Possibilities for Our Times.* London: Penguin Books. See in particular chapters 12, 13, 14. Sachs' discussion of capital is more nuanced than I have presented. On pages 244 and 245 he speaks of six kinds of capital: human capital (health, nutrition, skills); business capital (machinery, facilities, motorized transport, industry, services); infrastructure (roads, power, water and sanitation, airports and seaports, telecommunications); natural capital (arable land, healthy soils, biodiversity, well-functioning ecosystems); public and institutional capital (commercial law, judicial systems, government services); knowledge capital (the scientific and technological know-how that raises productivity).

a subsistence level for the workers, and unleash their role as entrepreneur. But somehow a third class of individuals comes into being who are motivated by future possibilities, not current consumption.

But if foreign aid is used it must play other roles than just to keep the population alive. It must also be future orientated. In particular, it could be invested in human capital—health, knowledge, education—and particularly in the human capital of women. With higher education levels women will be able to engage more productively in the labour market. This would increase their productivity outside of the home and raise their bargaining power within the household because they now control some part of the household's income. Having a child will be more costly because the mother's opportunity costs of child care and time spent in household production have risen, the opportunities foregone by spending time in the labour market being relatively more valuable. As a result she will have fewer children, and the extra market resources will be used to invest in the quality of the lives the children lead, rather than in the number of children. Fertility rates will fall, population levels will stabilize or even fall, and income per person will rise.

Our economy has changed and looks more like the situation depicted in Figure 3.2. In the new equilibrium the overall employment level of the society is the same, N^*, but the amount of output produced has risen to Q^H. Furthermore all of this increased output has gone to the workers who now receive the wage W^H. This is the result of two things. First, the marginal product of labour has increased, reflecting an upward shift of the production function with a higher physical and human capital stock. Second, this shift in the productivity of labour meets a labour supply curve that is non-responsive in the amount of labour offered. At the same wage rate firms would be willing to hire more workers as they are more productive. Labour has become scarce, and this implies an increase in the bargaining power of labour, which in turn results in wage rate increases.

This new labour supply curve reflects the new behaviour of the population as a result of higher levels of human capital. A vicious circle has been turned into a virtuous circle. As time goes on more investments in capital and technology will lead labour to be more productive and continue to shift the demand for labour outward. If labour supply remains fixed, wages and hence living standards will continue to rise in the new equilibrium. But this also implies that we have more than one good in our economy. Now there are at least two goods: say agricultural products, and manufactured goods like the machinery being used on the farms.[15] But where does this machinery come from? Perhaps some of it came from foreign aid, but over the long term it can have only two sources: either the population made it on its own, or it gets it from other countries by trading some of its agricultural output on international markets. So now we have to confront another very important public policy question: would this society be better off under autarky or under international trade?

---

[15] If development goes even further, our economy may urbanize as population leaves the countryside and agriculture for the cities where it produces manufactured goods and services like education and health care.

Figure 3.2: Escaping the Malthusian trap by investing in physical and human capital
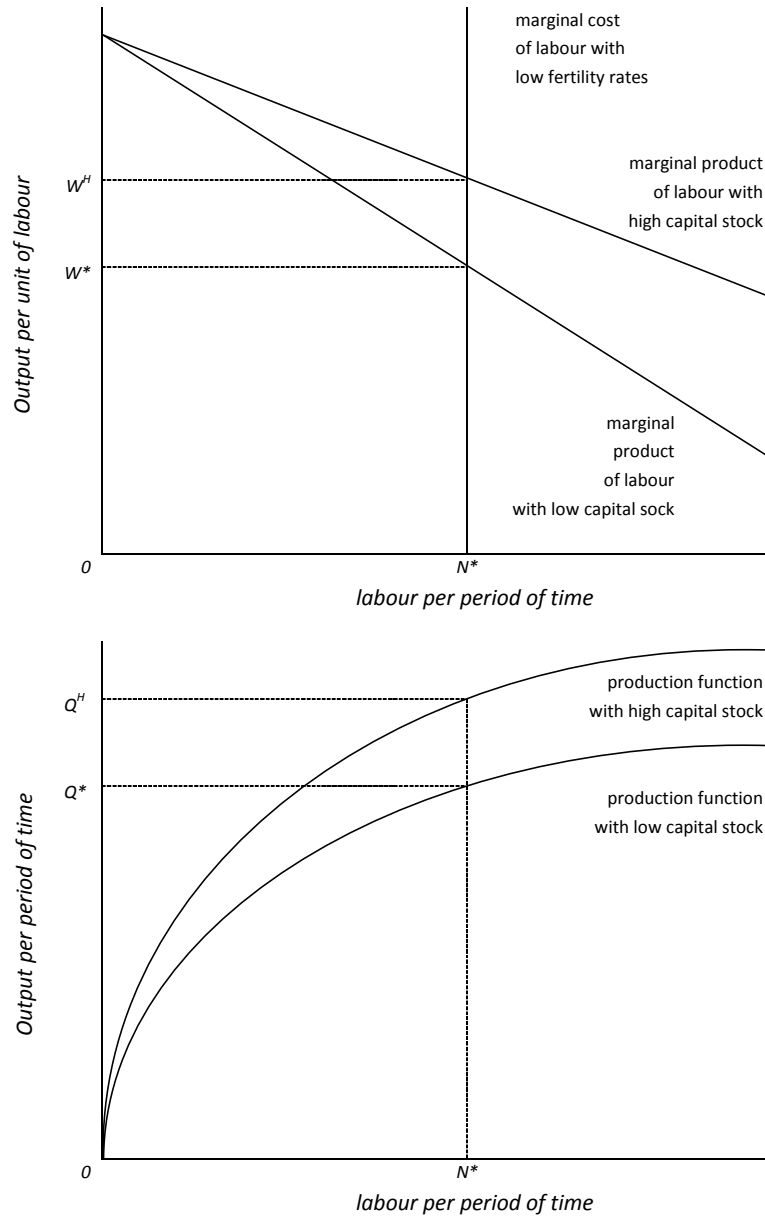
3.3   A two-good model and the gains from trade

As I have already suggested, the strong majority of economists would answer this question by saying that free trade is preferred. What is it that economic theory teaches us about this issue?

The case for free trade rests upon a logic derived from scarcity, opportunity costs, and marginal reasoning but now applied to a two-good world. With two goods there is the possibility of exchange, of trading. In effect, we are asking the more general question: are there gains from trade?

David Ricardo is the starting point for the economic analysis of international trade. Ricardo made a very strong statement about the advantages of free trade, at a time when it was a politically charged issue.

> Under a system of perfectly free commerce, [Ricardo wrote,] each country naturally devotes its capital and labour to such employments as are most beneficial to each. This pursuit of individual advantage is admirably connected with the universal good of the whole. By stimulating industry, by rewarding ingenuity, and by using most efficaciously the peculiar powers bestowed by nature, it distributes labour most effectively and most economically: while, by increasing the general mass of productions, it diffuses general benefit, and binds together by one common tie of interest and intercourse, the universal society of nations throughout the civilized world. It is this principle which determines that wine shall be made in France and Portugal, that corn shall be grown in America and Poland, and that hardware and other goods shall be manufactured in England.[16]

Whoaaa! This is as provocative now as it was 200 years ago. That trade is a positive sum game, win-win, making everyone better off is still a hard to accept proposition. This is because Ricardo's description of the underlying principle—that of comparative advantage—is still not fully part of public discourse. But the logic remains as valid now as it was then, and is the reason his name is still the starting point for any serious discussions of international trade.

Ricardo was writing in the aftermath of the Napoleonic wars. England had been pretty well continuously at war with France from the French revolution in 1789 until the defeat of Napoleon in 1815. These wars had the effect of disrupting international trade. At that time England was an exporter of manufactured goods and an importer of agricultural goods, but maritime transport was disrupted by state sponsored pirates and a French blockade. This restricted the supply of agricultural products coming into England, and as a result made the price of food relatively high. English landowners were the main beneficiaries, and manufacturers—who had to pay more for the labour they employed while at the same time selling less of their production—the losers. But after the war food prices fell as foreign supplies of agricultural products were brought back

---

[16]David Ricardo (1817). *On the Principles of Political Economy and Taxation.* Volume I of The Works and Correspondence of David Ricardo. Piero Sraffa (editor). Cambridge: Cambridge University Press. Pages 133-34.

onto the English market. This threat to their fortunes led the politically influential landowners to get legislation passed that, in essence, imposed a tariff on grain imports, with the result that supply was again restricted and relative prices maintained at higher rates. Ricardo was arguing for the repeal of these "Corn Laws," and he cast his argument in a way that showed society as a whole would gain from "free commerce," or "free trade" as we would call it today.[17]

He used a hypothetical example of trade between Portugal and England that involved just two goods, cloth and wine, and one factor of production, labour. This set of assumptions has become known as the Ricardian model of international trade, and it explains trade by differences in the relative productivities of labour that arise from differences in natural endowments. Two countries trade with each other because they are different in these endowments. We can think of the endowment as being associated with geography or climate, though its sources might be more general also reflecting cultural and institutional advantages or the results of public policy. Think, for example, of the advantage that a country whose population is English speaking might have in a world that increasingly uses English as the language of trade, commerce, and cultural exchange.

The important point is that in Ricardo's example the two countries differ in their opportunity costs of producing the goods they are interested in consuming. The example is set up in a very specific way.

> England may be so circumstanced, that to produce the cloth may require the labour of 100 men [yes men again!] for one year; and if she attempted to make the wine, it might require the labour of 120 men for the same time. ... To produce the wine in Portugal, might require only the labour of 80 men for one year, and to produce the cloth in the same country, might require the labour of 90 men for the same time.[18]

These labour productivities are listed for the sake of clarity in Table 3.1. The first thing to notice is that both countries are capable of producing both goods on their own. It is easy to rationalize trade if a country has no capacity to produce a good. The only way it can have any of the good is to trade for it. Bananas don't grow in Canada; if Canadians want to eat bananas they have to trade something for them. But that is not the case here.

Further, it might also be easy to rationalize trade if one country was very productive in one good, and the other country very productive in the other. Okay I suppose if you really tried hard, and spent a lot of time and money you could grow bananas in Canada. But the country is much more efficient at designing and manufacturing Blackberries. Similarly, I suppose if you really tried hard, and spent a lot of time and money you could produce telecommunications equipment in Costa Rica. But the country is much more efficient at growing

---

[17]This summary of the political context is taken from the brief discussion in Paul R. Krugman and Maurice Obstfeld (2003). *International Economics: Theory and Policy.* Sixth Edition. Boston: Addison-Wesley. Page 59.

[18]Ricardo (1817), page 135.

Table 3.1: The Ricardian model of trade: labour requirements per year to produce one unit of output

	Portugal	England
wine	80	120
cloth	90	100

bananas. It is easy to see that these two countries should trade manufactured goods for agricultural products. But again this is not how Ricardo set things up.

What Table 3.1 shows is that Portugal is more productive than England in the production of both goods: it takes less labour to produce one unit of wine in Portugal, and it also takes less labour to produce one unit of cloth. Does it make sense for these countries to trade even if one of them is more productive in everything? What does England have to offer Portugal? And if there was free trade don't the Portuguese pose a threat to the jobs of all English workers, whether they work in vineyards or in factories?

Even under these conditions the two countries will find it mutually advantageous to trade. To see this play with the numbers a bit. Assume each country wants to consume one unit of wine and one unit of cloth. Then under autarky this would require the English to use a total of 220 units of labour, and the Portuguese to use 170 units. Total world output is two units of wine and two units of cloth, based on the work performed by a total of 390 workers.

Contrast this with a situation in which each country produces only one good, specializing in the good for which it is relatively more productive, and then they trade one unit of wine for one unit of cloth. It is true that Portugal is more productive in both goods, but comparatively speaking it is more productive in wine than in cloth. It is true that England is less productive in both goods, but not as much so in the production of cloth. So Portugal produces all of the world's wine. Two units of wine require 160 units of labour. And similarly England produces all the world's cloth, these two units requiring 200 units of labour. Total world output is still two units of each good, but now only 360 units of labour was used to produce it. The Portuguese have used 10 units less of labour than under autarky; the English 20 units less. They each have as much to consume as under the previous scenario of no trade, but they have used less labour. The labour they saved could be used to produce a bit more of something thereby increasing overall production and consumption. When production is specialized on the activity that each does best, and trade permitted, everyone is better off.

This happens because the opportunity costs of producing an extra unit of the goods differs in the two countries. The opportunity cost of one unit of wine in Portugal is $\frac{8}{9}$ths of a unit of cloth. To produce one extra unit of wine the Portuguese must transfer 80 workers from cloth production to wine production.

As a result the production of cloth falls; but it falls by less than one unit, to be precise it falls by 80/90, or $\frac{8}{9}$ths of a unit. In England the opportunity cost of an extra unit of wine in terms of cloth is greater. Fully 120 labourers must be transferred from cloth production in order to get an extra unit of wine. But if this many labourers leave cloth production the amount produced falls by more than one unit: it falls by 120/100, or 1 and $\frac{1}{5}$th units. While the opportunity cost of wine production is lower in Portugal than in England, the opposite is the case with respect to cloth production. In Portugal an extra unit of cloth implies the country must forego $\frac{9}{8}$ units of wine; in England only $\frac{5}{6}$ths of a unit of wine needs to be given up.

Ricardo's hypothetical example illustrates that the same amount of output—two units of wine and two units of cloth—can be produced using less labour if each country specialized in just the good for which its opportunity costs are lower compared to the other country: wine for Portugal, and cloth for England. The important phrase is "compared to"; and we therefore speak of having a comparative advantage in the production of a good. To determine where the comparative advantage lies requires us to examine all four of the labour productivities.

This is why the English have something to offer the Portuguese. It is true that the Portuguese have an absolute advantage in both goods—something we can determine by pairwise comparisons of the labour productivities across the two countries—but this does not make a case against international trade. Rather, the English offer the Portuguese the opportunity to focus on the good in which they are relatively more productive, and with the extra output they produce make use of the English production technology to make the amount of the other good they desire.

In a similar way imagine the situation from the perspective of the English. Why are the Portuguese not a threat because they are more productive, in an absolute sense, in both goods? For England, trade should be seen as just another way of producing wine. England can produce wine using English labour, or using English labour to produce the requisite amount of cloth to buy it from the Portuguese. In this model there is only one resource, labour, and while it can be used in two ways—in the production of wine or in the production of cloth—its use is subject to constraints of technology, as described in Table 3.1, and in the amount that is available. Labour is scarce, and as a result the act of producing wine and cloth involves a trade-off. The terms of this trade-off are going to be different according to whether the goods are produced domestically or obtained through international trade. That is what a comparison of opportunity costs between the two countries tells us.

There is a more analytical way of illustrating this that uses the production functions in each country to derive what in economic analysis is refered to as the 'production possibility frontier'. Let's refer to the domestic labour requirements to produce one unit of wine and one unit of cloth—the productivity of labour—as a_W and a_C. In the case of England a_W is equal to 120 and a_C is equal to 100. To be clear, these numbers refer to the productivity of labour; the lower they are the less labour it takes to produce one unit of output, and hence the greater

Figure 3.3: Production possibilities frontier and consumption possibilities with international trade for England in the Ricardian model of trade
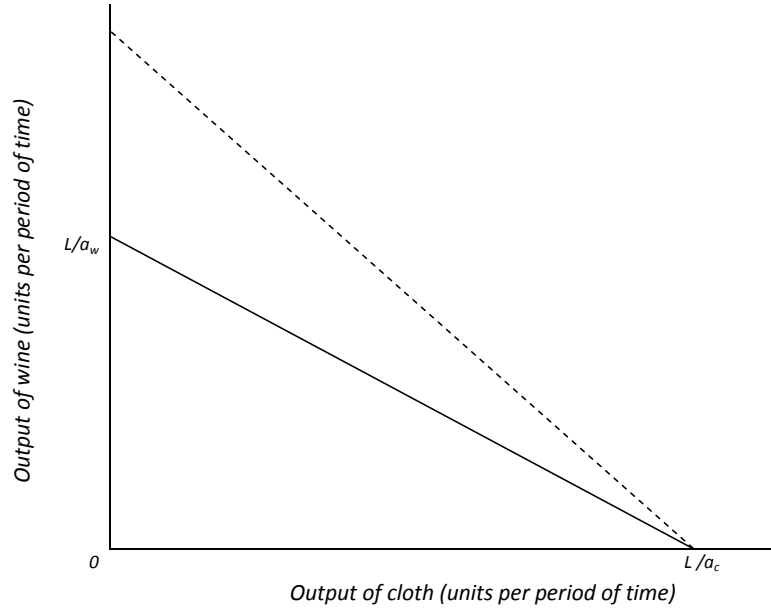
the labour productivity. They are constant because there is only one factor of production, and the state of technology is assumed not to change. In addition, let's call the total amount of labour in the society L; the amount of labour used in the production of wine L_W; and the amount used in the production of cloth L_C. The total amount of labour used in the production of these two goods cannot exceed the total amount available. That is $L_W + L_C \leq L$, where $\leq$ is the mathematical symbol meaning "less than or equal to." If all of the labour was applied to the production of cloth then $Q_C = L/a_C$ units of cloth could be produced. Similarly if it was all applied to the production of wine then $Q_W = L/a_W$ could be produced. Any combination of these two amounts could also be produced subject to the amount of labour used not exceeding the total available. This is illustrated in Figure 3.3 as the solid straight line, which is referred to as the production possibilities frontier. It represents the upper bound of the set of all possible combinations of the two goods society could produce, given the total amount of resources available. The efficient use of these resources

will place the society on this frontier, an inefficient use somewhere below it.[19]

There are two issues raised by the production possibility frontier, one technical and one social: how do we get to the frontier? where on the frontier do we choose to be? The first is a question of engineering and management; the second a question of social choice. We assume that the former is dealt with, and as positive analysts we avoid addressing the second. But if we can do anything to get beyond the frontier, beyond what is technically feasible, we will be better off. International trade offers that possibility.

The production possibilities frontier in this model is a straight line because there is only one factor of production, there is no diminishing marginal productivity. The slope of this line represents the opportunity costs of cloth in terms of wine, how much wine must be given up for an extra unit of cloth. This is L/a_W divided by L/a_C or simply a_C/a_W, the relative productivities of labour in the production of the two goods. This is the trade-off that England faces in producing an extra unit of cloth by reducing the production of wine and reallocating labour between the sectors. But the labour saved in reducing wine production by one unit could produce more wine than the amount foregone if it were in Portugal, and the only way of getting it there is by actually producing the cloth, thereby embodying the labour in it, and trading for the wine. Portugal offers a different opportunity cost between these goods, and through trade in effect offers England the chance of consuming beyond the possibilities that its technology and labour constraint impose. These new consumption possibilities through trade are illustrated by the dotted line in the figure.

This is possible because England has a comparative advantage in the production of cloth. Another way of saying this is $a_C/a_W < \alpha_C/\alpha_W$ where the α's represent the labour requirements in Portugal. It may be that α_C is less than a_C, and that α_W is less than a_W, in other words that Portugal has an absolute advantage in the production of cloth and also in the production of wine. But it is not these absolute advantages that determine the gains from trade, it is the relative or comparative advantage. Portugal offers England a relatively less costly way to "produce" wine, and at the same time the English offer the Portuguese a relatively less costly way to "produce" cloth. While the English have a comparative advantage in the production of cloth, the Portuguese have a comparative advantage in the production of wine. The contrast between the two lines in Figure 3.3 tells us that for every possible combination of wine and cloth produced under autarky there is through international trade the possibility of

---

[19]Convince yourself that another way of stating the labour constraint is as equation 3.1.

$$a_W Q_W + a_C Q_C \leq L \qquad (3.1)$$

When we re-arrange these terms to express the equation in terms of one of the outputs, say wine, we can give expression to the society's production possibilities set.

$$Q_W \leq \frac{L}{a_W} - \frac{a_C}{a_W} Q_C \qquad (3.2)$$

Technical efficiency implies an equality in this relationship, and it takes the form of the equation for a straight line. This is exactly what is illustrated in the figure.

getting more of one good without reducing consumption of the other. In this sense there are gains to trade that will make the society better off compared to autarky.

3.4 The gains from trade and the distribution of the surplus

All of this is an ingenious piece of logic with a certain political appeal. Everyone wins under free trade. And on the basis of this logic Ricardo became an influential politician who contributed to the repeal of the Corn Laws. There is a certain truth to this model that has stood up for 200 years or so. The underlying assumption that countries trade with each other because of differences in relative productivities leads to the prediction that we should observe trade of different goods. The exports of a country should be made up of goods with relatively high labour productivity. Indeed, much trade is of this sort. For example, China has a comparative advantage in labour intensive manufacturing, while the United States has a comparative advantage in human capital intensive goods and services. The trade patterns between this pair of countries and many other pairs reflect these relative advantages.[20]

However, on other important aspects the model is entirely mute. While it gives some indication of inequalities at the world level between countries, it says nothing, for example, about the distribution of the surplus within a country: what happened to the landowners and their rents? The only factor of production in this model is labour, so by construction the model cannot speak to this issue. This simplification allows a focus on the idea of comparative advantage and the resulting gains from trade, but as you can imagine this is not exactly what the landowners of Ricardo's time had in mind. A move to free trade may generate a surplus, but public policy may still be interested in who gains, who loses, and by how much.

To the extent that we wish to speak to this concern we have to reconsider the model, and its underlying assumption of a single factor of production. This is a good example of some of the issues raised in Secton 2.2 of Chapter 2. A model is not inherently right or wrong; it is has to be judged in terms of the purpose to which it is put. Ricardo's model is a very simple representation of reality, and in its simplicity it generates important predictions about the underlying reasons countries trade with each other: in particular trade patterns reflect differences in inherent productivities. But it assumes a single factor of production with constant marginal returns. Labour is the only factor of production, and as a

---

[20]But this prediction of the model is refuted by the facts in other cases. For example, trade between many rich countries occurs with similar goods. An important part of the trade between Canada and the United States is in cars and car parts that flow in both directions across the border. This reflects the fact that the marginal productivity of labour may rise as the scale of production expands, that there are increasing returns to scale. We leave the rationale that countries may trade to exploit these sort of economies aside, but this is discussed in Paul Krugman (2009), "The Increasing Returns Revolution in Trade and Geography." *American Economic Review*, Vol. 99, Number 3, pages 561-71.

result the model offers no predictions concerning the impact of international trade on the distribution of the surplus. The gains from trade make the country as a whole better off, but also each and every individual.

If this is the issue we are interested in then we need to modify the underlying assumptions to permit other factors of production, and a certain specificity in how they can be used in production of the goods. The way to think about this is that we are in the world described by Figure 3.1 but with three differences. First, the total labour supply is fixed. Second, there are two production functions, one for each of two different goods—say an agricultural product and a manufactured product—with both characterized by a diminishing marginal product of labour. Third, there are other factors of production specific to each good, say land that is used in the production of the agricultural good, and physical capital that is used in the production of the manufactured good. Land and capital do not move between these two sectors of the economy, but labour does.

Given that industries differ in the factors of production that they use, a change in the overall mix of the goods produced by the economy as a result of international trade will also have implications for domestic factor markets. In Ricardo's time the repeal of the Corn Laws would have lowered the price of corn, raised that of manufactures, and led to a decline in the demand for land, an increase in the demand for capital, and finally a movement of labour from the production of agricultural goods on land to the production of manufactured goods in factories. This would have implications for the distribution of society's surplus as the demand for land would fall along with the price of agricultural goods. Landowners would lose; capitalists would gain; and the impact on workers as a group would be ambiguous.

There is also a more analytical way of illustrating this. The production possibilities frontier of this society is determined by what happens to the production of each of the goods as labour is shifted from one sector to the other. It is a technical relationship reflecting the technology of production. But it now looks different than in Figure 3.3 because the technology of prodcution is different. If both sectors are characterized by the type of production function illustrated in the bottom panel of Figure 3.1, with the only difference being that the fixed factor differs—in agriculture it is land, and in manufacturing it is capital—there is a diminishing marginal product of labour in each sector and the production possibilities frontier is no longer a straight line. It is bowed out, or convex, as depicted in Figure 3.4. This shape reflects the diminishing marginal product of labour in each sector. It implies that the opportunity cost of producing more of one good varies with the amount already produced, and is represented by the ratio of the marginal products of labour in the two sectors. For example, if a lot of the manufactured good is being produced the marginal product of labour in manufacturing will be low. At the same time this implies not so much of the agricultural product is produced, and that the marginal product of labour in that sector is high. To obtain one more unit of the manufactured product will require that labour be reallocated from agriculture to manufacturing, but since this implies a movement from a high to a low productivity sector more than one unit of the agricultural product will have to been given up. The opportunity

Figure 3.4: Production choices in a two good specific factors model closed to international trade
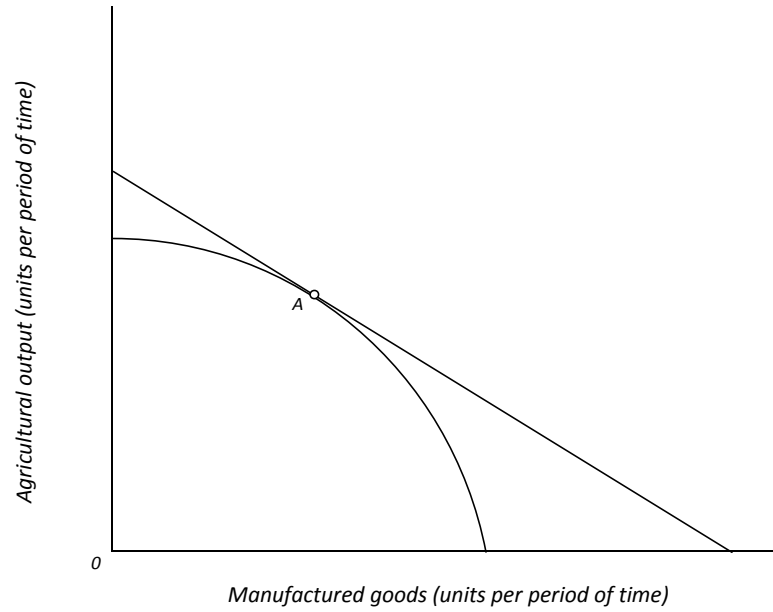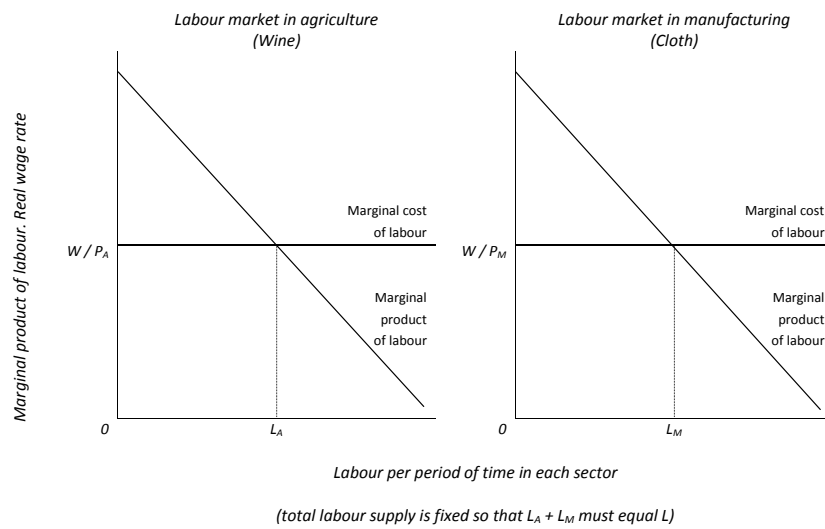

Figure 3.5: Labour markets in the two good specific factors model

cost of an extra unit of manufactures is high compared to agriculture. This is reflected in an increase in the absolute value of the slope of the production possibilities frontier as the amount of manufactures produced increases.

What is the mix of outputs produced? Where on the production possibility frontier are we? Once again, this is not just a technical issue, but rather also a matter of the preferences of the individuals in this society. Since we are no longer in a one-good world our analysis must deal with one more complication: prices, the rate at which one good is exchanged for another. These preferences will be reflected in the markets for these goods, a discussion of which we defer to the next chapter. For present purposes it is only important to realize that this will imply a set of relative prices for the goods. In Figure 3.4 these prices are represented by the slope of the solid straight line, and at point A this is the same as the opportunity costs in production. The combination of the two goods produced at point A is the best in the sense of maximizing the surplus in society: at this point the marginal benefits equal the marginal costs of prodution for both goods.

In turn, the relative price of the goods implies a particular allocation of labour, something determined in labour markets. Labour is allocated in the two sectors according to the supply and demand in each of the factor markets. Labour demand is characterized by a diminishing marginal product, and as a result the number of workers landowners and factory owners are willing to hire depends upon the money wage rate and the price of the good being produced. The labour markets for both goods are depicted in Figure 3.5. The price of the agricultural product is indicated as P_A and that for manufatured good as P_M, while the money wage rate is W. The decision on how much labour to hire depends upon the real wage rate the landlord or the capitalist must pay relative to the price of their product, that is the money wage rate relative to the price they receive for the output. I have drawn these to be the same in each of the sectors by assumption. This leads to L_A workers in agriculture and L_M in manufacturing. As before, the total amounts of each good produced by these workers are the areas underneath each of the labour demand curves, and are also represented as point A in Figure 3.4. As mentioned, the straight line in Figure 3.4 also represents the relative prices of the two goods: the absolute value of its slope is P_M/P_A. I have drawn the surpluses produced in each sector, the areas of the triangles above the marginal cost of labour and below the marginal product to be equal, though they are strictly not comparable because they represent physical quantities of different types of goods: wine and cloth for example. But clearly the interests of landowers and of capitalists depend upon the amount of surplus generated in each sector, and indeed if we assume, as a starting point, that the prices of the two goods are the same the figure would suggest that the total value of the surplus accuring to each is the same. In this sense the distribution of the wealth is relatively equal.

Compare this situation with an economy that is exactly the same, except that it is open to international trade. Now the relative prices of the goods will be different as they are the result of the overall world demand and supply of the goods. To continue with the example faced by Ricardo, the price of

the agricultural product is lower, while that of the manufactured good higher. This is illustrated in Figure 3.6 by the dotted straight line and the comparison with B, implying production of more manufactures and less agricutural output. How does this come about? The new set of product prices imply that relative scarcites in the labour market change. The lower price of the agricultural product implies that the real wage landlords must pay in terms of their product is higher. At the new real wage rate they do not demand as much labour. In the manufacturing sector the opposite is the case, the real wage is lower since the price of manufactures is higher, and as a result labour demand is higher. The excess supply of labour in agricultural production must be reallocated from the land to the factories.

The model continues to illustrate that there are overall gains to trade. The society could sell the increased production of the manufactured good on international markets and consume the same amount as at point A. This would be illustrated by a movement up the dotted line in Figure 3.6. But if it does this it will have more of the agricultural product than it did at point A: it has reached a point beyond the production possibilities it had under autarky.

At the same time, however, this overall gain for the society as a whole does not imply that all indivdiuals have gained. This is illustrated in Figure 3.7, which makes clear that this society is characterized by more inequality between landlords and capitalists. The surplus accuring to landlords is lower, while that of capitalists is higher. The impact on workers is ambiguous. It will depend upon the proportion of each of the goods in their consumption, and upon the final level of their money wage rate, W, which may also change in the process. For example, if they consume only food then they are concerned only with P_A relative to their money wage. If W is the same then a lower price of agricultural products implies that their real wage W/P_A is higher, and since in equilibrium everyone is employed the fraction of output going to them has gone up. Workers will come under pressure from landlords to lower the money wage. But the opportunity to work in the expanding manufacturing sector has made labour relatively more scarce, and given workers another option. They exercise this change in their bargaining power by leaving agricluture, and this process stops when their real wage adjusts to be the same in both sectors.

The model now generates a prediction concerning distributional issues, and in this way can be helpful in understanding public policy issues and the politics associated with a move to free or freer trade. International trade increases the share of total surplus obtained by the factor that is used relatively more intensively in the production of the good that is exported; it decreases the share of the surplus obtained by the factor that is used relatively more intensively in the production of the good that competes with imports. The factor of production that is mobile between the two sectors may gain or lose. This is different than the model put forward by Ricardo, which predicts not only that both countries gain from trade but also that every individual does as well.

Much still depends upon how things unfold in subsequent periods. If the extra surplus generated from international trade is saved and invested in a way that increases future productive capacity then it will contribute to the kind of

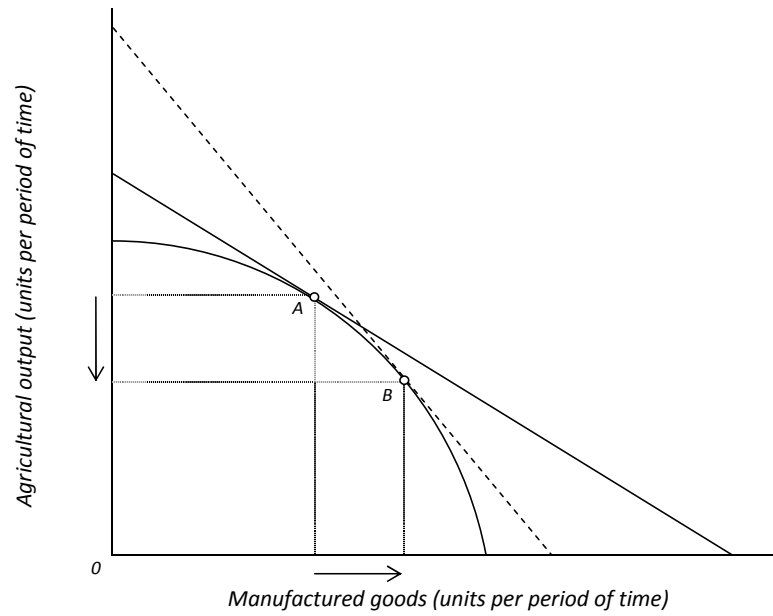Figure 3.6: Production choices in a two good specific factors model with international trade

Figure 3.7: Labour markets with different relative prices in the goods market
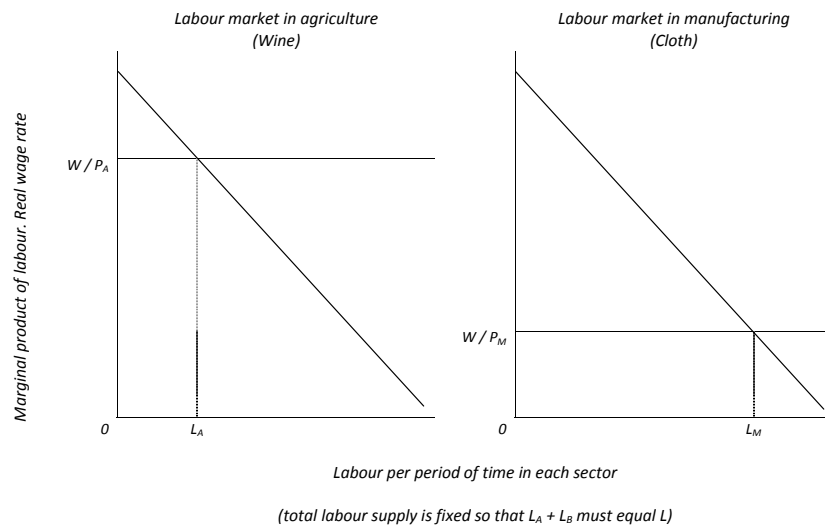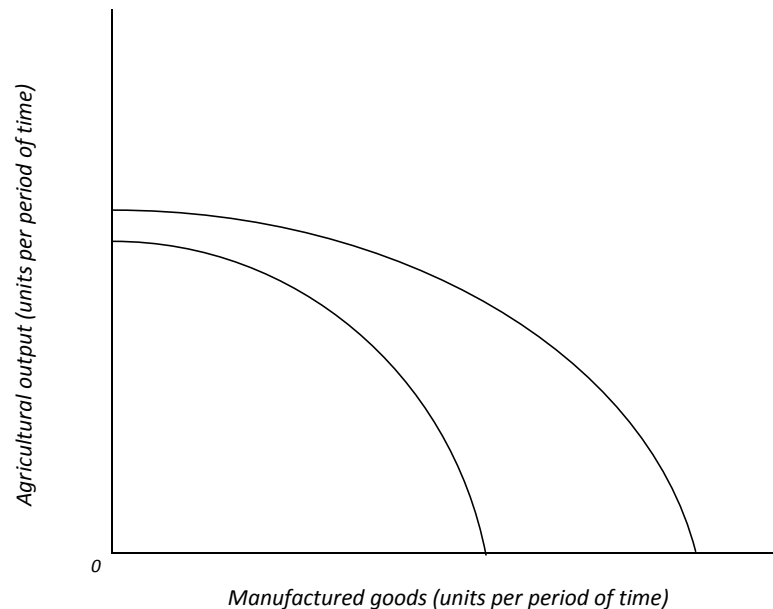
Figure 3.8: Production possibilities for the two good model with growth

virtuous circle originally illustrated in Figure 3.2. In our two good model the upward shift in the production function illustrated in this figure takes the form of an expansion of the production possibilities frontier, which I illustrate in Figure 3.8. In this particular case investment increases the production possibilities more in manufacturing than it does in agriculture, but it could be the opposite depending upon how the investments are made. This particular configuration might in a coarse way illustrate the nature of economic growth in China, with a rapidly expanding manufacturing sector drawing labour from the countryside and agricultural production to the cities and manufacturing, just as easily as it could have illustrated the same process during Ricardo's time. If this expansion occurs at a rate faster than population, incomes per person will rise, and they might rise at such a rate that the distributional conflicts originally hinted at by the model fade into history.

3.5 Comparative statics and public policy

Evaluating changes in public policy in this way requires a certain caution on the behalf of the economic theorist because the method being used involves a comparison of different equilibrium outcomes, rather than charting out a process of change from one equilibrium to another. This method of analysis is called comparative statics. The label is telling. It involves a comparison of market

outcomes that are at rest. It is not a method that incorporates historical time. An appreciation of this limitation allows us to interface more naturally with actual political processes: our economic models illustrate that there are potentially significant gains in moving from one equilibrium to another, that these gains represent incentives in market-based economies that generate a force for change, but at the same time our models do not describe the process of change. They leave the door open to politics as the art of moving—or for that matter not moving—from one equilibrium to another.

For example, the analysis in the previous section assumes that labour can be re-allocated between the two sectors without any costs. It is cold comfort to tell someone like Billy Bragg that this is a long-run model describing the nature of two different equilibria after all adjustments and changes have taken place. The assumption that workers can move costlessly between sectors—that there is no specificity in the use of workers in each sector associated with where they are located, the skills they have, or their age—is a reflection of the focus on long-run outcomes. But how long is the long-run? Is it a week, or is it a life-time? There is nothing in our model permitting us to answer this question because strictly speaking it is not a model in historical time, but rather a comparison of two economies that are identical with the exception of the value of a single underlying parameter.

The proper analogy for our methods is not history, but rather mechanics. When we speak of an economy reaching an equilibrium it is like talking about a pendulum that swings back and forth and eventually comes to rest, or a marble rolling back and forth in bowl and eventually becoming motionless at the bottom. This is a process that occurs in space rather than in historical time. If we perturb the pendulum the process will repeat itself in a way that does not depend upon the history of past pertubations. So 'long run' should be thought of as a position of rest that is attained in self-contained 'model time'. Societies and economies are not like that. They exist in 'real' or historical time with a present that depends upon a particular history, and facing a future that is unknown but dependent upon that history. The consequences of any change we introduce depend upon the history of previous changes that brought the economy to a particular point in real time. And they also depend upon the expectations of future developments. In real time the underlying parameters could themselves also change before the new equilibrium is reached.

So strictly speaking a discussion of two positions, like that of A and B in Figure 3.6, is not one involving movement from A to B, but rather a comparison of an economy closed to international trade to a hypothetical and identical economy that differs only in that it is open to international trade. Time, dynamics, movement don't enter into the procedure. This is a laboratory experiment that occurs in our minds. We are offering a comparison of static positions. As such this method of economic theorizing has a certain blind spot, and invites a need for knowledge of history, institutions, and politics. But even so it is very helpful in clarifying the incentives, motivation, and behaviour of different groups in response to economic developments or public policy changes. In fact, in some cases we may even understand the long run in a way that allows us to argue that

insitutions and public policy are endogenous, themselves changing in response to underlying economic incentives.

If we appreciate the two great questions that economics helps answer, then we also appreciate that any public policy change will imply not just a change in the overall size of the surplus but also a change in its distribution. Every change in public policy implies that there will be winners and losers, and often the arguments these groups make for or against a particular policy will be couched in terms of the public good in a way that masks underlying normative positions. Indeed, as has been suggested, Ricardo did as much himself, a wealthy stock-broker somewhat on the outside of mainstream British society whose personal interests were much more closely aligned with the growing entrepreneural class than with the landed gentry. His focus on comparative advantage placed the emphasis on the good of the whole, abstracting entirely from the issue of income distribution. And this from an economist who began his great book on the principles of economics by stating that distributional issues were at the very core of the subject.[21]

But this said, the point we should take away from the principle of comparative advantage is that the case against international trade cannot be made solely on distributional grounds. Changes in the distribution of a soceity's surplus occur for all kinds of reasons, and are not unique to international trade. They also result from changes in technolgoy, changes in consumer preferences for different consumption goods, the exhaustion of old resources, and the discovery of new resources. Preventing these changes because they imply changes in income distribution would likely have a determintal impact on the overall growth of the surplus that would outweigh the problem being addressed.

One aspect of the political process is that losses are typically highly concentrated and immediate, while gains are diffuse and in the future. This can imply a bias in social decisions because groups that are negatively affected can be highly organized. The introduction of international trade or a move to freer trade may have a significant effect on their livelihood. Groups that gain may not be organized at all as the benefit at the individual level is small even though it touches a large number. This may lend an imbalance in the political process and skew social choices toward the interest of the few at the expense of the

---

[21]The opening paragraphs of *On the Principles of Political Economy and Taxation* read:

> The produce of the earth—all that is derived from its surface by the united application of labour, machinery, and capital, is divided among three classes of the community; namely, the proprietor of the land, the owner of the stock or capital necessary for its cultivation, and the labourers by whose industry it is cultivated.

> But in different stages of society, the proportions of the whole produce of the earth which will be allotted to each of these classes, under the names of rent, profit, and wages, will be essentially different; depending mainly on the actual fertility of the soil, on the accumulation of capital and population, and on the skill, ingenuity, and instruments employed in agriculture.

> To determine the laws which regulate this distribution, is the principal problem in Political Economy...

whole.

Our analytical appartus helps us to understand contempary debates over free trade: whether they represent fundamental flaws in logic and a misunderstanding of the importance of opportunity costs; whether they draw attention to just losses or to just benefits without an appreciation of both marginal costs and benefits; whether they focus just on the short term discounting long term gains; and even whether they represent outright self-interest disguised as a claim for the public good.

The argument it makes for international trade rests upon the result that the total gains are large enough to compensate the losers for their losses. The degree to which these losses are in fact compensated, and how, is the outcome of the poltical process. If this compensation is in some sense deemed inadequate, or does not take place at all, it is a reflection of the fairness of the political process, not an argument against trade.

Chapter 4

Prices in perfectly competitive markets

1. Prices: Nominal, Real, and Relative. 2. Exchange, specialization, and the market. 3. Demand schedules and demand curves. 4. Supply schedules and supply curves. 5. Equilibrium and comparative statics. 6. Applying a model of perfect competition to agricultural markets.

IN DISCUSSING the Malthusian trap of poverty and despair we assumed a one-good, closed economy. Since there was only one good, corn, we could have a simplified discussion, abstracting from exchange and therefore from prices. If there is only one good, there is nothing to trade, and no need for prices. The way out of this trap was to relax this assumption along both of its dimensions. On the one hand, we suggested that the economy needed in some sense to be more future orientated, to forgo current consumption, to save and invest. In order to capture this we assumed more than one good, this other good we called 'capital'—using the word in a very general sense to mean both human and physical capital. As a result we need to be concerned with how a society's total wealth is divided between those two goods, and therefore at what ratio they trade. Our argument was that if current production can somehow be allocated in part to more future orientated activities that increase productivity, rather than solely to current consumption that augments the population, the society will become richer. In some sense we are making a trade between current consumption and future consumption, and this trade will be governed by a relative price. So in order to assess the terms of this trade we need to know the price at which consumption in the present trades for the possibility of consumption in the future. This price is the interest rate, and we will address it as a part of our discussion of macroeconomics in Chapter 9.

But our argument was also that trade could occur across space, not just across time. If our economy desires more capital, then in addition to producing it domestically we could also trade for it with other countries. But here again we

require a set of relative prices to govern the terms of the exchange. In fact, these prices were in the background of the simple Ricardian model of international trade. They were more explicitly necessary in our discussion of the specific factors model of trade, where another commodity also slipped into the analysis: money.

So whether through time or across space, we are left to deal with the question of what determines the rate at which one good will exchange with another, or in other words the price of one good in terms of another? This is a question that we need to explicitly answer, particularly if we are making a recommendation that an economy engage in international trade. The price that the good being exported commands in international markets will determine the value of the imports a country will be able to afford, and the extent to which trade actually raises material well-being.

It is in an institution called 'the market' that the buyers and sellers of goods and services come together and determine not only how much will be traded, but also at what relative price. And this is exactly what this chapter is about: the theory of value—at what ratio do goods exchange? What determines the price of a good relative to another? Buyers and sellers have of course different, indeed directly opposing, interests: as a consumer my material welfare will depend upon getting the lowest possible price for the things I buy, but as a producer it depends upon getting the highest possible price for the things I sell. How do markets work to determine prices, and how are they manipulated through the political process to the benefit of some and the detriment of others? These are the questions motivating our discussion.

The theory of value is at the heart of economics and its development: from Adam Smith, who published *The Wealth of Nations* in 1776—a book many take to mark the birth of economics as a science—through the works of David Ricardo and Karl Marx who followed in his tradition during the early to mid 1800s, and finally to the work of William Stanley Jevons and others in the late 1800s. For the so-called 'classical' theorists like Ricardo and Marx the value of a commodity was determined by the costs of producing it. This supply-side model required a comparison to some commodity that had an inherent and invariant value. For Ricardo this commodity was gold; for Marx it was the amount of labour used in production, and hence his 'labour theory of value.' In actual fact such a commodity never existed. Later researchers like Jevons, but also those from continental Europe, focused attention on the satisfaction or utility derived from consumption. This demand-side model saw value as being determined by how much satisfaction was derived from a commodity. But value is not so subjective as to be just in our minds.

This debate and its resolution are summarized in a metaphor used by Alfred Marshall, the Cambridge University professor whose famous text book was first published in 1890, went through 8 editions, and was taught well into the 1920s to generations of economics students. Marshall wrote:

> We might as reasonably dispute whether it is the upper or the under
> blade of a pair of scissors that cuts a piece of paper, as whether

> value is governed by utility or cost of production. It is true that
> when one blade is held still, and the cutting is effected by moving
> the other, we may say with careless brevity that the cutting is done
> by the second; but the statement is not strictly accurate, and is to
> be excused only so long as it claims to be merely a popular and not
> a strictly scientific account of what happens.[1]

The relative price of a commodity and the quantity of it exchanged are determined simultaneously by the supply side—the marginal costs of production—and the demand side—the marginal benefits associated with consumption. The two blades of Marshall's scissors are, of course, the supply curve and the demand curve.

Marshall's methodology puts to one side a question that was an element of the way many, Marx being only one example, phrased and indeed continue to phrase the issue: what is a fair, or just, price? But this way of thinking is immediately disconcerting for the analyst interested in 'positive' economics. 'Fair' for whom? 'Just' in what sense? Clearly we have to take a philosophical stance on what is a 'good' society, and evaluate outcomes in light of that standard. But whose standard, whose idea of the good, should we base this evaluation upon? We are moving along the lines of an argument about whether actual prices should be higher or lower, and this is an exercise in normative economics.

Before we can have this sort of discussion, about what prices should be, we need to understand what determines actual prices. It is important to know why actual prices are what they are, and how they change. In this sense Marshall's framework is helpful, but only if we use it in the context for which it is appropriate. This chapter focuses on the determination of relative prices in the context of perfectly competitive markets. In this environment, demand and supply curves are the tools to undersand prices and the quantities exchanged, and with this model we will then be in a position to evaluate policies intended to change market outcomes according to a sense of what is good, just, fair, or simply to the benefit of some and the detriment of others.

But before we do any of this we need to understand a basic technical issue about the measurement of prices in economies that use money. What exactly do we have in mind when we say the price of something is so many US Dollars, so many UK Pounds, or so many Euros? And what sort of price is our model attempting to determine? We are interested in the price of one commodity relative to another, but the prices we observe in reality are the prices of a commodity relative to a unit of the currency used in that society. In some sense a Dollar, a Pound, or a Euro might be taken to represent a fixed amount of all other commodities: a very convenient stand-in for comparative purposes. But this requires us to acknowledge that the value of money may change. A generalized decline in the purchasing power of money is called 'inflation' and a generalized increase is called 'deflation'. Our analysis of relative prices must be concerned with 'real,' or in other words inflation-corrected, prices if we are interested in

---

[1] Alfred Marshall (1920), *Principles of Economics.* 8th edition. London: MacMillan and Co., page 348.

Figure 4.1: Coffee prices on the German International Exchange: Brazilian Natural Arabicas, 1976 to 2009
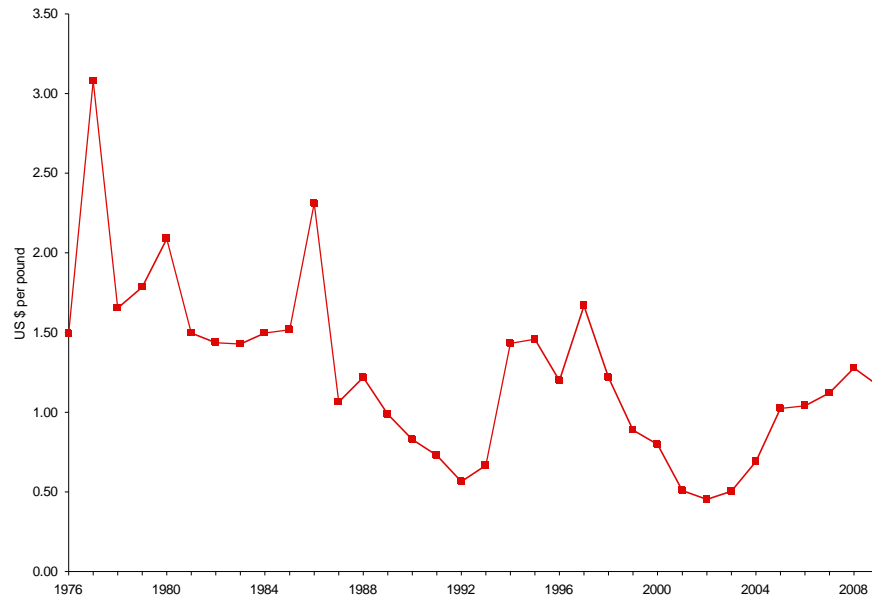
Source: International Coffee Organization, http://dev.ico.org/prices/p2.htm accessed on October 18, 2010, and historical statistics provided by the International Coffee Organization.

the welfare of consumers and producers. Appreciating this distinction is our first challenge in developing the so-called 'neoclassical' theory of value.

4.1 Prices: Nominal, Real, and Relative

As an example consider the case of my favourite drink, the cappuccino. It takes two principle natural ingredients to make a cappuccino: coffee and milk. Some countries specialize in the production of coffee, which is traded in large international markets. Milk on the other hand tends to be produced and consumed locally. Just as importantly, we should note that coffee and milk prices behave very differently, in a way that implies coffee producers are more than likely to be poor and to stay poor, while milk producers will be relatively rich and probably richer over time.

Figure 4.1 presents the price per pound of a particular type of coffee bean, in a particular market, since the mid 1970s. These are the prices that coffee retailers, be they Starbucks, Second Cup, or Timothy's World Coffee, have to pay in order to get the beans that they will ultimately transform into espresso,

Figure 4.2: Fresh milk prices relative to 2002 price: Canada, 1976 to 2009
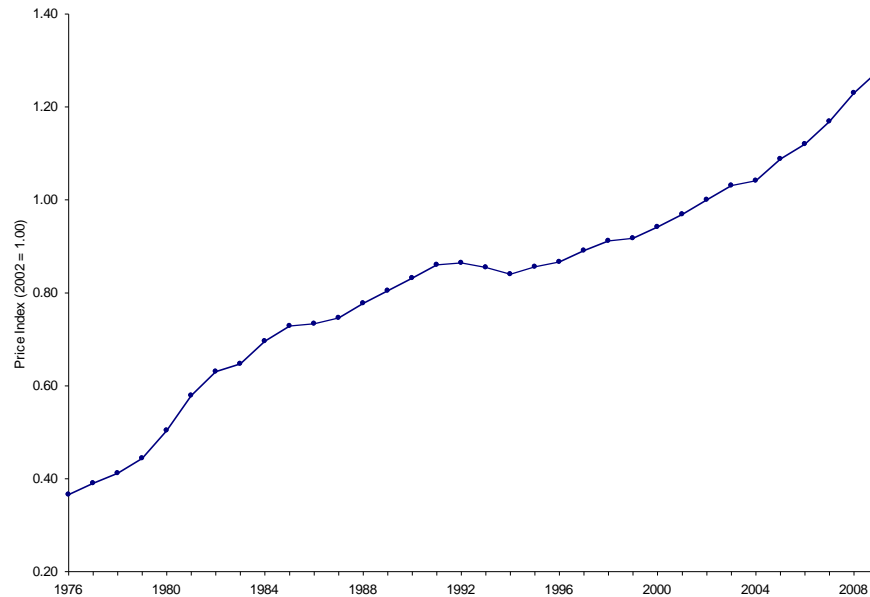
Source: Statistics Canada, CANSIM data base accessed October 18, 2010.

cappuccino, or simply a brewed cup. These are the prices that in turn coffee producers receive for the beans they harvest from the crops they have planted and nursed.

There are three important things to notice about this time series, all of them relate directly to the well-being of coffee producers, and also indirectly coffee consumers. The first is the very sharp variation: a pound of coffee can cost significantly more in one year compared to the next, and it can drop just as significantly from year to year. The second is the trend, which is downward: over time coffee is cheaper and cheaper. So on two important dimensions of individual well-being, coffee producers are in a precarious position: they have little security as the price of the good they sell fluctuates dramatically; and their standard of living is gradually diminishing as the price trends downward.

These two patterns are brought into even sharper clarity by contrasting them with the price of milk, which is presented over the same time period in Figure 4.2. In contrast to coffee, milk prices do not fluctuate significantly from year to year, and seem to increase steadily throughout the thirty or so years covered by the graph. Milk producers have experienced both security and an increase in their economic well-being.

In fact, the way I have chosen to present milk prices introduces the third important thing about Figure 4.1. Figure 4.2 actually offers not the price of a

litre of milk, but rather an index of this price: the price of milk in any particular year relative to what it was in 2002. That is why the value in the graph for 2002 is exactly 1.00: it is the actual price in 2002 divided by itself. All the other prices are calculated in this way; by dividing the actual price in a particular year by the price that prevailed in 2002. The way to read this information is as follows: if a certain quantity of fresh milk cost $1.00 in 2002, then the graph shows how much that same amount of milk would cost in any other year. In other words, for every $1.00 spent on milk in 2002 you would have paid less than 40 cents to get the same amount in 1976, and you would have paid $1.28 in 2009. The reason I have chosen to present milk prices this way, in terms of a fixed quantity rather in terms of its money price, is that the value of money is also changing through time. It is as if the measuring rod we are using changes year by year, and if that is the case then we cannot be certain whether we are measuring the real price of the good or the degree to which the units of measurement have changed. In fact, over this time period money has lost value. We refer to this as inflation. By using an index based upon a fixed quantity of a specific good of a given quality—in this case, the price of a particular amount of milk in 2002—we can be confident that we are measuring the evolution of the relative value of milk over time.

In this sense, Figure 4.1 could well be a misleading representation of coffee prices if there have been changes in the general level of prices over time, which in fact is the case. When we correct for the decline in the value of money a different picture emerges, one that suggests that the situation for coffee producers is much worse. In Figure 4.3 I present two time series of coffee prices, repeating the information from Figure 4.1 in the line labelled 'Nominal prices', and offering an inflation-corrected series labelled 'Real prices'. When we account for the fact that a dollar in the 1970s could buy a lot more goods than the same dollar in the 2000s it is clear that the living standards of coffee growers has fallen to a much greater degree. In inflation adjusted terms, the price of a pound of coffee has gone from around $5.00—rather than $2.00—to less than $1.00. This is a very important caution all economists are immediately aware of when examining prices over time: are they measured in current monetary values, or have they been corrected for inflation?

When the prices of all goods rise, the purchasing power of money falls; conversely when prices fall across the board, the purchasing power of money increases. The word 'inflation' is used to refer to a general rise in prices and the concommittant fall in the value of money, while 'deflation' is used to refer to a general fall in prices and the resulting increase in the value of money. Movements in the general price level—inflation and deflation—are an important topic in macroeconomics; relative prices—the ratio at which one good exchanges for another—are the subject of microeconomics. When relative prices are expressed in comparison to money we call them 'real' or 'inflation corrected' prices. Sometimes the phrase 'constant prices' is also used to mean the same thing.

Money has sometimes been called a 'fictious' commondity.[2] Fictious it may

---

[2]Karl Polanyi (1957), *The Great Transformation: The Political and Economic Origins of*

Figure 4.3: Nominal and real coffee prices on the German International Exchange: Brazilian Natural Arabicas, 1976 to 2009
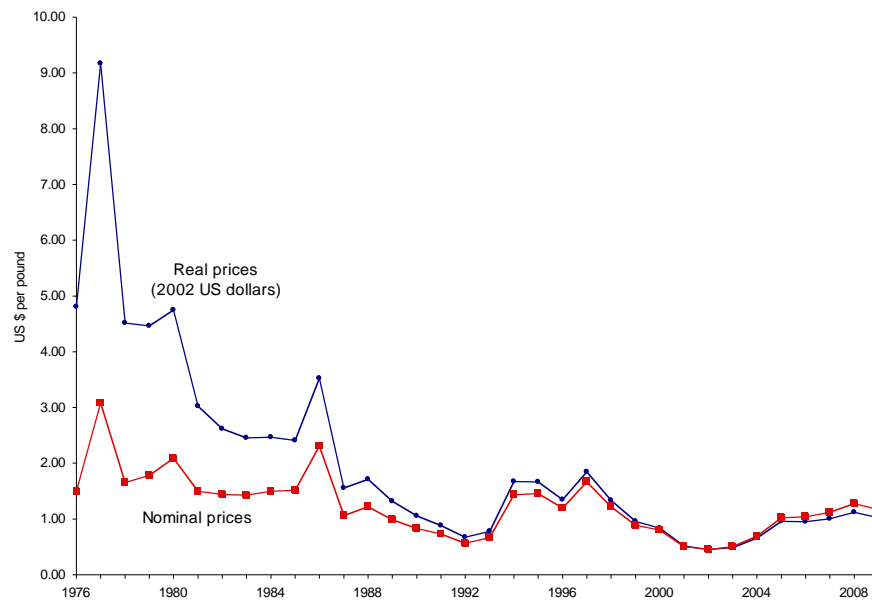
Source: International Coffee Organization, http://dev.ico.org/prices/p2.htm accessed on October 18, 2010 and historical statistics provided by the International Coffee Organization. Prices are deflated using the All Items CPI.

be, but it nontheless plays three important and interrelated purposes. The first is to act as the accounting unit in which values are measured. We may call currencies a Dollar, a Pound, a Euro, or a Drachma, but we have to make measurements in some unit, and each society is distinguished by how it refers to its accounting measure.

The second purpose of money is to act as a medium of exchange. Without money individuals in a society would have a much harder time engaging in trades because they would have to find another person who not only has what they want, but also wants what they have. This coincidence of wants implies that the economy will be forced to engage in barter, which could be very inefficient and put a brake on any possibility of economic growth. No advanced economies rely upon barter. If all members of an economy reach a social consensus that a particular object—and it could be any object—will act as a medium of exchange then it is only necessary to sell to someone who has a desire for the good offered, and only necessary to buy from someone who offers the good desired. Historically, the evolution from barter to monetary economies was associated with objects that had an inherent desire by all members of society becoming the medium of exchange. For example, iron or gold because they were malleable and could be formed into tools, or were otherwise realtively rare and whose ownership reflected social status, played this role.

In order to be a unit of account and a medium of exchange, money must also act as a store of value. This is the third purpose of money. In the time between I sell what I have to offer and I find what I desire, I have to hold money. Therefore, I need to have confidence that money keeps its value; that I can use it to store the value of the goods I have sold until I trade this value to buy the goods I desire. From this perspective money is nothing more than the physical embodiment of a social consensus. If all members of society agree that an object, any object, will act as a store of value, then that object will be money. If this is the case then it also would be more efficient not to tie up large amounts of useful objects to play this role. And so, rather than being coins made of silver, copper or gold, money has evolved to be paper, or even in the extreme, entries in accounts that are transferred digitally between buyers and sellers.

The value of money can change dramatically. Indeed, some countries have experienced hyper-inflation, extraordinary rates of inflation with prices rising by hundreds or even thousands percent over the course of a year or shorter periods. But rates of inflation as low as two to five percent per year can erode the purchasing power of money significantly over a decade or so. As such, even if inflation is not extreme it is important because individuals care about relative prices when they engage in trades. These trades are mediated by money, and therefore individuals need to form expectations of future inflation rates in order to be able to conduct their transactions with confidence. Workers may strike a bargain with employers over their wage rate, but this bargain needs to incorporate an idea of what will happen to the purchasing power of their earnings

over the term of the contract. Firms need to make investments in capital and incur the costs of other inputs with an accurate anticipation of revenues that will accrue over the course of many years into the future. And to accurately form these expectations individuals and firms need a way of measuring inflation.

Price indices are used to estimate the extent to which the purchasing power of money changes. The most commonly used index is the Consumer Price Index, which is often referred to by its acronym: the CPI. This is a price index that measures the change in the cost of a fixed basket of goods and services purchased by the average consumer. It is commonly cited in popular discussions and in the media because it is the appropriate index to use when referring to price changes faced by consumers. One of the important responsibilities of national statistical agencies is to calculate and publish appropriate measures of inflation like the CPI. Statistics Canada, the national statistical agency for Canada, describes the derivation of the CPI in this way:

> It is obtained by comparing, through time, the cost of a ***fixed basket*** of commondities purchased by ... consumers in a particular year. Since the basket contains commodities of unchanging or equivalent quantity and quality, the index reflects only ***pure price movements***.[3]

Other statistical agencies like the Bureau of Labor Statistics in the United States, or the Office of National Statistics in the United Kingdom, also have this responsibilty, and the methods used across these and other rich countries are similar; reflecting a scientific consensus that is coordinated by the United Nations.

It should first be noted that the CPI is not the only measure of inflation, there are other indices. The Industrial Product Price Index measures price movements of goods that are produced by the manufacturing industry; the Raw Materials Price Index measures price changes of the raw materials purchased by producers; the New Housing Price Index measures changes in the selling prices of new residential houses; and so on for the Machinery and Equipment Price Index, the Non-Residential Building Constructtion Index, the Farm Input Price Index, the Farm Product Price Index and many others. The broadest price index measuring price changes of all goods and services in the economy is called the GDP Deflator.[4] [5]

---

[3]Statistics Canada (1996), *Your Guide to the Consumer Price Index*, Ottawa: Statistics Canada, Catalogue Number 62-557-XPB, page 1. Emphasis in the original. This publication is the major reference for the paragraphs that follow.

[4]GDP is the acronym for Gross Domestic Product, the total value of all final goods and services produced in the economy during a particular period of time, usually a year.

[5]Family expenditure surveys are conducted to determine the goods and services making up the basket upon which the CPI is based. The households in these surveys, who number in the tens of thousands, are chosen to be representative of the national population. They are asked to record or recall their expenditures on goods and services over a period of time, and with this information the items purchased by a typical, or average, household are determined. These family expenditure surveys are conducted at regular intervals, and the underlying basket of goods are updated to reflect changes in spending behaviour, and the availability and quality of

It should also be noted that these indices do not measure price levels. The CPI, for example, is not a measure of actual prices, just as the information in Figure 4.2 does not indicate the actual price of a litre of milk. Price indices show the rate of change of prices through time relative to some specified base year. In Figure 4.2 the base year was 2002. The information in this Figure suggests that in 2007 the milk price index was 116.9 (2002=100). The food price index, measuring the change in prices of the basket of all food items purchased by the typical consumer, was 111.8 in 2007, and the CPI was 111.5 in the same year and for the same base year. From this information we cannot say anything about the actual price of milk, or of the bundle of food items typically purchased. What they tell us is that milk prices rose by 16.9% between 2002 and 2007, while for all food items they rose by 11.8%, roughly the same as for all items purchased by the typical consumer. Milk prices rose by more than the price of food in general, and more than all items purchased by the average consumer.[6]

All of the goods included in the basket making up the CPI must be associated with a retail price, and the statistical agency collects these prices on a regular basis by actually visiting establishments and retailers across the country. For most commodities this occurs each month of the year, and these prices—over 60,000 of them in the Canadian case—are then used to derive an index for each commondity that is in turn appropriately weighted to represent that commodity's relative importance in total expenditures. These weighted indices are then combined to calculate the CPI, or for that matter any index based upon selected sub-components. The 'base year' of an index is the point in time that the index is given a value of 100. The choice of a base year is important for interpretation, but once all prices are expressed in the same base year it does not influence the calculation of the rate of change of prices. any year can be chosen as the base year, but generally we choose a year close to the present

---

goods. The goods and services included in the CPI are classified into eight broad categories: food; shelter; household operations and furnishings; clothing and footyear; transportation; health and personal care; recreation, education and reading; alcoholic beverages and tobacco. Milk for example is a product that belongs to the food major component of the CPI, and as such it is grouped with other items having this common end use or considered to be substitutes for each other. See Statistics Canada (1996) for more details.

    [6]The other caution we should exercise is to note that the CPI is not strictly speaking an index of the "cost of living." Statistics Canada (1996, page 3) explains the distinction between a cost-of-living index and the fixed basket concept underlying the CPI as follows:

> ... the objective behind a cost-of-living index is to measure price changes experienced by consumers in maintaining a constant standard of living. The idea is that consumers would normally switch between products as the price relationship of goods changes. If, for example, consumers get the same satisfaction from drinking tea as they do from drinking coffee, then it is possible to substitute tea for coffee if the price of tea falls relative to the price of coffee. The cheaper of the interchangeable products may be chosen.

> We could compute a cost-of-living index for an individual if we had complete information about that person's taste and consuming habits. To do this for a large number of people, let alone the total population ... is impossible. For this reason, regularly published price indexes are based on the ***fixed basket*** concept rather than the cost-of-living concept.

because it relates more directly to our current experiences.

Table 4.1 presents the value of the all items CPI between 1990 and 2010 with 2002 as the base year. It also offers the index for the food component of the all items index, and individual indices for two of the goods that are part of this category: fresh milk and coffee. The most common calculation performed with indexes of this sort is the determination of percentage changes between two points in time. For example, over the 20 year period covered by this table consumer prices increased by 45.9%, milk prices by somewhat more at 53.6%, and coffee prices significnatly less by 27.2%.[7]

But another very important calculation is to use the CPI to compare money values over time, and that is exactly how it was used in deriving Figure 4.3. Since the CPI measures the change in the overall price level it can be used to measure changes in the purchasing power of money, and in fact it is one of the most commonly used indexes for this purpose. For example, Table 4.1 uses a base year of 2002=100, and shows that in 2009 the value of the CPI was 114.4. This implies that consumer prices increased 11.4% between these years. In other words $114.40 dollars in 2009 are equivalent to $100 in 2002; every dollar in 2002 has the same value as $1.14 in 2009. Or this can be expressed the other way around by dividing 100 by 114.4, which implies that 87.4 cents in 2002 would be equivalent to $1.00 in 2009. The change in the purchasing power of money can be calculated not just from the base year, but for any two periods. For example $100 in 1990 is equivalent to $\frac{114.4}{78.4} \times \$100 = \$145.92$ in 2009.

Monetary values are often expressed in 'current' dollars, that is in the dollar value prevailing at the time it is observed. We also refer to current dollars as 'nominal' dollars; both words mean the same thing, referring to the observed monetary values in the market. In the autumn of 2010 a small cappuccino cost $3.10. This is its current or nominal price. Monetary values expressed in nominal dollars cannot be compared over time if the value of money is changing, and for this reason we use price indices to derive a series of comparable values all expressed in terms of the value of money in some specific year. These derived values are referred to as 'constant' or 'real' dollars. The labelling is admittedly a bit confusing: 'real' dollar values are never actually observed in reality, they are something we derive.[8]

Consider another example. I bought a house in 1992. At that time I paid $100,000. I sold my house in 2002 for $200,000. On this sale I made $100,000 in nominal terms, but this clearly overstates the value of my gain because I spent $100,000 current 1992 dollars on the house, and received $200,000 current 2002 dollars; dollars, that because of the general rise in prices, are not worth as

---

[7]The change in the CPI is derived by using the values for 1990, 78.4, and 2009, 114.4, in the following calculation: $\frac{(114.4-78.4)}{78.4} \times 100$ or equivalently $[\frac{114.4}{78.4} \times 100] - 100$. The calculation can be done over any two years, with the index for the more recent year always being divided by the older index. It is not appropriate to simply calculate the difference between the indexes as the result will depend upon base period.

[8]The CPI is often used to make these derivations, but in doing so it should be clear that this is done because we wish to express the situtation facing consumers. We may have other purposes in mind, and in these cases we would be required to use the appropriate index.

Table 4.1: Consumer Price Indexes (2002=100)

Year	All Items	Food	Fresh Milk	Coffee
1990	78.4	79.6	83.2	92.7
1991	82.8	83.4	86.0	88.4
1992	84.0	83.1	86.4	80.5
1993	85.6	84.5	85.5	77.5
1994	85.7	84.9	84.0	97.9
1995	87.6	86.9	85.6	133.0
1996	88.9	88.0	86.6	119.1
1997	90.4	89.4	89.1	120.2
1998	91.3	90.9	91.2	120.0
1999	92.9	92.0	91.7	114.5
2000	95.4	93.3	94.2	108.6
2001	97.8	97.4	96.9	105.2
2002	100.0	100.0	100.0	100.0
2003	102.8	101.7	103.1	100.9
2004	104.7	103.8	104.2	102.3
2005	107.0	106.4	108.8	106.6
2006	109.1	108.9	112.0	111.0
2007	111.5	111.8	116.9	111.0
2008	114.1		123.0	112.8
2009	114.4		127.8	117.9
2010				

Source: Statistics Canada (2007), *Canadian Economic Observer, Historical Supplement*, Ottawa: Statistics Canada, Catalogue No. 11-210-X, Table 12, and CANSIM.

much as much as they were 10 years earlier. To calcuate my gain it is necessary to express all monetary values in real terms, say in 2002 dollars. From the information on the CPI in Table 4.1 the value of a dollar in 1992 expressed in real 2002 dollars is $\frac{100}{84.0} = 1.19$, implying that \$100,000 in current 1992 dollars is equivalent to \$119,000 in real 2002 dollars. So in fact my monetary gain over this period is \$200,000–\$119,00=\$81,00 2002 dollars. This result is based upon a calculation that holds the purchasing power of money constant.

Monetary values expressed in real terms determine my material well-being, and therefore it is important that I base my economic decisions on them rather than on nominal values. The prices we are concerned with in microeconomics are referred to as 'relative' prices: the price of one good divided by the price of another. Real prices expressed in a given base year are relative prices: they are the price of a good relative to the price of all other goods, which is what money is standing in for. To obtain real prices we need only divide nominal values by the value of the appropriate price index. This is exactly what was done to calculate the real price of coffee in Figure 4.3. As such, a particular quantity of coffee is expressed in terms of the quantity of other goods it can command.

Relative prices are important in microeconomics because under certain circumstances they represent an opportunity cost, a marginal evaluation. If one kilogram of coffee costs \$1.00, and one litre of milk costs \$1.00, then the relative price of coffee in terms of milk is one kilogram per litre. If I purchase a kilogram of coffee I give up the opportunity to purchase a litre of milk. If there is a fall in the value of money so that all prices are higher, and as a result coffee costs say \$100 per kilogram and milk \$100 per litre the relative price of coffee in terms of milk remains one kilogram per litre. It is still the case that if I purchase a kilogram of coffee I give up the opportunity to purchase a litre of milk. The inflation of all prices does not change anything, and will not influence the decisions of consumers and producers because it has not altered opportunity costs. On the other hand if the marginal valuation of one of these goods changes, or if the marginal costs in their production changes then the opportunity cost of coffee in terms of milk will change, and this ratio becomes an important signal that will affect consumption and production decisions.

As micro-economic analysts we have to be clear that the only role of money is as a stand-in for all other goods, and therefore that money prices should be measured in constant dollar values. Our concern is with relative prices, and how they are determined. This is what we mean by a theory of value. We put aside the causes and consequences of inflation and deflation, and address them as a topic in macroeconomics. For now we assume that the rate of inflation is either zero, or perfectly predictable by all participants in the microeconomy in order to focus on what determines relative prices.

4.2 Exchange, specialization, and the market

It takes quite a few things to make a cappuccino.[9] First, and most obviously, both milk and coffee. But also other ingredients. Certainly other natural resources like water and sugar, and also a nice china cup and saucer, a spoon of just the right size to fit along side, and a machine of stainless steel, tubing, and gauges that can be used to drive the hot steam at the proper temperature through the beans that have, in their turn, been ground to an appropriate texture using another machine. All this to say nothing of the barista's skills, who is capable of putting all these things together, including heating the cup beforehand with hot water, and then capping them with the foamed milk that gives the drink its name. But this is not the only person involved. In fact, no one person can make this drink; it takes, without exagerration, the coordinated efforts of a whole number of specialized operations and labourers in several countries. Who grew the beans? Who sorted them, roasted them, packaged them? Who was involved in getting them here? And so on for the milk, never mind even thinking about how the machine was designed, the parts manufactured and assembled, and the whole thing transported to the particular establishments that want one just like it.

And so there the cappuccino sits on the shiny table top in front of you: a beautiful symbol of the benefits that flow from the coordinated efforts of specialized labour. But it got there because of the opportunity to trade. Specialization makes sense only if exchange is possible. The opportunity to trade allows us to obtain a greater variety of goods, and in that way increases our material well-being. I may have milk, and you may have coffee. But I will forever be drinking hot milk, and you at best espresso, if we cannot somehow communicate, bargain, and exchange. If that is the case, then not only do we both have milk and coffee, but also the possibility of a cappuccino is born.

But the opportunity to trade does more: it permits specialization in production. Specialization offers us the chance to increase our productivity, and can thereby further increase both the quantity and the variety of goods available to us. If I cannot engage in trade then I need to produce both milk and coffee, and in doing so I probably will not be as good in the production of either than if I had focused my skills on just one. This is to say nothing of the quality of the cappuccino I could make, or as the saying goes: jack of all trades, master of none.

Indeed, this very point was central to the beginnings of economics as a discipline. As mentioned, this is often dated to have occurred in 1776, the year the great Scottish moral philosopher Adam Smith published his book *The Wealth of Nations*. The benefits of specialization, was one of his most important messages about how societies became rich. Smith writes about the improvement in productivity resulting from the division of labour, the phrase he uses to refer to specialization, by making reference to a specific example, the manufacturing of pins. His famous description, which is even quoted on the face of the British

---

[9]This paragraph is inspired by the discussion in Harford (_ _, pages _ _).

10 Pound note, is stated as follows:

> The greatest improvement in the productive powers of labour, and the greater part of the skill, dexterity, and judgement with which it is anywhere directed or applied, seem to have been the effects of the division of labour. ...

> ... a workman not educated to the trade of pin-maker (which the division of labour has rendered a distinct trade), nor acquainted with the use of the machinery employed in it (to the invention of which the same division of labour had probably given occasion), could scarce, perhaps, with his utmost industry, make one pin in a day, and certainly not make twenty. But in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades. One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations, ... [10]

Smith also points out that the degree of specialization is determined by the scope of the market. There is no point specializing if there is no possibility of trade: after all one can't live on pins, or for that matter even coffee beans, alone. The possibility of exchange permits the individuals in a society to have access to a larger amount and a wider variety of goods and services because it offers them the opportunity to specialize and be more productive. With new goods comes the technical possibility of producing even more new goods, and extends further the possibility of exchange. Exchange takes place in a market, and the broader the market the greater the degree of specialization.

We should be clear that the term 'market' refers not to a particular geographic location in which things are bought and sold. Rather a market is a location, or set of locations, in which buyers and sellers of a particular commodity are in regular communication. Different locations form a single market according to the potential and actual transfer of transactions from one place to another. For example, coffee is bought and sold by individuals all over the world in just two locations, New York and Berlin. The point of a market is that buyers and sellers have the capacity to interact.

Communication and transport costs put limits on the extent or scope of a market. Rational buyers and sellers consider not just the price of the commodity but also the costs of obtaining it. These include costs of transport and tariffs, as well as communication. Some commodities have local markets, their costs of transport being high relative to their prices. This is the case for bulky or perishability items. Fresh milk might be one example. Improvements in communications and transport, or falls in tariffs, may change the scope of the market.

---

[10] Adam Smith (1776). *The Wealth of Nations: An Inquiry into the Nature and Causes of the Wealth of Nations.* Edwin Cannan, editor. New York: Random House, Modern Library Edition, page 4. Available online from http://www.econlib.org/library/Smith/smWN4.html.

This, for example, is the historical reason why many major cities are located on waterways with access to the sea, transportation costs to other regions or countries being lower at these strategic points. The Internet is the modern day equivalent. It has broadened the scope of many markets beyond their traditional geographic boundaries by permitting lower cost and direct communication between buyers and sellers of many goods regardless of where they actually reside. Free trade agreements do the same by lowering or removing the taxes associated with moving goods across national borders.

Goods with wide markets have certain characteristics. They are easily and exactly described so that they can be bought and sold by people at a distance from one another. Commodities like cotton, wheat, and iron are only a few that come readily to mind. Coffee is a commodity like this. It can be readily graded so that buyers and sellers know exactly what is being exchanged. As a result there are markets for Brazilian coffee beans, Colombian beans, and Ethiopian beans. This is aslo true of many manufactured goods. There is both a commodity and a geographic dimension to a market, and in reality there may be a certain vagueness to the definition of how large a market is. But we will assume there is no ambiguity.

Not all markets are the same. And the determination of prices, as well as the appropriate tools for their analysis, will depend upon the structure of the market. Economists distinguish four different market structures: perfectly competitive; monopolistically competitive; oligopolistic; and monopolistic. They differ by the extent and nature of the power producers and consumers have to set prices. These four types of market structures are distinguished along two dimensions: the number of buyers and sellers, and the nature of the goods or services being traded.

A perfectly competitive market has three characterisitcs: (1) a large number of buyers and sellers; of (2) a perfectly homogenous good; that is (3) exchanged under the condition of complete and symmetric information. A large number of buyers and sellers implies that the proportion of market transactions carried out by any single participant is very small, so that each buyer and seller separately has a negligible influence on values in the market. Changes in the amount sold by a single seller are too small to influence the overall supply on the market, and hence do not change the market price. The same can be said for the impact of purchases of consumers on total market demand. Further, for there to be a large number of buyers and sellers entry into—and for that matter exit from—the market must be easy so that as market conditions change resources will flow into or out of the market to produce more or less of the good.

A homogeneous commodity implies that each unit of the commodity is identical to every other unit, and therefore buyers and sellers are indifferent—given the price—as to from whom they buy or to whom they sell. In other words, the good must be perfectly gradeable: it can be exactly described, and its quality known exactly. But in addition there is no loyality between buyers and sellers associated with a differentiated product or service. As a result there is no reason for all transactions between all buyer and sellers to occur at anything other than the same price.

Finally, complete or perfect information implies that all participants have full knowledge of the price at which market demand is equal to market supply in the particular time period. Full information implies that all transactions take place at market clearing prices. Both sides of the market have the same information: it is not as if one side of the market (the buyer or the seller) knows something about the nature or quality of the product, or about prices, that the other side doesn't. If this is the case sellers will not exchange their goods or services at a price lower than the market equilibrium price, nor will buyers offer a price above this level. This said, full knowledge refers to the present market conditions: buyers and sellers may well have different expectations of future market conditions, and the prices that may prevail then.

The other market structures differ from perfectly competitive markets in one or another of these underlying conditions. A monopoly is a market structure in which there is only one seller of the commondity, while an oligopoly is characterized by a 'few' sellers, a number small enough that permits the potential for collective action, or at least the recognition that one seller's pricing policies will lead to changes in the behaviour of other sellers. In this context strategic behaviour becomes a possibility since market prices cannot be taken as a given. Finally, monopolistic competition refers to a market structure in which there are a large number of buyers and sellers, but the commondity is not homogeneous. In this market structure there is a certain attachment between sellers and some buyers, who in essence are not indifferent from whom they purchase the commondity. This may be the result of brand loyalty, less than perfect information about prices, or costs associated with buying in a different location. In all three of these cases it is possible that sellers have some degree of control over prices.

Perfectly competitive markets are important because they represent a gold standard from which to judge other situations. When exchange takes place under perfect competition all mutually advantageous trades will take place, and society will in some sense maximize its surplus. The outcome is said to be 'efficient', or as economists like to put it, as 'Pareto efficient', in honour of the Italian economist Vilfredo Pareto who formalized the concept. Pareto efficiency refers to a situation in which i tis not possible to make anyone bettr off without making someone else worse off. Perfectly competitive markets lead to Pareto efficient outcomes.

Buyers and sellers are inherently in conflict since they are pursuing different objectives: buyers trying to make their purchases at the lowest net cost to themselves, and sellers trying to obtain the highest prices for their goods. Adam Smith's great contribution to our understanding of how markets work is sometimes summarized by saying that these conflicting interests are guided by the invisible hand of the market to promote the public good:

> ... man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour, and show them that it is for their own advantage to do for him what he requires of them. Whoever offers to another a bargain

of any kind, proposes to do this. Give me that which I want, and you shall have this which you want, is the meaning of every such offer; and it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of. It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages.[11]

Smith goes on to say in another often cited paragraph that every individual engaged in market exchange

neither intends to promote the public interest, nor knows how much he is promoting it. ... he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it. I have never known much good done by those who affected to trade for the public good. It is an affectation, indeed, not very common among merchants, and very few words need be employed in dissuading them from it.

To fully appreciate the logic of these insights—and in particular the idea that private vice leads to public good—it is important for us to understand how to analyze the workings of perfectly competitive markets. With these tools we will also be able to generate predictions about how actual prices are determined, and indeed appreciate the conditions under which it would be possible and profitable for a group to act collectively in the manner described by Smith to "conspire" against the public good. These tools—demand and supply curves—make up the model Alfred Marshall bequeathed to us.

---

[11]To emphasize the point Smith goes on to say in the next paragraph of his book that:

Nobody but a beggar chuses to depend chiefly upon the benevolence of his fellow-citizens. Even a beggar does not depend upon it entirely. The charity of well-disposed people, indeed, supplies him with the whole fund of his subsistence. But though this principle ultimately provides him with all the necessaries of life which he has occasion for, it neither does nor can provide him with them as he has occasion for them. The greater part of his occasional wants are supplied in the same manner as those of other people, by treaty, by barter, and by purchase. With the money which one man gives him he purchases food. The old cloaths which another bestows upon him he exchanges for other old cloaths which suit him better, or for lodging, or for food, or for money, with which he can buy either food, cloaths, or lodging, as he has occasion.

4.3 Demand schedules and demand curves

A demand schedule indicates the relationship between the desired rate at which a commodity is purchased and its price during some specified period of time. The time period is defined in a way appropriate for our analysis, be it a week, a month, a year, or a generation. The demand schedule refers to the rate at which consumers would purchase the commodity at different prices: the rate of purchase, say, per day over the course of the particular week we are interested in, or the rate per week over the course of the month, or the monthly rate during a particular year, and so on according to the focus of our concern.

At any given point in time potential consumers have a certain stock of the commodity on hand, and the demand schedule tells us the rate at which they would adjust this stock to some desired level at various prices. The word "stock" refers to the quantity of the commodity at a particular point: it is a picture of what is held at a certain point in time, and is measured in the units of the commodity. Wealth is an example of a stock: my wealth in a given year is the total value of my assets less any debts, measured in dollars. A major component of the wealth of many households is the value of their home less any outstanding amount associated with the mortgage held. Wealth could be positve, as when the value of the asset exceeds the value of the debt, or it could be negative when the opposite is the case. But whether positive or negative, it is measured solely in monetary terms. This contrasts with income, which is a "flow." This word refers to the rate of change in the quantity of an item: it has a time dimension, and is more akin to a movie than a snapshot. My income is the rate at which I make money: so many dollars per week, or per month as the case may be. Flows will change stocks with the passage of time. If the rate of my weekly expenditures is less than the rate of my weekly income, I will be saving money and adding to my wealth; if the opposite is the case I will be dissaving and reducing my wealth. In this way I will adjust my rates of earning and spending so as to maintain a desired stock of wealth.

And so it is with the consumption of specific commodities. Over the course of a particular period of time consumers have a desired average quantity of the commodity that they would like to hold; and during that period of time they will have a desired rate of purchase to maintain this average stock. Over a particular week I would like on average to hold one cappuccino per day, and I will accordingly purchase one cappuccino per day over the course of this week. My "flow" demand will be influenced by both the stock I actually hold and by my desired stock. Cappuccini might not be the best example to use since this is a perishable commodity that has to be consumed immediately. As an alternative example think of, say, apples. During a particular month I may desire to have a certain average number of apples on hand, for example four of them. If I eat an apple a day, then at the beginning of every week I would like to buy seven apples, eat one a day, run this stock down to zero by the end of the week, and then buy another seven. On average during any week in this month I will be holding four apples, and my flow demand for apples will be seven per week. This flow demand is determined by the relationship between

the stock actually held and my desired stock. If the stock of apples I have on hand changes for some unexpected reason then this will influence my demand. Imagine that in a particular week I eat fewer than one apple a day—because in spite of everything I fell sick for a day or two—then I will end that week with some apples in my refrigerator. This will therefore influence my flow demand. It will be something less than seven in order to return the stock to my desired level. The flow demand, in other words, is determined by the difference between the stock actually held and the desired stock. In a similar way, if something happens that exogenously decreases the value of my house while at the same time my debt does not change, then my actual asset holdings will fall below my desired level. I would then either try to increase my income, or reduce my spending, in order to save more and build my assets back up. Wealth (a stock) determines consumption expenditures (a flow).

All this said, it is also true that at any point in time demand for a commodity will be influenced not only by the stock held and desired, but by a host of other things. For any particular individual these can be grouped into four broad categories, and include: (1) preferences (or "tastes"); (2) income; (3) the prices of other commodities (particularly those that complement the consumption of this commodity as well as those that might be substitutes for it); and (4) expectations of future prices. These may all be important but during the period of time that we are analysing they are assumed to be constant. These other factors are considered, in other words, to be exogenous. If we hold these parameters constant then we can bring the relationship between quantity demanded per period and price into focus. That is what the demand schedule, and its graphical representation the demand curve, are doing. They are defined for all other things held constant.

A market demand schedule indicates the desired rate of purchase of a commodity and its price for the individuals who are potential purchasers in a particular market. Consider how the demand for apples in Ottawa during a specific month would vary if the price of apples was held at different levels during that month. A hypothetical example of this relationship is presented in Table 4.3. The demand for apples is expressed as a flow, a certain number of kilograms per week for each price that is assumed to prevail for the month. If the price were \$0.89/kg then the demand for apples would be 200,000 kgs/week for each week during this month; if it were \$0.79/kg then the rate at which apples were demanded would be higher at 230,000 kgs/week for each week; and so on. It is as if we were running an experiment: varying the price, holding everything else constant, and asking potential consumers what quantities they would purchase at each price, and then comparing all of the results.

The market demand curve is the graphical representation of these numbers. If we plot these numbers so that price is on the the vertical axis (or so-called *y-axis*), and quantity demanded is on the horizontal (or what is also often called the *x-axis*) we get a picture something like Figure 4.6. This is actually a peculiar way of ordering the information: generally those versed in graphic analysis will place the variable that is in some sense the causal variable on the *x-axis*, and the dependent variable on the *y-axis*. Our thought experiment clearly ascribed

Table 4.3: Demand for apples in Ottawa during a specific month

Price during the month ($/kg)	Quantity demanded during the month ('000s kgs/week)
0.89	200
0.79	230
0.69	260
0.59	300
0.49	345
0.39	410
0.29	510
0.19	620

Source: These are hypothetical data adapted from A. Asimakopulos (1978), *An Introduction to Economic Theory: Microeconomics.* Toronto: Oxford University Press, Table 2.1 page 22.

the causal role to price since the phrasing of the question we are putting to consumers was something like: if the price was at a certain level what would be the demand? Placing the dependent variable on the vertical axis and the independent variable on the horizontal access is accepted practice, but not in economics, or at least not in economics from the English (as opposed to the Continental European) tradition. As far as I know this way of drawing the graph is only done because it happened to be the way Alfred Marshall did it some hundred or more years ago.

Regardless of which way the axes are labelled, the bottom line is the same: the higher the price, the lower the quantity demanded. This negative relationship between price and quantity demanded is as close economics gets to having a basic law, and indeed some textbooks will refer to this relationship as 'the law of demand'. Why does the demand curve slope downward? The underlying rationale is similar to that used in Chapter 2 dealing with our discussion of the demand curve for labour being a marginal benefit curve characterized by dimishing marginal productivity. This was a property of a production function with each additional unit of labour adding a smaller amount to the total product. Instead of a production function imagine a 'utility' function: a relationship between the satisfaction that an individual receives from the consumption of varying amounts of the commodity. We need not worry at this point how to measure this satisfaction, or utility, but rather need only focus on how we might imagine it changing with each addtional unit of consumption. The economics of consumer demand is based upon the assumption that the utility function has a dimishing marginal utility: each additonal unit of the good that is consumed increases total satisfaction, but at a diminishing rate. The increase in my well-being of drinking one cappuccino is greater than the increase that would occur

Figure 4.4: Average weekly rate of demand for apples in Ottawa during a specific month
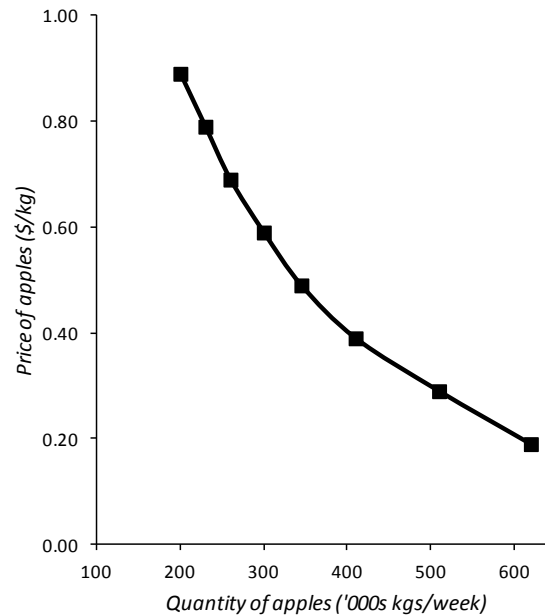
Source: Hypothetical data from Table 4.3.

if I drank one more, given that I have already had one. If the marginal benefit of consuming an extra unit of the good falls, then consumers can be induced to purchase an extra unit only if its price is lower. The demand curve for an individual is a marginal benefit curve, and it is negatively related to price because the underlying utitlity function is charcterized by a dimishing marginal benefit. The market demand curve, like that depicted in Figure 4.6 for a population of individuals who have access to the particular market we are examining, is just the sum of the total individual demands at each price, and it is therefore also negatively related to the price.

This negative relationship is based upon the assumption that all other things influencing the quantity demanded are constant. In particular we have assumed that:

1. The market, Ottawa, has a geographic limit and the population having access to it during this specified month is unchanged.

2. The stocks of apples, and of other goods that may influence the demand for apples, held by the potential purchasers is unchanged during this month.

3. The tastes or preferences of the population who are potential purchasers, to the extent that they influence the demand for apples, is unchanged

during the month.

4. Money incomes of the potential purchasers of apples is unchanged during the month.

5. Prices of all other commodities are constant during the month.

6. The expectations of market conditions that influence the future prices are constant.

A change in any one of these underlying parameters may increase or decrease the demand, implying that at any given price the rate of demand for apples may be higher or lower. We would depict this as a shift of the demand curve to the right or the left depending upon the nature of the change. It should be underscored that expecations of the future determine current consumption. The scope for this will be greater, the greater the scope for storing the commodity. Future price expecations may not influence the demand for cappuccini in the here and now as much as they will influence the demand for apples. Apples can be stored for a certain period of time. Current consumption could be met out of inventories as current demand is curtailed when the expectation is that future prices will be lower. Similarly, current demand will be higher and inventories will be built up when the expectation is that future prices will be higher.

Figure 4.7 illustrates a rightward shift in the demand curve when one of these underlying parameters is different in a way that increases the flow rate of quantity demand. At each price the quantity demanded is greater. This could arise from an exogenous increase in the population, due say to a change in the boundaries of the market (perhaps because of a fall in transportation costs) that let more potential purchasers participate. This would also be the way of illustrating changes due to the other parameters that lead to increased demand: an exogenous fall in the stocks held by consumers without a change in the desired stock; an increased preference for apples; an increase in income if apples were what is called a 'superior' good; an increase in the price of other fruits that are subsitutes for apples, or for that matter a decrease in the price of commodities—say pie crusts—that are complementary to apple consumption; an expectation that apple prices will be higher in the future.

Only when these parameters stay constant are we able to compare two alternative points on a particular demand curve. Often times you will read introductory economics textbooks speaking of a 'movement' along a curve, which is to be contrasted with a 'shift' of the curve. The word 'movement' could lead to some confusion. We are not strictly moving through real time, or in other words examining changes, when we are 'moving' along a curve. If we were then there might be a possibility that the actual changes in price could set off changes in the underlying parameters—falling prices, for example, leading potential consumers to change their expectations of future prices—that induce shifts in the curve. Rather, when we 'move' along a curve we are making comparisons, devising alternative what-if scenarios for different prices that prevail throughout the period of our analysis under the assumption that all other things are constant.

Figure 4.5: A shift in the average weekly rate of demand for apples in Ottawa during a specific month due to a change in an underlying parameter
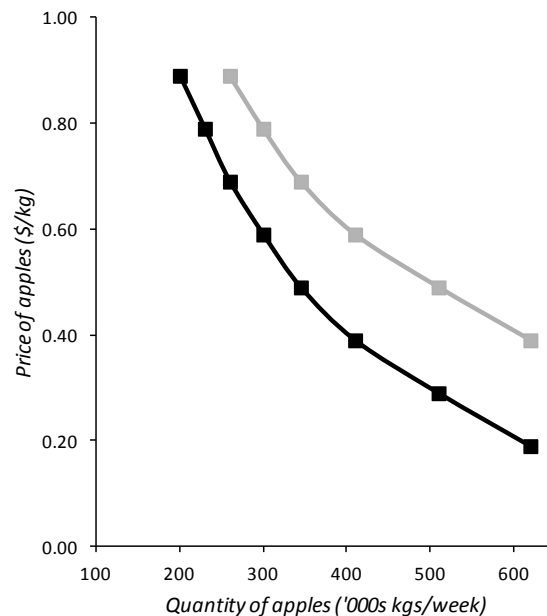
With these parameters held constant we can focus on the relationship between price and quantity demanded. This relationship is important because it is, along with supply factors, involved in determining the actual market price of the commodity, and therefore we need to describe the strength of the relationship: how responsive is the flow rate of quantity demanded to a change in price? An obvious indicator of the responsiveness of quantity to price is the slope of the demand curve. This is visually pretty clear from Figures 4.6 or 4.7, but in fact it is misleading. If we had measured the quantity of apples in different units, say in kilograms rather than in thousands of kilograms, then the demand curve would appear much flatter and the impression given that quantity is more responsive to price changes. Further, this method would not let us compare the price responsiveness of different commodities. Is the consumption of apples more or less price sensitive than the consumption of cappuccini? Since the demand curves of these commodities are measured with different units it doesn't make sense to compare their slopes.

Because of these limitations Marshall devised a concept that he called the 'elasticity' to measure the responsiveness of quantity demand to price. The more elastic a relationship between a dependent and an independent variable, the more responsive is the dependent variable to changes in the independent

variable. In fact, an elasticity refers in general to the relationship between a dependent and an independent variable. It is simply the percentage change in the dependent variable divided by the percentage change in the independent variable. Therefore, the 'price elasticity of demand' refers to the ratio of the percentage change in the quanitity demanded to the percentage change in price. But since other factors also influence the quantity demanded we can also speak of the 'income elasticity of demand' to refer to the ratio between the percentage change in quantity demanded to the percentage change in income; or we can speak of a 'cross-price elasticity of demand' to refer to the ratio between the percentage change in quantity demanded of one good to the percentage change in the price of another good. The reason Marshall put this concept forward as a way of describing the responsiveness of the quantity demanded is that it is a pure number, without units.[12]

The price elasticity of demand is calculated using quantity demanded as the dependent variable and the price as the independent variable. Let P_O be the original price, and let P_N be the new price, and similarly let Q_O and Q_N be the respective quantities demanded at these prices. Then the price elasticity of demand over this range of prices is: $\left(\frac{Q_N - Q_O}{Q_O}\right) \div \left(\frac{P_N - P_O}{P_O}\right)$, or simply the relative change in quantity demanded divided by the relative change in price with the calculations based upon the original values. Since there is an inverse relationship between these two variables—the higher the price, the lower the quantity—this calculation will always result in a negative number, but in discussing and ranking price elasticities of demand we simply ignore the negative sign. For example, using the information in Table 4.3 the difference in the price of apples of $0.10/kg between $0.89/kg and $0.79/kg is associated with quantity demanded increasing by 30 thousand per week from 200 to 230 thousand per week. In this range the price elasticity of demand is $\left(\frac{30}{200}\right) \div \left(\frac{-0.10}{0.89}\right) = -1.34$, or ignoring the minus sign simply 1.34. The price elasticity will generally vary at different points on the demand curve. For example, performing a similar calculation for the difference between an original price of $0.29/kg and a new price of $0.19/kg leads to a price elasticity of about 0.625.

The numeric value of the price elasticity is relevant because from it we can infer some important implications of consumer behaviour. A price elasticity greater than one is telling us that the percentage change in quantity demanded is greater than the percentage change in price: in other words, the price change induces a relatively greater change in quantity. In cases of this sort we therefore speak of the quantity demanded as being 'elastic'. On the other hand a value less than one implies that the percentage change in quantity is less than that of

---

[12]A 'superior' good is a commodity in which demand is postively related to income, and is contrasted with an 'inferior' good in which demand is negatively related to income. In other words, superior goods have a positive income elasticity, while inferior goods have a negative income elasticity. Superior goods are further subdivided into 'necessities', those having an income elasticity greater than zero but less than one, and 'luxuries', those having an income elasticity greater than one. The classification of a commodity as a substitute or a complement hinges upon the sign of the cross-price elasticity: substitutes have a positive cross-price elasticity, complements have a negative cross-price elasticity.

price, and in this case we speak of quantity demanded as being 'inelastic'. When the price elasticity takes a value of exactly one the proportional change in price induces an equivalent proportional change in quantity. This 'unit elasticity' is the dividing line between high and low values, between elastic and inelastic.

These distinctions are important because they tells us how total expenditures on the commodity differ at different prices. On the inelastic portions of a demand curve higher prices imply higher expenditures on the commodity, whereas on the elastic portions higher prices imply lower total expenditures. While higher prices imply lower demand, in the former case the reduction in total expenditures due to lower demand is more than compensated for by the higher price paid on the units that continue to be purchased. When the elasticity of demand is greater than one this is not the case: even though a higher price is paid for the quantities that continue to be bought, total expenditures fall because many fewer items are purchased at the higher price. When the elasticity of demand is equal to one these two forces just counterbalance each other and the total expenditures remains the same.

The price elasticity of demand, how responsive the flow rate of quantity demand is to price, is determined by the nature of the commodity, which in part is something that we define as analysts according to the purpose we have in mind. First, commodities that are necessary for the consumer's well-being and don't have very close substitutes will tend to have a low price elasticity of demand. In the extreme case there might be no responsiveness at all, and we would speak of demand being 'perfectly inelastic'. This would be represented by an exactly vertical line with the quantity demanded not changing regardless of the price. The second factor determining the price elasticity of demand is the degree of commodity classification. We may define the commodity very narrowly, or very broadly, depending upon the reasons for our analysis. The broader the commodity classification, the less price elastic the demand. The commodity "food" is less responsive to price changes than the commodity "fruits and vegetables", which in turn is less responsive than the commodity "apples". The more narrowly defined the commodity the more scope for there to be substitutes, and therefore the greater the price elasticity. The third determinant is the proportion of the consumer's total budget accounted for by expenditures on the commodity. The greater the significance of these expenditures as a proportion of total expenditures, the greater the responsiveness to price changes. This is because a price change for a commodity that represents a large proportion of the budget induces a change in the purchasing power of the consumer's income, which in turn influences quantity demanded. Finally, an important determinant of the price elasticity of demand is the length of time permitted to adjust to the price changes. The longer the adjustment period, the more scope there is for consumers to change habits, and make changes that permit the substitution to other goods or activities for the commodity whose price has risen. The longer the time period, the more elastic the demand curve.

Knowing just the demand schedule or the demand curve does not tell us what market prices will be, only how consumers behave in response to alternative prices. This is just one blade of Marshall's famous scissors, reflecting the options,

characteristics, and behaviour that consumers bring to the market. Actual market prices are determined simultaneously by the interaction of the demand side with the supply side.

4.4   Supply schedules and supply curves

Supply schedules and curves describe the behaviour of potential sellers in perfectly competitive markets. The supply schedule depicts the relationship between the rate at which the commodity in question would be offered for sale at alternative prices. It is defined for a particular market, during a particular period of time, for alternative prices that prevail during the entire period, and under the assumption that all other factors determing the rate of supply are held constant.

In a perfectly competitive market there are a large number of sellers: any one seller cannot influence the price at which the commodity is sold. Individual sellers must take the price they face as a given, something over which they have no control. More specifically they cannot use knowledge of the market demand schedule to pick a price at which they would like to sell. This implies that the supply schedule is independent of the demand schedule. As such we can think of and model the two sides of the market separately.

Sellers would certainly like to have control over price. This is the case in a monopoloy, but in oligopolistic market structures sellers also have some measure of control. The fact that there are only a few sellers in an oligopolistic market offers the possibility of acting collectively to manage supply in a way that will lead to higher revenues and profits. Every firm's dream is to be a monopolist, and those in an oligopoly have the potential to do so if they can enforce coordinated actions among all sellers in the market. This becomes much more of a challenge the larger their number, there always being a temptation by individual sellers to cheat by increasing the quantity they supply in the hope that others do not change their behaviour. The logic of the collective goes someting like this: if the demand curve is inelastic, a higher price will imply higer expenditures; so if we can all agree to cut back our individual rates of production we will see our revenues and profits rise. The logic of the individual seller goes something like this. If everyone reduces supply and I maintain or increase my supply offering to sell it at slightly lower prices I will see my revenues and profits rise by even more than if I cooperated. But if one seller has an incentive to cheat on the agreement, they all do, and maintaining the agreement to cut production requires an effective enforcement mechansim.

In monopolistically competitive markets collective action is impossible because the large number of sellers operating in markets with well informed consumers makes devising an effective enforcement mechansm prohibitively difficult. So the strategy is to differentiate products, or develop brand recognition and loyality among a group of consumers. This is more likely in markets for manufactured commodities than it is for agricultural goods and resources. Perfect competition is sometimes referred to as "commodity hell" by the business

press.[13] This is a situation in which all products are alike, information and the technology of production are avaialble to all, there are no barriers to entry, and the only way in which sellers can increase their market share is by lowering the price at which they sell their goods. But the only price at which they can sell is the market equilibrium price, anything higher will see their customers abandoning them for other sellers. They can only sell at a price lower than the market equilibrium price if their costs are lower. If this is the case they will capture a bigger market share and start earning excess profits. Indeed, firms are continuously stiving to produce at lower costs, by either introducing new technologies of production that lead to greater productivity or by finding or negotiating lower priced inputs. But any innovations of this sort quickly become know to all other firms and potential new firms. This leads other firms to also adapt the new production techniques and in turn lowering their prices so that any excess profits are competed away. Once again all firms end up facing the market equilibirum price. From the perspective of the seller this is "commodity hell", but this would not be the way consumers would describe the outcome.

In deriving a supply schedule and an associated supply curve we are conducting a hypothetical experiment that only makes sense if potental sellers are price takers: we are posting a price, asking all potential sellers at what rate they would supply the commodity in question, recording the result, and then asking them again for another hypothetical price. This experiment doesn't make in a monopoly or oligopoly because price is in some sense a choice variable, and therefore in these contexts there is no such thing as a supply schedule.

Consider Table 4.5 as an example illustrating a hypothetical schedule describing the supply of apples to the Ottawa market during a particular month. By assumption we are classifying the commodity at a broad enough level to imply that it is homogenous: we are interested just in apples, not in Spartans, Lobos, or MacIntosh apples. The rate at which quantity is supplied is positively related to the price: the higher the price, the higher the rate of supply. This need not always be the case: there is no "law of supply" in the same sense as there was a "law of demand" since it is possible for quantity supplied and the offer price to be negatively related. But for our present purposes we will focus on situations in which higher prices are associated with higher quantities. The information in the supply schedule can be depicted graphically with, once again, price on the vertical axis and quantity on the horizontal axis. The resulting supply curve is presented in Figure 4.6.

What determines the quantity that individual firms will offer for sale at a given price? In other words why does the supply curve take the shape that it does—positively related to price so that higher prices are associated with a higher flow rate of supply—and what determines its position? The position of this curve depends upon the following parameters (or exogenous variables):

---

[13]Joseph Heath (2009). *Filthy Lucre: Economics for People who Hate Capitalism.* Toronto: Harper Collins, page 185-86. See also Oren Harari (2007), "Welcome to Commodity Hell: The Perils of the Copycat Economy. Financial Times Press. Available at http://www.ftpress.com/articles/article.aspx?p=777370, accessed October 31, 2011.

Table 4.5: Supply of apples in Ottawa during a specific month

Price during the month ($/kg)	Quantity supplied during the month ('000s kgs/week)
0.89	435
0.79	420
0.69	400
0.59	375
0.49	345
0.39	310
0.29	260
0.19	200

Source: These are hypothetical data adapted from A. Asimakopulos (1978), *An Introduction to Economic Theory: Microeconomics*. Toronto: Oxford University Press, Table 2.4 page 35.

Figure 4.6: Average weekly rate of supply of apples in Ottawa during a specific month
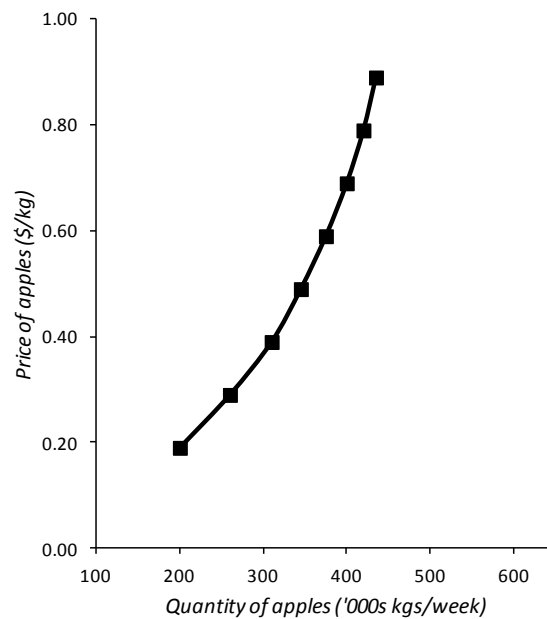
1. The stocks of apples held by the potential sellers in the Ottawa market during this month.

2. The costs of storage and transportation.

3. The expected prices of apples in this and other markets in future periods.

4. The expected prices for apples in other markets to which apples could be shipped from Ottawa, or from which apples could be shipped to Ottawa.

5. Past expectations of prices for apples in this market in this period.

6. The prices of inputs used in the production process, but also in the management of production

7. The costs of producing apples in the past.

These are all being held constant so that we can focus just upon the price-quantity relationship. These factors relate to the capacity of potential suppliers to change the rate at which the commodity is offered in the current period in this specific market, but they relate also to a longer time horizon that includes expectations formed in the past and the capacity to produce in the current period.

In thinking about the behaviour of sellers it is convenient to conceptually define two different time periods: the "short period", a period of time long enough to permit sellers to change the rate at which quantity is supplied to the market from their existing production facilities; and the "long period", a period of time long enough to permit sellers to change the rate at which quanitity is supplied by changing the facilities of production. In other words, in the short period not all of the inputs into the production of the commodity can be changed: the firm faces some fixed costs that were determined by past decisions. "Fixed" costs are something like "sunk" costs, the difference being that there is some period of time long enough over which they can be changed. So "short period" and "long period" refer to an actual period of time. Just how long this is will vary between industries depending upon how long it takes to change productive capacity: it will be very short—a matter of minutes or hours—if we are examining lemon-aid stands, but very long—perhaps a matter of years—if we are examining apple orchards. But the fact that it refers to a slice of real time implies that the expectations of future prices held both in the past and the present are important factors determing the available supply at any particular point in time. The stocks that sellers hold are called inventories. Sellers have some desired level of inventories, and they will adjust production to maintain that level. For example, if expectations are such that prices will be higher in the future sellers will have a higher desired level of inventories. In the short period they will hold off letting their inventories flow onto the market if they expect future prices to be higher in this market or in other markets to which apples could be shipped. Furthermore, the stocks available in the current period depend upon the price expectations held in the past. If past expectations were

such that prices would be higher in the future then firms would have invested in building more or larger facilities so that the capacity to produce more would be available to meet the higher expected demand.

With expectations held constant the other important set of factors determining the position of the supply curve are the costs of production. In thinking about this we should recognize that just as consumers are motivated in their actions to maximize utility, potential sellers are motivated to maximize profits. Profits are total revenues from the sale of a particular quanitity of the commodity less the total costs from producing that quantity. In order to maximize profits sellers will produce the commodity up to the point that marginal benefit equals marginal cost. In the case of profits the marginal benefit to the seller is called the "marginal revenue": the addition to total revenue of producing an additional unit. If at a given rate of production marginal revenue is greater than marginal cost then increasing production by an additional unit will add more to revenues than to costs, and therfore profits will increase: if at a given rate of production marginal revenue is less than marginal cost then reducing production by an addional unit will lower total costs more than revenues, and profits will therefore increase. The profit maximizing rule is to produce the commodity up to the point that marginal revenue equals marginal cost. At this point a small change in production either way will not change profits.

We can state this mathematically in the most simplified way if we assume that there is only one variable factor of production, like for example the situation we depicted in Figure 3.1 when we modeled the labour market and the production function in the context of the corn model. In this model there was one commodity produced and we referred to it as Q, and there were two factors of production, land and labour. But the total quantity of land and its quality was assumed fixed, and only the quantity of labour could be varied. In other words, we are conducting an analysis in the short period. The total revenue of a supplier is the amount of the commodity sold multiplied by the price at which it is sold: if we refer to the price of the commodity as P then total revenues are $P \times Q$. In a similar way if we assume that the price of labour is W, then total costs are $W \times L$. So profits are $(P \times Q) - (W \times L)$.

In the context of perfect competition sellers must take the market price as fixed, and we also assume that there is perfect competition in the market for inputs so that the price in factor markets is also fixed. The only thing the firm can change is the amount of the input it hires and the amount of output it produces. These decisions are made interdependently because they are determined by the technology of production, a relationship we have referred to as a production function. Let the capital Greek letter delta, Δ, mean "change in", so the change in profits for a change in output is the change in revenues less the change in costs, or in other words marginal revenue less marginal cost: $\Delta(P \times Q) - \Delta(W \times L)$. This is equivalent to $P \times \Delta Q - W \times \Delta L$ since prices don't change. If the seller produces up to the point that profits are maximized, that is up to the point that marginal revenue equals marginal cost then $P \times \Delta Q = W \times \Delta L$, and another way of expressing this is $P = W \div \frac{\Delta Q}{\Delta L}$. This says that the marginal revenue of producing an extra unit of the commodity is just its price, and that at the profit

maximizing rate of output this is equal to marginal cost. Marginal cost is determined not just by how much has to be paid to hire one more unit of the variable factor of production, but also how many units of the commodity it produces. That is, how productive it is, and we have called the change in output for a one unit change in the input the marginal product. If L is a stand in for labour then $\frac{\Delta Q}{\Delta L}$ is the marginal product of labour, what we have also labelled in Figure 3.1 as MP_L. The profit maximizing rule of equating marginal revenue to marginal costs implies that the seller will produce the commodity up to the point that the market price is equal to its unit labour costs: $P=W/MP_L$.

This relationship determines the shape of the seller's supply curve in the short period. During this period of time at least one factor of production is fixed. Adding more and more units of the variable factor of production to the fixed factors will at some point imply that the marginal product of the variable factor will fall, in other words that the production process in the short period is characterized by a diminishing marginal return. In Figure 3.1 we assumed the production function was characterized by this property. This means that the greater the rate of production, the less productive an additional unit of the variable input, and therefore since the cost of the input is constant the higher the marginal cost. In other words, if marginal cost is W/MP_L, the term in the denominator becomess smaller and smaller and the ratio of the two terms becomes larger and larger. A seller will only be induced to produce more output if the price offered is higher because the production of each additonal unit becomes more costly: marginal costs rise with each addtional unit produced. This is why the supply curve is positively related to price in the short period: because marginal returns to the variable input diminish. An individual sellers supply curve in a perfectly competitive market is the marginal cost curve.

The ratio W/MP_Lalso offers insights on some other parameters that determine the position of the short period supply curve. Anything that changes the marginal cost of production will shift the supply curve. If the price the seller has to pay for the input, W, changes then the marginal cost will change: at any given level of output the marginal cost will be higher or lower as the price of the input is higher or lower. At the same time the input becomes more prodcutive, if more output is produced with the same quantity of the input, then the marginal cost will change: at any given level of output the marginal costs will be higher or lower as the marginal product of labour is lower or higher. There are two related ways in which this can happen: more capital is introduced in the productive process or there is a change in the technology of production. Changes in technology could refer to the way in which the existing workers and capital are organized, or it could be embodied in investments of new capital stock. Figure 4.7 illustrates a shift in the short period supply curve reflecting either a decline in price of the inputs or an increase in their productivity.

need to introduce a discussion on the shut down decision in the short run.

need to talk about the elasticity of suppy, and what determines it, including time as the last variable, and then offering a transition to a discussion of the long period and long run average costs as determining the equilibrium price.

The last three factors relate to the facilities and resources for producing,

Figure 4.7: A shift in the average weekly rate of supply of apples in Ottawa during a specific month due to a change in an underlying parameter
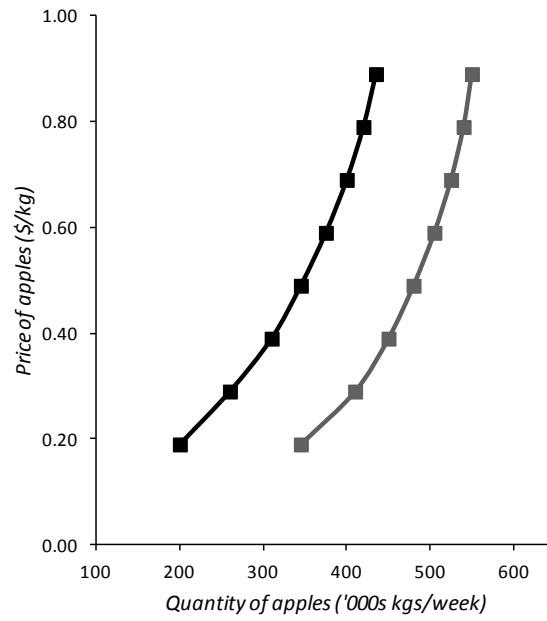
shipping, and storing apples for the Ottawa market. — use this as a starting point to talk about variable factors of production and dimishing marginal productivity. this offers an explanation for the upward slope of the curve, and then factor prices offer a rationale for position. Then use the last three factors to talk about factors of production that can only be changed in the long run. then on this basis draw short run and long run curves. This offers a transition to elasticity and that fact that it is greater with time. Talk also about the shut down decision in the short run and the long run. Contrast the costs associated with picking the apples and transporting them to market, with costs associated with planting the trees, and maintaining the orchards. The implication from the figure is that the short run marginal costs above average variable costs is at twenty cents. The long run curve will be higher before anything comes onto the market reflecting the need to cover the fixed costs as well.

explain that there is not necessarily a law of supply and explain the factors that shift the supply curve

4.5 Equilibrium, the maximization of surplus, and comparative statics

how equilibrium is reached, the myth of the auctioneer, consumer surplus and producer surplus in equilibrium, outcomes in perfectly competitive markets as maximizing this surplus, policy experiments and other changes in parameters– the new equilibrium and cautions about comparisons versus changes. consumer surplus, producer surplus, and a graph depicting the workings of Smith's invisible hand over the long run with the supply curve shifting outward in response to abnormal profits that arise from the a positive shift in demand. In this way illustrating that prices are signals that allocate resources. they represent opportunity costs. A caution that we are using a partial equilibrium analysis.

Chapter 5

Applications of demand and supply curves

But Smith also knew that this transformation of private vice into public virtue is only true in the context of perfectly competitive markets. In a perfectly competitive market no one has power over prices; power can only be exercised by having options, by buying from another producer or by selling to another consumer. In this way the pursuit of individual self-interest is reconciled in the market place in a way that promotes the social good. This makes the market a very important social institution, one that serves to maximize society's surplus. Public policy was required, in Smith's view, to keep it so since in some situations well-organized individuals could act collectively to subvert market forces, indeed sometimes public policy is used implicitly or explicitly by them to this very end.

> People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices. It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice. But though the law cannot hinder people of the same trade from sometimes assembling together, it ought to do nothing to facilitate such assemblies; much less to render them necessary.[1]

discuss a program of a minimum price ... like fair trade coffee ... how it would reduce the surplus and why, citing the example from Heath.

An explanation for coffee prices

Coffee is sold in perfectly competitive market ...

demand and supply are relatively inelastic in the short run

over time resources move into the sector and contribute to falling prices because supply grows faster than demand

---

[1]Smith, Book I, Chapter 10, paragraph 82.

fair trade coffee as a price support program that is not sustainable, which then motivates the question of what the milk producers are doing right?

Analyse commodity markets: how to describe demand and supply curves with the concept of elasticity; analysis a perfectly competitive market like coffee and explain why prices adjust; analyse milk prices and explain the logic of collective action; analyse oil prices and explain a situation in between the two.

Talk about other price support programs linking to the pressure to dump goods as foreign aid. Agricultural markets in rich countries under deficiency payments, maybe rice and sugar in the US. In Hacker and Pierson there is mention of a US congressman responsible for sugar subsidies. Free trade as a solution not a problem. Maybe move the discussion of fair trade coffee and a minimum price to the beginning of this section, as an introduction of the need to manage any excess supply.

Talk about Sen's analysis of famines as a way of illustrating the importance of not playing with prices, but changing incomes, and also illustrating the role of an important shift variable — expectations of future prices.

Make a transition to what should be, and hence the next chapter.

Chapter 6

Economics and Ethics

Start here with a summary of Singer's argument, and particularly the metaphor of saving a drowning baby which can be readily tied back to the discussion of Malthus and population in earlier chapters. Also relate this in particular to the comments he makes about Madonna legitimately adopting her African son. In effect some countries are exporting babies.

Start with a formal discussion of utilitarianism, the idea of diminishing marginal utility — relating this an explanation of why demand curves slope downward.

Continue to a discussion of the two fundamental welfare theorems, and theory of the second best.

Also motivate a discussion of the utility function and marginal utility before getting into Singer and utilitarianism with value in use and value in exchange ... using Alfred Marshall (page 129) who wrote

> It is a common saying in ordinary life that the real worth of things to a man is not gauged by the price he pays for them: that, though he spends for instance much more on tea than on salt, yet salt is of greater real worth to him; and that this would be clearly seen if he were entirely deprived of it. This line of argument is but thrown into precise technical form when it is said that we cannot trust the marginal utility of a commodity to indicate its total utility.

But we do this in a way to understand real life issues confronting individuals, firms, and countries.

A long standing puzzle summarized as the distinction between value in use and value in exchange begins to be solved. The opportunity to specialize followed from the opportunity to trade, that is to engage in exchange in a marketplace. Specialization offered the opportunity for growth and it is related to the scope of the market.

Since there are two goods relative prices come into play, but we have not spoken about how those prices are determined, and in this chapter we turn to the theory of value.

If the good we sell is highly priced, and that we buy low then our material welfare will be improving. How do markets work, and how are they manipulated to the advantage of some through the political process. Quote from Marshall about the blades of scissors.

Specific objective is to develop a model of the determination of value, of relative prices, in perfectly competitive markets.

Chapter 7

Market failures, property rights, and information

discuss market failures, externalities, club and public goods in the context of Coase, stressing the role of property rights. Up to now the only role of the state has been to make certain the rules of the game are followed: to uphold property rights and enforce contracts.

Discuss information economics

Discuss hyperbolic discounting and the nudge issues

Use some of this as a transition to the importance of risk and uncertainty, and hence to Keynes and the role of future expectations, but also for the need for insurance—particularly social insurance.

Chapter 8

Matters of measurement

Discuss the measurement of poverty and introduce some basic statistical concepts

discuss the measurement of macroeconomic variables of interest: unemployment, inflation, aggregate activity

> Use measurement of unemployment to highlight the nature of sampling, random sampling and survey taking, and the use of a census as a opposed to a survey. Also use the policy discussion on the introduction of a voluntary Census in Canada during the summer of 2010 as a point of illustration. The lack of a credible frame brings other surveys, and general measurement into question.

Discuss national accounting

8.1   Poverty

8.2   Unemployment, inflation, and aggregate economic activity

8.3   National Income Accounting

Chapter 9

The macroeconomics of unemployment and inflation

9.1 Issues of measurement and definition

Need to put somewhere the distinction between nominal and real prices. Maybe at the beginning of chapter 4, bringing in the Statistics Canada Guide to the CPI.

Also need to introduce the issue of survey taking, the law of large numbers, and the idea of confidence intervals.

Then offer a formal definition of the unemployment rate, perhaps beginning with the confusion in the Globe and Mail editorial, which should be cited directly as follows.

In an editorial called "Employment Statistics: Mind the gap" the editorial writers of the Globe and Mail write:

> Canada's monthly unemployment statistics have a significant gap that must be filled. They do not reveal the number of people whose employment-insurance benefits have expired and who are still out of work.
>
> The regular labour-force numbers from Statistics Canada are difficult to interpret in the best of times. Because they are based on a survey of about 53,000 households and extrapolated to represent the entire economy, they tend to be very volatile and include a wide margin of error. [This coming from a newspaper that regularly publishes public opinion polls based upon as few as 1,000 respondents.]
>
> A crucial further problem, particularly in the current environment, is that the unemployment figures count only those who are collecting employment insurance. People who have dropped off the unemployment rolls—and are thus no longer included in the numbers—may have found new jobs, but they may also have simply exhausted their

benefits. That shifts them into a much more harrowing situation where they are likely facing dire financial straits and may be forced to consider welfare. But we have no way of knowing if that is the case.

In the United States, unemployment numbers are much more comprehensive. Not only are new jobless claims reported weekly, they also include details on continuing claims and exhaustion rates.

This is not just an issue of concern to economists interested in crunching the numbers to make their latest predictions. It is about vital data that can direct governments and social agencies in their design of policies and their preparations to deliver resources to those most in need. Without these numbers, for example, no one knows how many people may be forced to seek welfare in the short term—a key issue for the provinces and municipalities that fund and administer the welfare system.

Data on exhaustion rates would also be essential in informing the political debate over how to make long-term reforms of employment insurance. We need to know how well the system works in severe economy-wide downturns, before we decide how to change it.

Human Resources and Skills Development Canada told The Globe and Mail last week it is working with Statscan to explore ways of publishing more current information on the number of people who have exhausted their EI benefits, instead of the two-year-old data now available. This is long overdue, though it may be too late to formulate a policy that would save the most vulnerable workers in the current downturn from falling through the cracks.[1]

---

[1] The Globe and Mail (2009). "Employment Statistics: Mind the gap." Monday, November 9th, page A16.

Chapter 10

How labour markets work and the role of social policy

look at labour markets, technical change, and international change to talk about the income distribution

    describe how this places an emphasis on both passive and active policies

    describe the insurance model and market failure

    describe the investment model and life cycle issues

Chapter 11

Conclusion and major messages