



Introduction to economic analysis

Alberto Bisin

Dept. of Economics

NYU

January 12, 2011

Contents

Preface	ix
1 Introduction	1
1.1 Theory, Models, and Empirical Analysis	4
1.2 Debates	6
1.3 Predictions	7
1.4 Empirical analysis	11
1.5 Some successful results	13
1.6 Fun readings	14
1.7 Useful references	15
2 Rational choice	17
2.1 Preferences	18
2.2 Rationality	19
2.3 Utility representation	21
2.4 Choice	24
2.5 Social choice	25
2.6 Classic references	26
2.7 Useful references	27
2.8 Problems	28
3 Consumer choice	35
3.1 The Budget constraint	36
3.2 Indifference curves	38
3.2.1 Monotonicity	41
3.2.2 Convexity	41
3.3 Types of preferences	43
3.3.1 Perfect substitutes	45

3.3.2	Perfect complements	45
3.4	Consumer problem	46
3.5	Demand Functions	49
3.6	Problems	50
4	Equilibrium and efficiency	61
4.1	The Economy	62
4.1.1	Notation and Definitions	62
4.2	Pareto efficiency	63
4.2.1	The Social Planner problem	64
4.3	Competitive equilibrium	66
4.4	Welfare Economics	67
4.4.1	The Edgeworth box	68
4.5	Externalities	70
4.6	Problems	70
5	Equilibrium in the labor market	79
5.1	Competitive labor markets	79
5.2	Search frictions	80
5.2.1	Pareto inefficiency of search equilibrium	82
5.3	Empirical implications of search equilibrium	82
5.4	Problems	83
6	Time and uncertainty	85
6.1	Choice over time; a.k.a. Saving	86
6.1.1	Intertemporal preferences for consumption	86
6.1.2	The intertemporal budget constraint	87
6.1.3	Choice over time, for given r	88
6.1.4	Determinants of savings	90
6.1.5	The interest rate r	90
6.1.6	Inflation and Monetary Neutrality	91
6.2	Choice under Uncertainty; a.k.a. Portfolio choice	93
6.2.1	Probability	93
6.2.2	Expected Utility	94
6.2.3	Risk Aversion	95
6.2.4	State contingent consumption	95
6.3	Problems	96

7	Growth	121
7.0.1	Production	122
7.0.2	Growth and Savings	123
7.0.3	Other Determinants of Growth	125
8	Finance	129
8.1	Asset Pricing and the CAPM	129
8.1.1	Portfolio Choice	130
8.1.2	Time and Risk Correction	130
8.1.3	Beta Representation	132
8.1.4	The Term Structure of Interest Rates	134
8.1.5	The Market Rate of Return and the betas	135
8.1.6	Equivalent Risk Neutral Representation	136
8.2	Efficient Market Hypothesis	137
8.2.1	Idiosyncratic Risk Does Not Affect Prices	137
8.2.2	Prices Adjust Immediately to All Available Information	137
8.2.3	Risk Adjusted Prices Are Not Predictable	138
8.2.4	Arbitrage Opportunities Are Not Present in Financial Markets	138
8.3	Modigliani-Miller Theorem	140
8.3.1	Modigliani-Miller More Formally	140
8.3.2	Other Implicit Assumptions	142
8.4	Problems	143
9	Incentives	157
9.1	Principal-Agent Problems	158
9.1.1	The symmetric information benchmarks	159
9.1.2	Moral hazard	161
9.1.3	Adverse selection	165
9.1.4	Social security	169
9.1.5	Other examples [to be added]	172
9.2	Problems	172
10	Fiscal policy	175
10.1	Ricardian Equivalence - a.k.a. Arithmetics	175
10.1.1	Borrowing-Constraints	177
10.1.2	Distortionary Taxes	178
10.1.3	Krugman,, and the army of fiscal policy hawks . .	181

10.2	Optimal taxation	186
10.2.1	Distortionary Taxation	186
10.2.2	Public Goods Provision	187
10.2.3	Time Inconsistency	191
10.3	Problems	192
11	Empirical economics	193
11.1	Correlation and regression analysis	195
11.2	Controlled experiments	198
11.3	Natural experiments	199
11.4	Structural econometrics	202
11.4.1	Problems with regression analysis	202
11.4.2	Back to models	205
11.5	Problems	205
12	Game theory	207
12.1	Strategic Games	208
12.1.1	Nash equilibrium	209
12.1.2	Examples	209
12.1.3	Mixed Strategies	213
12.2	Dynamic Games	214
12.3	Bargaining	215
12.4	Reputation and repeated games	215
12.5	Problems	216
13	Mathematical Appendix	217
13.1	Sets	217
13.2	Numbers	218
13.3	Vectors	219
13.4	Functions	219
13.4.1	Some Properties of Functions	220
13.4.2	Optimization	224
13.4.3	Solving the Cobb-Douglas Consumer's Problem	229
13.5	Linear algebra	232
13.5.1	Reduction to echelon	233
13.5.2	Independent columns	234
13.5.3	Inverse matrix	235
13.5.4	Determinant	235

CONTENTS

ix

13.6 Problems	235
13.6.1 Sets and numbers	235
13.6.2 Derivatives	236
13.6.3 Maximization	240
13.6.4 Linear algebra	245

Preface

Introduction to Economic Analysis is the introductory class I taught for several years at NYU. It is the first class of the theory concentration major in Economics. Though this is not a plain vanilla Economics 101, it is nonetheless intended for students who wish to begin their formal study of economics. Its goal is to develop concepts and analytical tools that economists use to understand general social and economic phenomena. This book originates as lecture notes for this class.

As such the book relies on a relatively high level of abstraction and focuses on the concepts and the techniques of economic analysis rather than on the understanding of specific economic phenomena or institutions. It is particularly well suited for students/readers who are interested in pursuing careers or higher degrees in economics or in quantitative fields such as finance. But readers who favor an analytic approach to their intellectual endeavors might find the book enjoyable.

Some math background is needed to read this book. I offer no apologies. Economists very often express their ideas using mathematical concepts simply because these allow them to express themselves more precisely than with spoken language. Virtually, all the math used in the book is (or should be) taught in high school. We shall explain thoroughly any mathematical concept we use and we shall refer to the Math Appendix for details when necessary. How much the reader will get out of the book will depend more on how much fun he/she has doing math than on how much math he/she knows.

My gratitude goes to James Ramsey, Director of Undergraduate Studies at NYU, for convincing me to teach in the undergraduate program and for having first envisioned and then supported the whole idea of a course like *Introduction to Economic Analysis*. My gratitude goes also to the many students who have suffered through various versions of my notes over the years. I am proud to report that two of them are now successful academic economists, Ahu Gemici at NYU and Laura Pilissoph at the University of Chicago. The book also owes a lot to an exceptional list of T.A.'s who have helped me and who wrote (and solved) most of the problems: Shachar Kariv, Mark Dean, Victor Archavski, Ignacio Esponda, Luke Geldermans. They are also all by now successful academic economists.

The reader with some knowledge of the methodological distinctions in the economics profession will certainly note a certain *Chicago freshwater*

style prevailing in the book. My graduate school years at the University of Chicago in the early nineties have fundamentally marked the way I do (and teach) economics. I enthusiastically take this occasion to thank Gary Becker, Jose' Scheinkman, Mike Woodford, Bob Lucas, John Cochrane, ... for their help and their inspiration, as well as my closest peers there, Danilo Guaitoli, Afonso Mello Franco, Giorgio Topa, Fulvio Ortu.

Jess Benhabib, Andy Schotter, Chuck Wilson in particular, have shaped my

The book is dedicated to my son Vittorio. He was just learning to read when I started designing *Introduction to Economic Analysis*. Now he could read this book.



Chapter 1

Introduction

[...] the difference between economics and sociology is very simple. Economics is all about how people make choices. Sociology is all about why they do not have any choices to make. (James S. Duesenberry).



Adam Smith

David Ricardo

Economics is a broad-ranging discipline in scope. Consequently, various definitions of economics have been attempted, with mixed success. Perhaps the most influential definition is due to Lionel Robbins (in his 1932 essay on *The Nature and Significance of Economic Science*):

economics is "the science which studies human behavior as a relationship between ends and scarce means which have alternative uses."

The point of this definition is to stress the issue of scarcity, essentially to point out that economists would be useless in Nirvana, where everybody is happy with an unlimited supply of anything they might desire. Without scarcity there are no choice problems; and no choice - no economics!

More pragmatically, it is nowadays commonplace to implicitly define economics functionally, as follows:

economics is the study of a specific set of phenomena we call ‘economic;’ demand and supply of commodities, market equilibrium and prices, effects of monetary and fiscal policies, [.....].

But this (as well as any) functional definition is very narrow. Economists actively study phenomena which are traditionally attributed to sociology, anthropology, political science, law, and even biology. Examples include: crime, family, fertility, primitive societies (like hunter-gatherers), voting, comparative analysis of political and legal institutions, social norms, social networks, genetic and cultural evolution of preferences, and so many others we do not have space to list.¹

If a functional definition of economics is narrow, as the boundaries of the subjects studied in the discipline keep expanding, then we better rely on a methodological definition. An appropriate one could go something like this:

economics is the study of aggregate phenomena as equilibrium outcomes of individual choices.

Individual choice and *equilibrium analysis* are in fact at the core of any economic analysis of whatever issue. These terms are somewhat jargony and hence it's better to be clear about how we interpret them at the outset.

Individual choice. Individuals - not groups - make choices. When choosing individuals evaluate *rationally* costs and benefits: they will choose the best bundle of goods for their money; they will sell assets they know are over-valued and will buy assets they know are under-valued; they will engage less

¹Applications to economics outside of typically economic phenomena has a long tradition in the discipline. The economist which is mostly responsible for this, who has studied strange subjects when it was not cool to do so but who ultimately has received a Nobel prize for his efforts along these lines, is Gary Becker, from the University of Chicago. His view has been popularized recently by Steven Levitt, also from the University of Chicago, in a recent series of books/blog entries/NYTimes columns/movies/... under the title of *Freakonomics*.

in criminal activities if the probability of detection is higher, if punishment is more severe, if more alternatives to crime are offered ex-ante, if crime is not "cool" in their social group; they will migrate when expecting improvements in their economic and social lifestyle. Rationality can be substantially relaxed - and it is nowadays in economics. But it still represents the standard assumption in most of the discipline. *Rational choice* is referred to (and criticized) as the method of economics in sociology and political science.

Equilibrium. Individuals in an economy or society interact through markets and through different institutions (families, firms, schools, peers). Choices of different individuals are connected in the economy, and an economic analysis of a specific phenomenon considers all the relevant connections, the direct and indirect effects of a change in the determinants of such phenomenon, for example. For instance, a preference shift of young people in favor of beef over chicken will have an effect in their demand, which will in turn have an effect on the relative price of beef and chicken, which will have an effect on the demand of all people, and of the producers of beef and chicken. An equilibrium is the level of demand of young and old people, the price, the supply of beef and chicken after all these effects have been taken into account. Similarly, a technological innovation which changes the demand for skilled workers will have an effect on their wage, and on the wage of un-skilled workers too; in turn this will have an effect on their supply (e.g., because more people will acquire the demanded skills following an increase in the wage rate for skilled workers).

Often the notion of *economic equilibrium* is criticized by other social scientists (and even by some economists who should know better) on the grounds that modern economies should be better represented as in a sort of "permanent disequilibrium." This critique is based on a fundamental misconception of what an economic equilibrium is, a misconception induced by the meaning that the word "equilibrium" has in common parlance (but not in economics!). An economic equilibrium is not in any sense a state in which economic variables are constant or slowly moving or anything like that. It is in fact possible to show that the notion of economic equilibrium, even when combined with individual rationality, is perfectly consistent with any sort of complex dynamics of e.g., aggregate capital or GNP. Even *chaotic dynamics*, that is dynamics which do not converge to a constant not a cyclical state and which depend greatly from initial conditions, are possible.²

²The word "chaotic" also has a precise meaning in the mathematics of dynamical

Let's try to see the economic method (individual choice+equilibrium) at work on an issue which many would characterize as sociological, just to beat the dead parrot some more. Consider the issue of the effects of a minimum drinking age law. Such a law might reduce alcohol consumption in minors. You impose a constraint on a group (minors) - making it harder for them to drink - and they'll drink less. But this argument is based on group choice - the group drinks less. If we think of the group as composed by individuals - each one choosing rationally - we need to ask ourselves why it is rational for each one of them to drink. A possibility is that alcohol consumption is a manifestation of preferences for risky and rebellious behavior. In this case, making drinking illegal would make it even more a risky rebellious choice and hence would give minors a new incentive to drink. In this case a minimum drinking age law might even exacerbate under-age drinking. But suppose that the law punishes drinking with such severity that in fact it reduces the incentives of minors to drink. In this case, in equilibrium, they will satisfy their demand for risk and rebellion with some other substitute behavior. Let's say bungee-jumping from tall buildings. Then, to evaluate the effect of a minimum drinking age law it is not enough to see the effects on drinking, but rather the equilibrium effects on bungee-jumping - which might be worse than drinking - might be considered.... Now is the time for you to be impressed about how deep the economic method (individual choice+equilibrium) can get with ease and speed.

1.1 Theory, Models, and Empirical Analysis

An economist is the only professional who sees something working in practice and then seriously wonders if it works in theory (Ronald Reagan)

Nothing is less real than realism. Details are confusing. It is only by selection, by elimination, by emphasis that we get to the real meaning of things (Georgia O'Keeffe)³

Models are theoretical exercises of abstraction: *ignoring many details in order to focus on the most important elements of the problem.*

systems that is not well represented by its meaning in common parlance.

³Thanks to Andrea Moro for suggesting, and in turn to Narayana Kocherlakota for discovering, the quotation. .

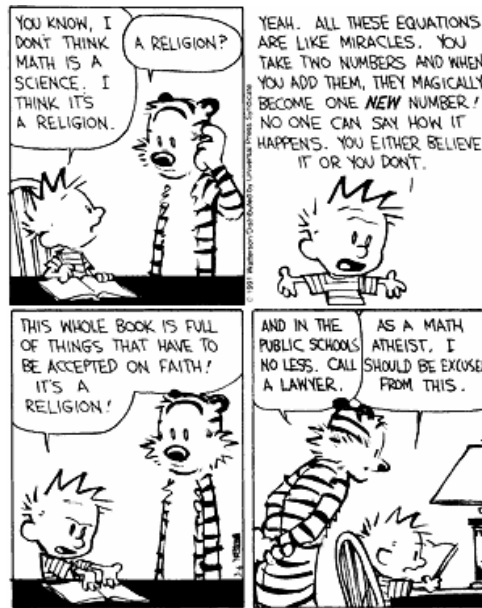


Figure 1.1: Math in economics

There is no such thing as the right degree of abstraction for all analytic purposes. The proper degree of abstraction depends on the objective of the analysis. A model that is a gross oversimplification for one purpose may be needlessly complicated for another. A map might be an appropriate metaphor for a model: we rarely need 1:1 maps; and sometimes we need a map of the whole American continent, sometimes one of the upper east side of NYC.

Models are not necessarily mathematical models. The following example (taken from Krugman, 1995) illustrates this point:

Dave Fultz at the University of Chicago in the late 40's showed that a dishpan filled with water, on a slowly rotating turntable, with an electric heating device bent around the outside of the pan provides a good representation of the basic pattern of weather. The dishpan was build to model the temperature differential between the poles and the equator and the force generated by the earth's spin (abstracting from most of the the intricacies and complexities of the earth geography) and was successfully shown to exhibit phenomena which could be

interpreted as tropical trade winds, cyclonic storms of the temperate regions, and the jet stream.

But often models are in fact mathematical models: the most sophisticated weather forecasts nowadays require the estimation of the parameters of a large number (very large, hundreds) of equations. Most economic models, in particular, are mathematical models. This is in part due to the fact that math is a very efficient language for abstract arguments (especially, because it facilitates the manipulation of complex logical arguments and the identification of logical and conceptual mistakes in abstract arguments). But also, mathematics, especially when coupled with fast computers, allows the constructions of models which can be used as laboratories, that is, mathematical imitation-economies that generate simulated data which can be compared with actual data from real economic systems (Lucas, 1980 develops this point with clarity).

Models can therefore be tested with the methods of statistics and econometrics. While models are not easily "falsified" because their implications are mostly not deterministic but rather stochastic (statistical, if you wish; more on this later), econometric analysis supports some models against others and in the end guides economic theorizing. For instance in the theory of economic decision and in game theory, ample evidence of deviations from rationality has significantly affected the models economists use. Also, the econometric evidence regarding real effects of monetary policy (e.g., of liquidity injections) has severely affected macroeconomic theory as well as actual monetary policy.

1.2 Debates

Why does public discussion of economic policy so often show abysmal ignorance of the participants? Why do I so often want to cry at what public figures, the press, and television commentators say about economic affairs (Robert M. Solow).

Politicians and reporters are fond of pointing out that economists can be found on both sides of many issues of public policy. If economics is a science, why do economists quarrel so much? After all, physicists do not debate whether the earth revolves around the sun or vice versa.

The question reflects a misunderstanding of the nature of science! Disputes are normal at the frontier of any science. Clearly, nowadays physicists

do not argue whether the earth revolves around the sun but they did (quite vociferously). They do now argue however how useful is superstring theory, though their disagreement on this matter goes mostly unnoticed to the public because only few of us understand what they are talking about. When physicists debate about issues that politicians and the general public care about and think they understand, the outcome is the same, as is it obvious e.g., on the issues regarding the causes (and even the existence) of global warming.

Disputes about economics are generally aired to the public and thus all sorts of people are eager to join in the ensuing debates, relying on their common sense. Unfortunately, common sense is not always a reliable guide in economics since many economic relationships are counterintuitive. Hopefully, by the end of our course we will have a better sense of when common sense works and when it fails.⁴

1.3 Predictions

Many economic commentators have observed in the course of the last few years that few economists have anticipated the financial market crisis in the fall of 2008. A cursory look at the time series in the This is fact true, though a handful of economists did in fact publish warnings about an impending crisis. These economists include Nouriel - a.k.a. Doctor Doom - Roubini (NYU Stern School of Business and RGE Monitor), Robert Shiller (Yale University), Raghu Rajan (University of Chicago Booth School of Business).

A cursory look at Figure 1.2, which report the time-serie of the Dow Jones Industrial Index (one of the most common U.S. stock market indexes), may make it seem as if at least in the last quarter of 2008 the impending crisis had been clearly manifesting itself. In fact crying wolf in the last quarter of 2008, predicting a crisis would have been easy enough (with a guarantee of get it right with about 50% probability). Many have done this in fact. But this is not what we mean by predictions in economics. What we mean is what for instance R. Rajan at the Federal Reserve Bank of Kansas City's Annual Jackson Hole Symposium in 2005, while head of the IMF. He did not just predict a crisis, but he provided a coherent analysis of the reasons for a possible impending crisis. This is the crux of the matter.

⁴Besides, economists tend to agree on many issues, not lastly e.g., the economic gains of free trade.



Figure 1.2: Dow Jones Industrial index

Let me try and better explain with a metaphor what I mean when I talk of *generating predictions out a coherent analysis*.

Inter Milan will soon be playing A.C. Milan at San Siro. I predict Inter Milan will win. How did I come up with this prediction? Three possibilities:

1. I am an A.C. Milan fan and superstitiously (like any Italian) I believe that saying that Inter Milan will win brings bad fortune to it.
2. A tarot reader/an oracle told me.
3. By selling Kaka' A.C. Milan has no valid attacking midfielder. A.C. Milan's game will have to start in the defensive side of their midfield, with A. Pirlo. Inter Milan, on the other hand, can play a solid 4-3-1-2, with W. Sneijder behind the attackers. This will allow Inter Milan to press A.C. Milan in its half, stopping A.C. Milan's source of play in A. Pirlo. Inter Milan's two new great forwards, D. Milito and S. Eto'o will do the rest.

Ok, you do not need to know everything about Italian soccer to understand what I am talking about. And you guessed it, generating predictions out a coherent analysis means doing as in predictions of type 3.

Generating predictions out of a coherent analysis, not just predicting, is the objective of economics as a discipline. You might already guess that by *a coherent analysis* economists mean the outcome of a model. It must also be added that economic models are typically stochastic models, that is they predict stochastic processes over time. In other words, economic models do predict the time path a quantity of interest, say the Dow Jones index, but predict the probability associated to each possible time path of the index.⁵ If an event, say a financial crisis, has a very small probability of occurring - it does not mean it can't occur. If it does occur, it does not mean that the model predicting a very small probability associated to the event is wrong. Several investment banks had in fact predicted very accurately what would have happened if real estate prices would have stopped growing (or, God forbid, started to decline); they just associated a very low probability to the event that real estate prices would start declining.

Stochastic models are a way to explicitly admit one's inability to predict with certainty. In this sense, stochastic prediction are, in some sense, an admission of partial failure.

Consider the example of a fair coin throw. It is typically modelled as a stochastic process in which the event Heads occurs with probability $1/2$ and the same for Tails. But I could construct a model of the dynamics of a coin, depending e.g., on the initial position of the coin (Heads or Tail), the weight of the coin, the force the thumb operates on the coin, ... With such a model, and with the observation of my determining variables, I could probably predict the outcome of the coin throw with more precision, or maybe even predicting it with relative certainty. You see then the sense in which modeling the coin throw as a stochastic process is an admission of failure: I cannot construct a model of the dynamics of a coin (or I rely on approximations, which is something else we do not want to get into here).

Finally, a quick and dirty comparison between predictions in economics and in other disciplines can be useful as an illustration. A couple of examples will be enough.

Classical physics is very successful in developing deterministic predictions (and in testing them). But remember however that not all physics is about forecasting the trajectory of a bullet, which can be done with a certain deterministic precision thanks to Newton. Quantum theory broke the certainty

⁵Even in the context of soccer predictions, when I say *I predict Inter Milan will win* I mean something like *I predict that Inter Milan has higher chances of winning*.

spell for physicists - to the point of leading A. Einstein to famously complain that *God does not play dice*. And then physicists are better scientists (no doubts about this). Or may-be they also have it easier (Max Plank apparently thought so: Professor [Max] Planck, of Berlin, the famous originator of the Quantum Theory, once remarked to me that in early life he had thought of studying economics, but had found it too difficult!," John Maynard Keynes, 1933.)

Also, in other hard sciences, meteorologists produce stochastic predictions about the weather, vulcanologists about eruptions, geologists about earthquakes,...

Anthropology as a discipline is methodologically centered on collecting data and stressing differences across human cultures. For the most, anthropologists refuse to order data through models. Anthropological accounts tend to be exceptionally interesting, but, without the use of models, predictions are impossible (nor are they an aim of anthropologists). Ever heard anybody saying, *In that unexplored forest there has to be two, possibly three, hunter and gatherer tribes whose system of religious beliefs is animistic and one which is evolving towards monotheism?* Nonetheless, we owe to anthropologists a much deeper understanding of human nature; this is what they are after, with great success.

Sociology and most of political science⁶ have similar, while much less extreme, methodological characteristics as anthropology. A large part of political science, for instance, is *comparative*, essentially a detailed institutional analysis of different political systems.

In medicine, a doctor can (and will) tell you that smoking increases your chances to develop lung cancer, but cannot tell you if and when you will develop one, nor that you won't if only you stop smoking. Similarly, an economist can (and will) tell you that allowing large leverage ratios in financial markets (investing with borrowed money) increases the chances of a financial crisis, but cannot tell you if and when one will occur, nor that it won't if only we strike a stringent regulation on financial markets.

Having said all this, there is another reason why only few economists have anticipated the financial market crisis in the fall of 2008: our theory of asset prices predicts that asset prices are not predictable, in the sense that

⁶A sizeable and growing part of political science has essentially adopted the method of economics. Rational choice political scientists, so they are called, are not distinguishable from economists, though they sit in different academic departments. This is not the case for sociology. Rational choice sociologists are few still leave in reservations next to Native Americans.

we can only say that riskier assets have higher returns on average (this is not a joke, nor an empty theory - you will understand what I mean later on in the course when we'll do finance). Furthermore, economists do not have an established theory of asset pricing bubbles and hence tend to underestimate the probability that a sustained growth in an asset price is a bubble. Many economists are back to the blackboard studying bubbles these days. As it should be.

1.4 Empirical analysis

Many observers (social scientists, journalists, people you cannot refrain from talking when they should not,...) complain that economics as a discipline reduces to useless theoretical/mathematical exercises. This is simply false. It is most probably the distorted view of a tribe whose system of beliefs include tenet that *any math is too much math!*

Even a superficial look at a good economics journal (*Econometrica*, *American Economics Review*, *Journal of Political Economy*, *Quarterly Journal of Economics*,...) will give you the (correct) impression that economics is an applied discipline, addressing well-defined real world problems. For instance, these are the titles of the last issues, June and August 2010, of the *Journal of Political Economy* and *Quarterly Journal of Economics*, respectively:

Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors

Innovation, Firm Dynamics, and International Trade

Housing Externalities

Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings

Competition and the Structure of Vertical Relationships in Capital Markets

Estimating Welfare in Insurance Markets Using Variation in Prices

Measuring Beliefs and Rewards: A Neuroeconomic Approach

Monetary Non-Neutrality in a Multisector Menu Cost Model

Regulation and Distrust

Improved Access to Foreign Markets Raises Plant-Level Productivity... for Some Plants

Sex and Science: How Professor Gender Perpetuates the Gender Gap

Can Exchange Rates Forecast Commodity Prices?

Counterparty Risk in Financial Contracts: Should the Insured Worry about the Insurer?

The Geographic Determinants of Housing Supply

School Choice with Consent

Muslim Family Law, Prenuptial Agreements, and the Emergence of Dowry in Bangladesh

No papers on Hilbert spaces nor on Martingales or Matrix algebra *per se* (though all these mathematical concepts are indeed used by economists when deemed useful). If you opened the journals, however, you would find a lot of equations - often statistical analyses of data (which economists call *econometrics*).

Statistical analyses in economics are commonplace. They are also typically difficult. This is because economists generally do not have access to controlled randomized experiments, the typical instrument of empirical analysis of the hard sciences and of medicine and pharmacology. Let me illustrate.

A test of the effects of a new drug e.g., for hypertension involves selecting a group of subjects with the same medical characteristics - a fraction of which are chosen randomly to receive the drug (the others typically receives a placebo). Since the subjects have *ex-ante* the same characteristics, any differential outcome in the two groups after the experiment is reasonably interpreted to be an effect of the drug. An economist willing to test the effects of monetary policy e.g., on employment, on the other hand, cannot use anything like a controlled randomized experiment: all individuals in the economy are affected by the policy, no two groups with the same characteristics can be selected.

As a consequence, economists have had to develop sophisticated methods to try and identify relationships of cause-effect in the data.⁷ We shall talk about these methods in a subsequent chapter.

1.5 Some successful results

At the cost of appearing cocky, let me finish the chapter by listing some examples of important non-obvious results of economics which illustrate its intellectual success as a discipline:

- The characterization of *gains from trade* in general and of international trades in particular. This is the *Invisible hand* result (or First Welfare Theorem).
- *Comparative advantage* (and not absolute advantage) determines trade flows between countries: a country which is less productive than any other one at producing all tradable goods will still find international trade markets where to sell some of its products.
- Risk adjusted returns in financial markets are unpredictable. This is the *No-arbitrage* theorem.
- Several "neutrality" or "equivalence" results (careful! these results hold under restrictive assumptions - somewhat like "in the vacuum" results in physics - and hence never in real economies, but are fundamental benchmarks that, when not recognized, induce important logical mistakes):
 - i) Doubling the amount of money in an economy has no real effects; equivalently, dividing all prices by a third (multiplying the value of a Dollar by three) has no real effects. This is called *Monetary neutrality*.

⁷Recently, economists have been able to design randomized experiments regarding the effects of small scale policies in development, like, for instance, the effect of a new school in a isolated village. In these contexts they can in fact select villages with similar characteristics and then choose randomly some of them to set up a school in. The Abdul Latif Jameel Poverty Action Lab at MIT is at the forefront of these methods.

- ii) Given government expenditures in the present and future, how they are financed, e.g., by taxing now, or by taxing later, or by printing money (inflation) has no real effects. This is called *Ricardian equivalence*.
- iii) Given a firm's production plan in the present and in the future, how is the firm financed, e.g., by equity or debt, has no effect on its value. This is the *Modigliani-Miller theorem*.
- The Malthusian theory of fertility, that is, economies will not grow because fertility growth will eat up all income growth, is inconsistent with individual rationality.
- Evolutionary theory in biology can be accurately represented by the methods of game theory. The same for Foraging theory.
- Altruistic (or cooperative) behavior is consistent with individual rationality in a well-defined series of conditions. This is the *Folk theorem*.
- Value and equilibrium prices coincide (once "value" is properly defined). This is the *Theory of value*.
- The study of the effects of economic policy is logically and empirically flawed when not embedded in equilibrium analysis. This is called the *Lucas Critique*.
- Many empirical relationships and many stable correlations have been uncovered by means of statistical and econometric techniques. Example include, the effects of taxes on labor supply, the determinants of business cycles, the determinants of asset prices, and many many many others.
-

1.6 Fun readings

P. Krugman, *Development, Geography, and Economic Theory*, 1995, MIT Press; Ch. 3.

G.S. Becker, *The Economic Way of Looking at Life* 1992, Nobel Lecture, posted on my web page.

R.E. Lucas, ‘Methods and Problems in Business Cycle Theory,’ *Journal of Money, Credit, and Banking*, 12, 1980, Part 2.

On Freakonomics, read Ariel Rubinstein’s critique:

A. Rubinstein, ‘Freak-Freakonomics,’ *The Economists’ Voice*, 3(9), Article 7, 2006.

On the financial crisis of 2008 and the debate on economics it has generated you might want to read

P. Krugman on the *NYTimes Magazine*, September 6th, 2009;

who has a very different take on the position of economists with respect to the crisis than the one offered here, but he reaches similar conclusions regarding the economics of bubbles. See also

P. Krugman, *How did economists get it so wrong?*, New York Times Magazine, Sept. 2nd, 2009.

J.H. Cochrane, *How did Paul Krugman get it so wrong?*, University of Chicago, mimeo.

R.G. Rajan, ‘Has financial development made the world riskier?’, presented at the Kansas City Fed Meeting at Jackson Hole, 2005.

A. Rolnick, ‘Interview with Thomas Sargent,’ *The Region*, September 2010.

1.7 Useful references

C. Simon and L. Blume, *Mathematics for Economists*, Norton, New York, 1994.

A. Schotter, *Microeconomics: A Modern Approach*, Addison Wesley Longman, New York, 2000.

H. Varian, *Intermediate Microeconomics: A Modern Approach*, Norton, New York, 1999. R.P. McAfee and T. Lewis, *Introduction to Economic Analysis*, Flatworld Knowledge, 2008.

- M. Osborne, *An Introduction to Game Theory*, Oxford University Press, Oxford, 2004.
- K. Binmore, *Fun and Games: A Text on Game Theory*, D. C. Heath and Company, Lexington, MA, 1992.
- H. Gintis, *Game Theory Evolving*, Princeton University Press, Princeton, NJ, 2009.
- R.E. Wright and V. Quadrini, *Money and Banking*, Flatworld Knowledge, 2008.

Chapter 2

Rational choice



Jeremy Bentham

Vilfredo Pareto

Gerard Debreu

Ken Arrow

In this chapter we shall study how (economists postulate) individuals make decisions. An important by-product of the theory is a precise definition of rationality.

The central figure of decision theory in economics is the individual *decision-maker* (DM). The typical example of a DM is the *consumer*. We shall assume that:

the DM has well defined *preferences* over a *choice set*;

the DM is *rational*.

We will show that these assumptions, especially rationality, once it is defined properly, imply that the DM's preferences can be represented by a

utility function and that the DM chooses her preferred element in the choice set *as if* she were *maximizing* her utility function in the choice set.

2.1 Preferences

The choice set is represented by X , an arbitrary set of objects; e.g., a list or a set of *consumption bundles*. An example should clarify, $X = \{\text{Ducati motorcycle, trip to Vegas, iPad}+\$500 \text{ on iTunes, year tuition at NYU, ...}\}$. More abstractly, we could have $X = \{x, y, z, k, p, \dots\}$.

Preferences are represented by a *preference relation*, denoted by \succsim (note that it is not \geq), which allows us to compare any pair of alternatives x, y in the set of objects X , from the point of view of the DM.

Definition 2.1.1 We read $x \succsim y$ as "*x is at least as good as y*" or "*x is weakly preferred to y*," for the DM.

In words, we also say "*x is weakly better than y*," for the DM (the DM weakly prefers x to y). The term "weakly" is math jargon: *x is weakly better than y* means that either the DM likes x better than y or else he is indifferent between x and y . From \succsim we can derive two other important relations on X .

Definition 2.1.2 The strict preference relation, \succ , satisfies

$$x \succ y \text{ iff } x \succsim y \text{ but not } y \succsim x$$

where iff means *if and only if*.

Read $x \succ y$ as "*x is strictly preferred to y*," for the DM (the DM strictly prefers x to y).

Definition 2.1.3 The indifference relation, \sim , satisfies

$$x \sim y \text{ iff } x \succsim y \text{ and } y \succsim x$$

Read $x \sim y$ as "*x is indifferent to y*," for the DM (this is not proper English but economists often get away with it; "the DM is indifferent between x and y " is better said).

Imagine that we present a consumer with any pairs of alternatives, x and y , and ask how she compares them: Is either x or y weakly better than the other in your eyes? For each pair x and y , we can imagine four possible responses to our question:



Figure 2.1: Milton Friedman and George Stigler, both Nobel Prize winners in economics, strolling down University Ave. on the University of Chicago campus.

- i) x is better than y , but y is not better than x ,
- ii) y is better than x , but x is not better than y ,
- iii) neither is better,
- iv) I am unwilling to commit to a judgment.

Make sure you understand that the answer that x is better than y *and* y is better than x is logically possible only if when the DM says "better" she means "weakly better." We need to be precise with our DM when we ask her questions!

Finally, note that a preference ordering is not formally different from other more common binary orderings like e.g., the height ordering, which we use when we say that Uma Thurman is (weakly - and strictly) taller than Reese Witherspoon; or that George Stigler was (weakly - and strictly) taller than Milton Friedman.

2.2 Rationality

The hypothesis of rationality is embodied in two assumptions about the weak preference relation \succsim : *completeness* and *transitivity*.

Definition 2.2.1 (Completeness) For all $x, y \in X$, either $x \succsim y$ or $y \succsim x$ or both.

Completeness simply requires that any two elements of the choice set can be compared by the DM. As such it is hardly objectionable as a first order assumption. But can you think of situations in which it fails ? Try introspection.

Definition 2.2.2 (Transitivity) *For all $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$ then $x \succsim z$.*

Transitivity implies that a DM never faces a sequence of pairwise choices in which her preferences cycle, that is, such that $x \succsim y$, $y \succsim z$ and $z \succ x$ (we also write cycles more compactly as $x \succsim y \succsim z \succ x$). Is transitivity of preferences a necessary property that preferences would have to have? Try introspection again. Is it a plausible assumption?

To try and convince you that transitivity is actually a plausible assumption about preferences, notice the following:

Lack of transitivity implies the possibility of *money pumps*. Here's an example of a money pump.

Consider an economy with 3 goods, x , y , z , and 2 agents, 1, 2.. Agent 1 is endowed with one unit of z , one unit of y , and some money. Agent 2 is endowed only with one unit of good x . Agent 1 has non transitive preferences \succsim such that: $x \succsim y \succsim z$ and $z \succ x$. We do not specify any preferences for agent 2 [make sure that at the end of the example you understand why we do not need to]. Consider the following sequence of trades between agent 1 and 2: Agent 2 gives agent 1 a unit of good x in exchange for a unit of good y . - Agent 2 gives agent 1 a unit of good y in exchange for a unit of good z . - Agent 2 gives agent 1 a unit of good z in exchange for a unit of good x plus a (small but positive) sum of money. Note that the sequence of trades is constructed so that it is feasible (possible) from the given endowments, and so that agent 2 is willing to enter each of these trades. At the end of the sequence of trades agent 2 has the unit of x he was endowed with, plus a sum of money. Agent 1 has instead the units of y and z he was endowed with, but he has lost a sum of money. This trade cycle can be repeated indefinitely until agent 1 lost all his money.

On the other hand, experimental evidence is not always consistent with transitivity. An important phenomenon, called *endowment effect*, has been

first documented in laboratory experiments by Kahneman, Knetsch, and Thaler in the late 80's and often replicated since. Here's a version of the experiment.

A student is shown two objects, a pen and a coffee mug. She is asked to say which object she likes better and she answers "the pen." She is then asked how much she is willing to spend to buy the pen, and she answers "\$1." In the course of the experiment the student is given the mug as a present. At the end of the experiment she is asked if she wants to give back the mug in exchange for \$1, and she refuses. This behavior turned out to be relatively typical among subjects in the experiment (in the original experiment the subjects were Cornell University college students). It constitutes a violation of transitivity:

$$\$1 \succsim pen, pen \succsim mug \text{ and } mug \succ \$1.$$

In summary, completeness and transitivity are assumptions about DM's preferences, not statement of logical necessity. Convince yourself that completeness and transitivity are a reasonably accurate assumptions about a DMs' preferences; or that they are not. Economists tend to think that they are, in most relevant decision environments: a sub-field of economics, *Experimental economics*, is about testing these assumptions by means of behavior of subjects in the lab. This field now interacts extensively with psychology and even neuroscience (neuroscience labs typically contain *functional-MRI machines*, essentially big magnets, which allow the experimenter to observe brain activation of subjects while they solve choice problems).

We conclude re-iterating the main point of this section: *rationality in economics essentially requires assuming completeness and transitivity of individual DMs' preferences.*

Nothing more and nothing less.

2.3 Utility representation

Rationality of DM (completeness and transitivity of preferences) allows us to represent the preference relation, \succsim , by a utility function.

Definition 2.3.1 *A function $u : X \rightarrow \mathbb{R}$ is a utility function representing preference relation \succsim if, for all $x, y \in X$,*

$$x \succsim y \Leftrightarrow u(x) \geq u(y).$$

As you will painfully learn studying these notes, it is much easier to work with the binary relationship \geq on real numbers than with \succsim on arbitrary sets of objects. This is exactly what we are after by introducing the concept of a utility function. The following theorem is central to the analysis.

Theorem 2.3.2 (Representation) *A preference relation \succsim can be represented by a utility function if and only if it is complete and transitive.*

Proof. Assume the set X contains a finite number N of elements, (x_1, \dots, x_N) . [the general proof is more complicated and is omitted]
[only if] By contradiction. Suppose \succsim is not complete. Then there exist $x, y \in X$ such that neither $x \succsim y$ nor $y \succsim x$. But if \succsim can be represented by some function $u : X \rightarrow R$, then either $u(x) \geq u(y)$ or $u(y) \geq u(x)$ or both (the order \geq on real numbers is in fact complete). This is a contradiction. Suppose \succsim is not transitive. Then ... continue by analogy with the argument used above for completeness.

[if] By construction. Construct a utility function $u : X \rightarrow R$ as follows. Rank (from worst to best) in terms of preferences \succsim all elements (x_1, \dots, x_N) . Suppose, without loss of generality (make sure you understand this is in fact the case), that the rank is exactly x_1, \dots, x_N , that is, $x_i \succsim x_{i+1}$, for $i = 1, 2, \dots, N - 1$.

Here's the construction.

1. Let $u(x_1) = 1$;
2. Let $u(x_2) = u(x_1) = 1$ if and only if $x_2 \sim x_1$; otherwise, if $x_2 \succ x_1$ let $u(x_2) = u(x_1) + 1$;
3. Proceed as in step 2 for all $u(x_i)$, $i = 3, 4, \dots, N - 1$.

Now, you should be able to prove that the utility function, as we constructed it, is in fact a representation of the preference ordering \succsim . ■

Do it. To make sure you really understood the proof, answer the following questions:

- i) *Where did we use completeness and transitivity in the "if" part of the proof? I hid this step on purpose. Find it.*

- ii) Is it important that, in step 2, that, if $x_2 \succ x_1$, we let $u(x_2) = u(x_1) + 1$ as opposed to $u(x_2) = u(x_1) + 1.39672$ or some other real number?
- iii) Is it important that, if $x_2 \succ x_1$ and $x_3 \succ x_2$, the utility function we construct satisfies the property that $u(x_2) - u(x_1) = u(x_3) - u(x_2)$?

If you answer correctly these questions, you have understood that the *Representation theorem* implies that the only property of a utility function that is important is how it orders the bundles. The level of utility and the size of the utility difference between any two bundles doesn't matter. In the economics jargon, we say that

preferences are ordinal and not cardinal.

This is an issue that has bothered the classical (English, 18th and 19th century) economists quite a bit. More precisely stated, *ordinality of preferences* is implicitly defined by the following result.

Proposition 2.3.3 *If a utility function $u(x)$ represents the preference relation \succsim , any monotonic strictly increasing transformation of $u(x)$, $f(u(x))$, also represents the same preference relation.*

Proof. Suppose u represents some particular preference relation, \succsim . Then, by definition,

$$u(x) \geq u(y) \text{ if and only if } x \succsim y$$

But if $f(u)$ is a monotonic strictly increasing transformation of u , then

$$u(x) \geq u(y) \text{ if and only if } f(u(x)) \geq f(u(y))$$

Thus,

$$f(u(x)) \geq f(u(y)) \text{ if and only if } x \succsim y$$

and the function $f(u)$ represents the preference relation \succsim in the same way as the function u . ■

Of course *ordinality of preferences* also implies that it is impossible to compare two different agents' utility functions (*impossibility of interpersonal comparison of utility*, economists say). Statement like "*I like the Rolling Stones more than you do*" are absolutely meaningless in the context of economic decision theory. You understand that this is an important, if disturbing, property of our theory.. It in turn implies, for instance, that I cannot

say that society would do good by taking $\$x$ from a rich individual to give them to a poor individual, no matter how small is x (and how rich is the rich and how poor is the poor). (We shall discuss later on in the course how economists do/don't resolve this issue). Vilfredo Pareto (Italian economist and sociologist, end of 19th century; the father of the modern theory of preferences) explained this to his students by saying, "*my tooth-ache is different than yours.*"

2.4 Choice

A rational agent chooses then the element x of the choice set X to which is associated the highest utility $u(x)$. Formally, a rational agent choice problem is written:

$$\max_{x \in X} u(x)$$

Typically, however, the choice set faced by a rational agents is subject to some constraint. This is the case, for instance, for the classic example of a choice problem, the consumer problem. Let m be a fixed amount of money available to the consumer and let $p = (p_1, \dots, p_k)$ be the vector of prices for goods $1, \dots, k$. The set of affordable objects, for the consumer, is given by

$$B = \{x \in X : p_1x_1 + \dots + p_kx_k \leq m\}$$

and is called the *budget set* (and $p_1x_1 + \dots + p_kx_k \leq m$ is called the *budget constraint*). Note that we can also write $p_1x_1 + \dots + p_kx_k \leq m$ more compactly in vector notation as $px \leq m$. The consumer problem can then be written as:

$$\text{choose } x \text{ to maximize } u(x) \text{ subject to } x \in B,$$

or equivalently as:

$$\text{choose } x \text{ to maximize } u(x) \text{ subject to } x \in X \text{ and } px \leq m.$$

More compactly, we write:

$$\max_{x \in B} u(x).$$

2.5 Social choice

We might debate if completeness and transitivity represent accurate assumptions about a DMs' preferences, but we can show that they are certainly not accurate description of *group preferences*.

Consider an economy with 3 goods, a , b , c and 3 agents, 1, 2, 3. Suppose that each agent has complete transitive preferences over the 3 goods, and that no agents is ever indifferent between any of the possible binary comparisons. We can then represent their preferences by rankings. Agent 1's rank (from best to worst) is a, b, c ; agent 2's is b, c, a ; and agent 3's is c, a, b . Suppose that *group preferences*, represented by a preference ordering \succsim_G , are formed by majority voting from the preferences of the 3 agents. That is, the group prefers x to y if and only if at least 2 of the agents do. Construct now the preferences of the group, \succsim_G . Do it! It's instructive: associate a statement of the form $x \succsim y$ to any 2 elements x, y of the choice set $X = \{a, b, c\}$ - where $x \succsim_G y$ if and only if for at least two agents $x \succsim y$). If you have done this correctly, you will see that

$$x \succsim_G y \succsim_G z \text{ and } z \succ_G x.$$

That is, the group has non-transitive preferences.

The fact that group preferences can easily display non-transitivity is a fundamental result. It is due to the marquis de Condorcet (an 18th century French philosopher, mathematician, and political scientist). Instances of non-transitivity are usually called *Condorcet cycles*. The existence of Condorcet cycles means that we cannot treat groups as individuals, that we typically cannot disregard the fact that groups are formed of different individuals when we study group behavior: *groups do not necessarily display rational (complete and transitive) preference ordering*.

In other words, *social choice* and *individual choice* are distinct problems, which cannot in general be dealt with using the same conceptual tools. I can't resist noticing that this is a major blow to classical sociology, defined as the study of groups *per se*, independently of the individuals constituting the group.

But notice that I (as Condorcet) postulated that group preferences are formed by majority rule. May-be here is the problem. May-be there exist a different mechanism to aggregate individual preferences into group preferences which guarantees that the resulting preference ordering is always (that

is, for any underlying individual preferences) complete and transitive. Well, ... not really.

Theorem 2.5.1 (Arrow impossibility) *The only mechanism which aggregates individual preferences into a complete and transitive preference ordering, for any underlying complete and transitive individual preferences, is dictatorship, that is a mechanism which selects one individual of the group and identifies group preferences with the preferences of that individual.*

The theorem takes its name from Ken Arrow.¹ This result has opened up a whole new sub-field of economics: *social choice theory*.

2.6 Classic references

The first version of the coffee mug experiment is in:

Daniel Kahneman, Jack Knetsch, and Richard Thaler, "Experimental tests of the endowment effect and the Coase theorem," *Journal of Political Economy*, 98(6), 1990.

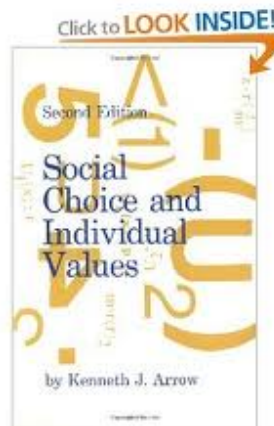
The first representation theorem is in:

Gerard Debreu, "Representation of a Preference Ordering by a Numerical Function", in Thrall et al., editors, *Decision Processes*, 1954.

The impossibility theorem is contained in Ken Arrow's dissertation:

Kenneth Arrow, *Social choice and individual values*, Cowles Foundation,

¹We shall see his name again in other parts of the course: he is certainly one of the most influential economists of this century.



Yale University Press, 1951.

2.7 Useful references

A beautiful very simple introduction to decision theory is in

David Kreps, *A course in microeconomic theory*, Chapter 2.1, Prentice Hall, 1990.

Two masterly treatments of the theory (somewhat difficult) are:

David Kreps, *Notes on the theory of choice*, Westview Press, 1988.

Ariel Rubinstein, *Lecture Notes in Microeconomic Theory: The Economic Agent*, Chapters 1-3, Princeton University Press, 2006.

A nice and simple introduction to behavioral anomalies can be found in:

Daniel Kahneman; Jack Knetsch; Richard Thaler, Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, *Journal of Economic Perspectives*, 5(1), 1991.

A relatively simple formal statement and proof of Arrow impossibility theorem are in:

John Geanakoplos, Three brief proofs of Arrow's impossibility theorem, mimeo, Cowles Foundation, Yale University.

2.8 Problems

After a brief review of decision theory, on preference relations, we list a series of problems. Some of the problems have solutions, the others are left as exercises for the reader.²

Problem 2.8.1 *Assume the weak preference relation, \succeq , is complete and transitive. Prove that this implies that the strict preference relation, \succ , is transitive. Is \succ also complete?*

Answer. Let $x, y, z \in X$ such that $x \succ y$ and $y \succ z$. Firstly we have

$$\begin{aligned} x &\succeq y \text{ and } y \succeq z \\ \Rightarrow x &\succeq z \text{ (by transitivity of } \succeq \text{)} \end{aligned}$$

Secondly we have

$$y \not\succeq x \text{ and } z \not\succeq y$$

From here we would like to conclude that $z \not\succeq x$, but we not to shown the transitivity of the negation of the weak preference relation. Suppose it was the case that $z \succeq x$. We also have from above that $x \succeq y$. By *transitivity* of \succeq it must be the case that $z \succeq y$. However $y \succ z$ implies $z \not\succeq y$, and $z \succeq y$ and $z \not\succeq y$ cannot hold at the same time. Thus $z \not\succeq x$. Summing up we have

$$\begin{aligned} x &\succeq z \text{ and } z \not\succeq x \\ \Rightarrow x &\succ z \text{ (by definition of } \succ \text{)} \end{aligned}$$

and so we have the transitivity of \succ . Try and answer the completeness question.

Problem 2.8.2 *Assume the weak preference relation, \succeq , is complete and transitive. Then the indifference preference relation, \sim , defined as above is transitive. Is \sim also complete?*

Answer. Let $x, y, z \in X$ such that $x \sim y$ and $y \sim z$. Firstly we have

$$\begin{aligned} x &\succeq y \text{ and } y \succeq z \text{ (by definition of } \sim \text{)} \\ \Rightarrow x &\succeq z \text{ (by transitivity of } \succeq \text{)} \end{aligned}$$

²Recall that X denotes a choice set and \succeq a preference relation. Let \succ and \sim be the associated strict preference and indifference relations, respectively.

Secondly we have

$$\begin{aligned} z \succeq y \text{ and } y \succeq x & \text{ (by definition of } \sim \text{)} \\ \Rightarrow z \succeq x & \text{ (by transitivity of } \succeq \text{)} \end{aligned}$$

Summing up we have

$$\begin{aligned} x \succeq z \text{ and } z \succeq x \\ \Rightarrow x \sim z & \text{ (by definition of } \succsim \text{)} \end{aligned}$$

and thus we have shown the transitivity of \sim . Try and answer the completeness question.

A few of the problems follow straightforwardly from the analysis of Chapter 1. Some others require a significant amount of ingenuity.

Problem 2.8.3 (i) Consider a group of individuals A, B and C and the relation at least as tall as, as in A is at least as tall as B . Does this relation satisfy the completeness and transitivity properties? (ii) Consider instead a group of individuals A, B and C and the relation taller than, as in A is taller than B . Does this relation satisfy the completeness and transitivity properties? (iii) Consider the same group of individuals A, B and C and the relation at least as smart as, as in A is at least as smart as B . Does this relation satisfy the completeness and transitivity properties?

Answer. (i) The relation *at least as tall as* is complete and transitive. To verify completeness, pick any two individuals A and B . Clearly, either individual A is at least as tall as individual B or individual B is at least as tall as individual A or both. For transitivity, pick three individuals A, B and C and suppose that individual A is at least as tall as individual B and individual B is at least as tall as individual C . Obviously, individual A must be at least as tall as individual C . Thus, the relation at least as tall as satisfies the transitivity property.

(ii) is left to the reader. (iii) is as (i) if smartness is defined, e.g., by the I.Q. number. But we can certainly fudge the definition so as to generate an incomplete or a non-transitive relationship. Try doing this.

Problem 2.8.4 Determine if completeness and transitivity are satisfied for the following preferences defined on $x = (x_1, x_2)$ and $y = (y_1, y_2)$: (i) $x \succsim y$ iff (if and only if) $x_1 \leq y_1$ and $x_2 \leq y_2$; (ii) $x \succsim y$ iff $x_1 \geq y_1$ and $x_2 \geq y_2$;

(iii) $x \succsim y$ iff $x_1 > y_1$ or $x_1 = y_1$ and $x_2 > y_2$; (iv) $x \succsim y$ iff $x_1 < y_1$ and $x_2 \leq y_2$; (v) $x \succsim y$ iff $\max\{x_1, x_2\} \geq \max\{y_1, y_2\}$; (vi) $x \succsim y$ iff $x_1 \geq y_1$ and $x_2 \leq y_2$.

Answer. (i) The preference relation $x \succsim y$ iff (if and only if) $x_1 \leq y_1$ and $x_2 \leq y_2$ is not complete. Consider the following counter example: $x = (0, 1)$ and $y = (1, 0)$. Clearly, neither $x_i \geq y_i$ for all i nor $y_i \geq x_i$ for all i . So, neither $x \succsim y$ nor $y \succsim x$. Hence, the bundles $x = (0, 1)$ and $y = (1, 0)$ can not be compared. The preference relation $x \succsim y$ iff (if and only if) $x_1 \leq y_1$ and $x_2 \leq y_2$ is however transitive. Pick $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $z = (z_1, z_2)$ and suppose that $x \succsim y$ and $y \succsim z$ towards showing that $x \succsim z$. By assumption, $x \succsim y$ then $x_i \leq y_i$ for all i and since $y \succsim z$, $y_i \leq z_i$ for all i . That is,

$$x_1 \leq y_1 \text{ and } x_2 \leq y_2$$

and

$$y_1 \leq z_1 \text{ and } y_2 \leq z_2$$

Hence,

$$x_1 \leq z_1 \text{ and } x_2 \leq z_2$$

Therefore, $x \succsim y$ and $y \succsim z$ imply that $x \succsim z$.

(ii) Preferences defined by $x \succsim y$ iff (if and only if) $x_1 \geq y_1$ and $x_2 \geq y_2$ are not complete. Consider the following counter example: $x = (0, 1)$ and $y = (1, 0)$. Clearly, neither $x_i \geq y_i$ for all i nor $y_i \geq x_i$ for all i . So, neither $x \succsim y$ nor $y \succsim x$. Hence, the bundles $x = (0, 1)$ and $y = (1, 0)$ can not be compared. On the other hand, preferences defined by $x \succsim y$ iff (if and only if) $x_1 \geq y_1$ and $x_2 \geq y_2$ are transitive. Pick $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $z = (z_1, z_2)$ and suppose that $x \succsim y$ and $y \succsim z$. Need to show that this implies $x \succsim z$. By assumption, Since $x \succsim y$ then $x_i \geq y_i$ for all i and since $y \succsim z$ then $y_i \geq z_i$ for all i . That is,

$$x_1 \geq y_1 \text{ and } x_2 \geq y_2$$

and

$$y_1 \geq z_1 \text{ and } y_2 \geq z_2$$

Hence,

$$x_1 \geq z_1 \text{ and } x_2 \geq z_2$$

Therefore, $x \succsim y$ and $y \succsim z$ imply that $x \succsim z$. Note that this is called *Pareto preference relation*.

(iii) Preferences defined by $x \succsim y$ iff $x_1 > y_1$ or $x_1 = y_1$ and $x_2 > y_2$ are not complete. For a counter example pick two bundles x and y such that $x = y$. For example, $x = (1, 1)$ and $y = (1, 1)$. Clearly, since $x_1 = y_1$ and $x_2 = y_2$ neither $x \succsim y$ nor $y \succsim x$. Hence, the two bundles $x = (1, 1)$ and $y = (1, 1)$ can not be compared by this preference relation. On the other hand, preferences defined by $x \succsim y$ iff $x_1 > y_1$ or $x_1 = y_1$ and $x_2 > y_2$ are transitive. Pick $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $z = (z_1, z_2)$ and suppose that $x \succsim y$ and $y \succsim z$. Need to show this implies $x \succsim z$. By assumption, Since $x \succsim y$ then

$$\text{either } x_1 > y_1 \text{ or if } x_1 = y_1 \text{ then } x_2 > y_2$$

and since $y \succsim z$ then

$$\text{either } y_1 > z_1 \text{ or if } y_1 = z_1 \text{ then } y_2 > z_2$$

Hence, it must hold that

$$\text{either } x_1 > z_1 \text{ or if } x_1 = z_1 \text{ then } x_2 > z_2$$

which implies that $x \succsim z$. Note this is called *Lexicographic preference relation*.

(iv) The preference relation $x \succsim y$ iff (if and only if) $x_1 < y_1$ and $x_2 \leq y_2$ is not complete. Consider the following counter example: $x = (0, 1)$ and $y = (1, 0)$. Although $x_1 < y_1$, $x_2 > y_2$. Therefore $x = (0, 1)$ and $y = (1, 0)$ can't be compared. However, it is transitive. Pick $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $z = (z_1, z_2)$ and suppose that $x \succsim y$ and $y \succsim z$ towards showing that $x \succsim z$. By assumption, $x \succsim y$ then $x_1 < y_1$ and $x_2 \leq y_2$ and since $y \succsim z$, $y_1 < z_1$ and $y_2 \leq z_2$. That is,

$$x_1 < y_1 \text{ and } x_2 \leq y_2$$

and

$$y_1 < z_1 \text{ and } y_2 \leq z_2$$

Hence,

$$x_1 < z_1 \text{ and } x_2 \leq z_2$$

Therefore, $x \succsim y$ and $y \succsim z$ imply that $x \succsim z$.

(v) The preference relation $x \succsim y$ iff $\max\{x_1, x_2\} \geq \max\{y_1, y_2\}$ is complete. Pick any $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Clearly, either

$$\max\{x_1, x_2\} \geq \max\{y_1, y_2\}$$

holds or,

$$\max\{x_1, x_2\} \leq \max\{y_1, y_2\}$$

holds or both. Hence, either $x \succsim y$ or $y \succsim x$ or both. It is also transitive. Pick any $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $z = (z_1, z_2)$ and suppose that $x \succsim y$ and $y \succsim z$. To show that transitivity we need that $x \succsim z$. Since $x \succsim y$

$$\max\{x_1, x_2\} \geq \max\{y_1, y_2\}$$

and since $y \succsim z$

$$\max\{y_1, y_2\} \geq \max\{z_1, z_2\}$$

So, we conclude that

$$\max\{x_1, x_2\} \geq \max\{z_1, z_2\}$$

which implies that $x \succsim z$. Therefore, $x \succsim y$ and $y \succsim z$ imply that $x \succsim z$.

(vi) is left to the reader.

Problem 2.8.5 For the following choice environments and preference relations, is the relation complete and transitive: (i) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $\max(x_1, x_2) \geq \max(y_1, y_2)$; (ii) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $x_1 \geq y_1$; (iii) $X \subseteq \mathbb{R}$, $x \succeq y$ if x is a rational number and y is not; (iv) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $x_1 > y_2$; (v) $X \subseteq \mathbb{R}$, $x \succeq y$ iff $x - y \geq 2$; (vi) $X \subseteq \mathbb{R}^3$, $x \succeq y$ iff $x_i > y_i$ for 2 of the 3 i ; (vii) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $x_1 + x_2 = y_1 y_2$; (viii) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $\sqrt{(x_1 - 3)^2 + (x_2 - 7)^2} \geq \sqrt{(y_1 - 3)^2 + (y_2 - 7)^2}$; (ix) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $x_1 + x_2 \geq y_1 + y_2$ and $x_1 x_2 \geq y_1 y_2$; (x) $X \subseteq \mathbb{R}^2$, $x \succeq y$ iff $x_1 + x_2 \geq y_1 + y_2$ or $x_1 x_2 \geq y_1 y_2$; (xi) $X \subseteq \mathbb{R}_{++}^2$, $x \succeq y$ iff $x_1 + x_2 \geq y_1 + y_2$ and $x_1 x_2 \geq y_1 y_2$.

Problem 2.8.6 For the following preferences over bundles containing both apples and oranges, determine whether or not there is a utility representation, and if there is, write down two utility functions that represent the preferences: (i) the DM prefers the bundle with the most apples; (ii) the DM prefers the bundle with the most items of fruit; (iii) the DM prefers the bundle that is closest to having the same number of apples and oranges; (iv) the DM flips a coin: if it comes up heads, he prefers the bundle with the most apples; if it comes up tails he prefers the bundle with the most oranges.

Problem 2.8.7 Consider the following examples of choice sets X and associated preference relations \succeq : (i) $X = \{1, 2, 3\}$ and \succeq is defined by

$$\begin{aligned} 1 &\succeq 1, 1 \succeq 2, 1 \succeq 3 \\ 2 &\succeq 3 \\ 3 &\succeq 1; \end{aligned}$$

(ii) $X = \{\text{all the people in the world}\}$ and \succeq is defined by

$$\begin{aligned} &\text{for any } x, y \in X \\ &x \succeq y \text{ if } x \text{ "shares at least one given name with" } y; \end{aligned}$$

(iii) $X = R$ and \succeq is defined by

$$\begin{aligned} &\text{for any } x, y \in X \\ &x \succeq y \text{ if } x \geq y; \end{aligned}$$

(iv) $X = R$ and \succeq is defined by

$$\begin{aligned} &\text{for any } x, y \in X \\ &x \succeq y \text{ if } |x - y| > 1; \end{aligned}$$

(v) $X = R$ and \succeq is defined by

$$\begin{aligned} &\text{for any } x, y \in X \\ &x \succeq y \text{ if } |x - y| \text{ is an integer multiple of } 2. \end{aligned}$$

Show the following: (i) is NOT complete and NOT transitive; (ii) is NOT complete and NOT transitive; (iii) is complete and transitive; (iv) is NOT complete and NOT transitive; (v) is NOT complete but is transitive. Furthermore, [Note that this part of the problem goes beyond the difficulty level of this course and is only for those who are interested] consider the following potential properties of the above preference relations: **Weakly complete:** \succeq on X is weakly complete if for all $x, y \in X$, either $x = y$, $x \succeq y$ or $y \succeq x$; **Negatively transitive:** \succeq on X is negatively transitive if for all $x, y, z \in X$ such that $x \not\succeq y$ and $y \not\succeq z$, it is the case that $x \not\succeq z$; **Reflexive:** \succeq on X is reflexive if for all $x \in X$, $x \succeq x$. **Irreflexive:** \succeq on X is irreflexive if for all $x \in X$, $x \not\succeq x$. **Symmetric:** \succeq on X is symmetric if for all $x, y \in X$ such that $x \succeq y$, it is the case that $y \succeq x$. **Asymmetric:** \succeq on X is asymmetric if

for all $x, y \in X$ such that $x \succeq y$, it is the case that $y \not\succeq x$. **Antisymmetric:** \succeq on X is antisymmetric if for all $x, y \in X$ such that $x \succeq y$ and $y \succeq x$, it is the case that $x = y$; **Acyclic:** \succeq on X is acyclic if for all $x_1, \dots, x_n \in X$, for $n \in \mathbb{N}$, such that $x_1 \succeq x_2 \succeq \dots \succeq x_{n-1} \succeq x_n$, then it is the case that $x_1 \neq x_n$. For each of the examples (i)-(v), show whether or not the preference relation has any of the properties listed above. That is, if the preference relation has the property, provide a proof; if not, provide a counter-example.

Problem 2.8.8 The DM cannot detect small differences. He consumes two goods: x and y and he strictly prefers bundle (x', y') to bundle (x, y) , $(x', y') \succ (x, y)$, if and only if

$$x'y' - xy > 1$$

(if the difference is less than one in absolute value he is indifferent). (i) Show that the strict part of his preferences is transitive, i.e., that if $(x_1, y_1) \succ (x_2, y_2)$ and $(x_2, y_2) \succ (x_3, y_3)$ then $(x_1, y_1) \succ (x_3, y_3)$. (ii) Show that the indifference part of the preferences is not transitive, i.e., that if $(x_1, y_1) \sim (x_2, y_2)$ and $(x_2, y_2) \sim (x_3, y_3)$ then it is not necessary that $(x_1, y_1) \sim (x_3, y_3)$.

Answer. (i) Consider three bundles (x_1, y_1) , (x_2, y_2) and (x_3, y_3) . If $(x_1, y_1) \succ (x_2, y_2)$, then it follows that $x_1y_1 - x_2y_2 > 1$. If $(x_2, y_2) \succ (x_3, y_3)$, then $x_2y_2 - x_3y_3 > 1$. Summing up these two inequalities we obtain that $x_1y_1 - x_3y_3 > 2$, which implies that $x_1y_1 - x_3y_3 > 1$. Notice now that this last inequality holds only if $(x_1, y_1) \succ (x_3, y_3)$. (ii) Consider the following counterexample: (x_1, y_1) , (x_2, y_2) and $(x_3, y_3) : x_1y_1 = 1$, $x_2y_2 = 1.8$ and $x_3y_3 = 2.2$. Observe that $(x_1, y_1) \sim (x_2, y_2)$ since $x_2y_2 - x_1y_1 = 0.8 < 1$. At the same time, $(x_2, y_2) \sim (x_3, y_3)$ since $x_3y_3 - x_2y_2 = 0.4 < 1$. However, $(x_3, y_3) \succ (x_1, y_1)$ since $x_3y_3 - x_1y_1 = 1.2 > 1$.

Problem 2.8.9 Consider the preference ordering over the composition of Congress induced by the following procedure: I flip a coin; if heads I strictly prefer a Congress with a Democratic majority; if tails I strictly prefer a Congress with a Republican majority. (i) Is this preference relationship complete? (ii) Is it transitive?

Chapter 3

Consumer choice



Alfred Marshall

John Hicks

Consider an agent choosing her consumption of goods 1 and 2 for a given budget and recall his choice problem in the form introduced in the previous chapter

choose x to maximize $u(x)$ subject to $x \in X$ and $px \leq m$.

This is the workhorse of microeconomic theory. We shall study it in detail.

3.1 The Budget constraint

There are only two goods, $X = \mathbb{R}_+^2$.¹ A consumer who consumes x_1 units of good 1 and x_2 units of good 2 is said to consume the consumption bundle $(x_1, x_2) \in X$. Any bundle can be represented by a point on a two-dimensional graph.

The prices of these goods, p_1 and p_2 , are given and known to the consumer who is a *price taker*. The consumer's income, that is, the amount of money the consumer has to spend, is m . Then the consumer's *budget constraint* can be written as $p_1x_1 + p_2x_2 \leq m$.

The *budget set* consists of all bundles that are affordable at a given prices and income. More precisely,

$$B = \{x \in X : p_1x_1 + p_2x_2 \leq m\}$$

Thus, the *budget line* is the set of bundles that cost exactly m :

$$p_1x_1 + p_2x_2 = m,$$

and can be written also as:

$$x_2 = -\frac{p_1}{p_2}x_1 + \frac{m}{p_2}.$$

Note that this is a linear function with a vertical intercept m/p_2 , a horizontal intercept m/p_1 and a slope of $-p_1/p_2$; as graphed, in Figure 3.1.

The slope of the budget line measures the rate at which the market is willing to exchange good 1 for good 2. Thus, economists say that the slope measures the *opportunity cost* of consuming good 1: the consumer has an alternative opportunity to consuming 1 unit of good 1, consisting in selling the unit of good 1, for an amount equal to p_1 , and then using the proceeds to buy good 2 - that is $\frac{p_1}{p_2}$ units of good 2.

Clearly, when prices and income changes, the set of goods that a consumer can afford changes as well. Increasing income shifts the budget line outward, while increasing the price of good 1 (resp. good 2) makes the budget line steeper (resp. flatter). Make sure you understand this. What do you think happens to the budget line when both prices are changed at the same time? For example, both prices become t times as large. Give a graphical and an analytical answer.

¹The reader should note that extending the analysis to n goods, $X = \mathbb{R}_+^n$, is straightforward, though the graphical representation of the problem is obviously lost.

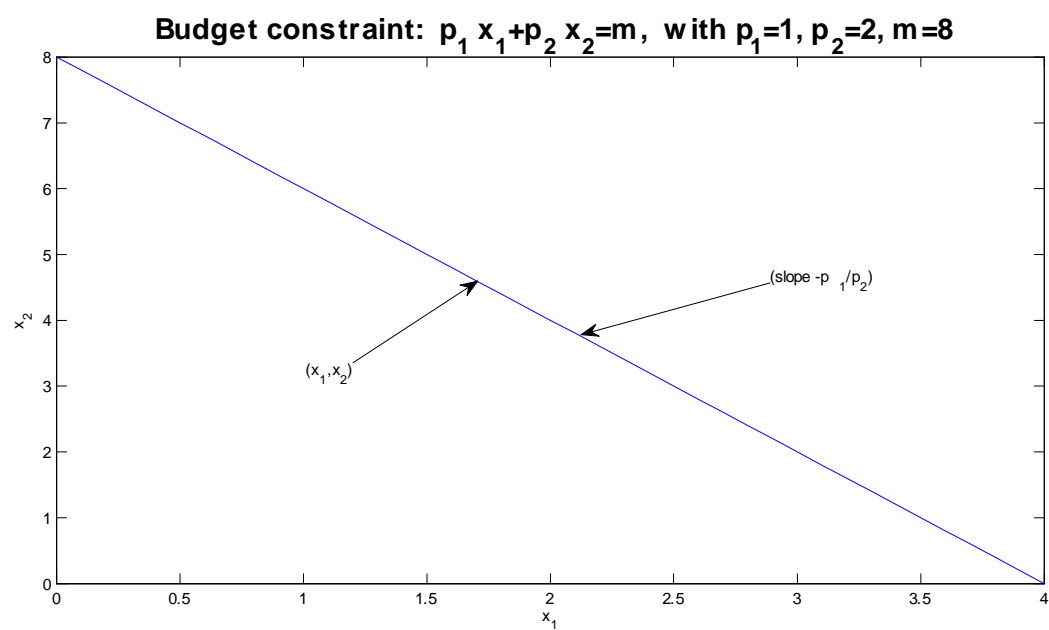
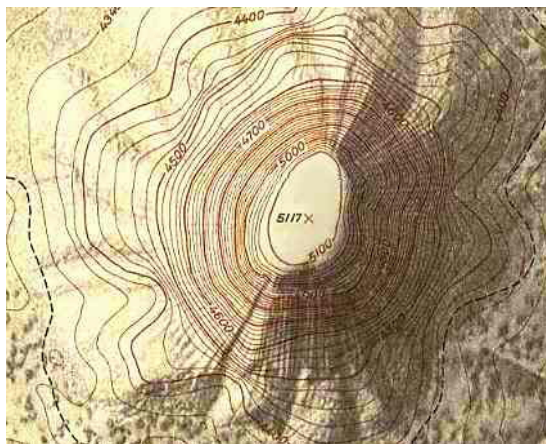


Figure 3.1: Budget constraint

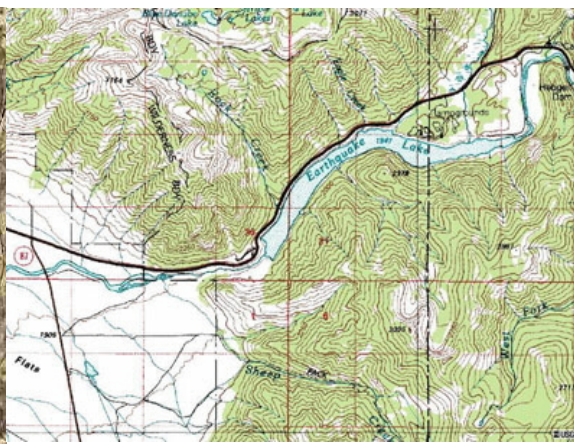
3.2 Indifference curves

Recall that the theory of consumer choice can be formulated by preferences that satisfy completeness and transitivity. A useful graphical description of preferences exploits a construction known as *indifference curves*.

Indifference curves are *level curves*, that is, they represent *the set of* (x_1, x_2) *for which the utility is constant at some level, say* U . This is not unlike topographic maps, as in the figures below.



A peak



Earthquake Lake, Montana - Idaho

Formally, the indifference curve at level U is represented by the equation

$$u(x_1, x_2) = U.$$

Can indifference curves representing distinct levels of utility cross ?

Note that the equation $u(x_1, x_2) = U$ is non-linear (since $u(x_1, x_2)$ is non-linear, in general) in (x_1, x_2) . As a consequence, a solution in the form $x_2 = f(x_1, u)$ is often impossible to write down (even when it exists). Nonetheless we can characterize the slope of the indifference curve, known as the *marginal rate of substitution*, by relying on the *Implicit Function theorem* of calculus.² The *marginal rate of substitution*, $MRS(x_1, x_2)$, is precisely defined as the

²The *Implicit Function Theorem* is one of the fundamental theorems of calculus, and one of the most used in economics. A (imprecise) statement of the theorem is as follows.

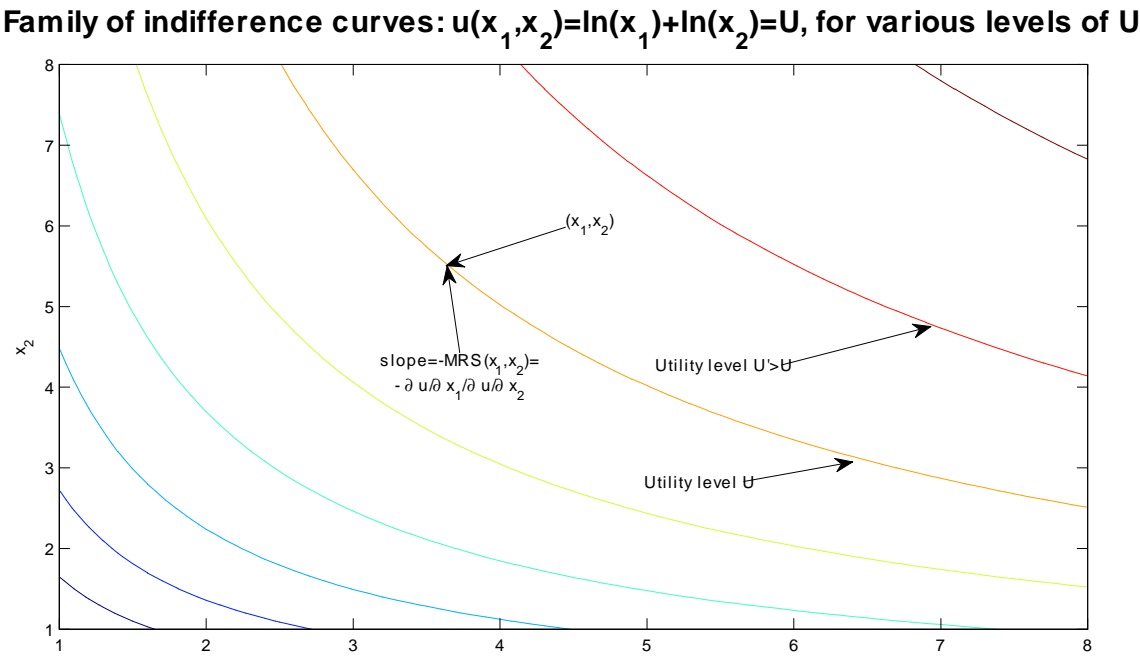


Figure 3.2: Indifference curves and the marginal rate of substitution

infinitesimal change in good 2 that keeps the utility unchanged after a unitary infinitesimal increase in good 1.

Formally, consider a change (dx_2, dx_1) that keeps the level of utility constant:

$$du = \frac{\partial u(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial u(x_1, x_2)}{\partial x_2} dx_2 = 0.$$

Rearranging, we obtain

$$\frac{dx_2}{dx_1} = - \frac{\frac{\partial u(x_1, x_2)}{\partial x_1}}{\frac{\partial u(x_1, x_2)}{\partial x_2}}.$$

But $\frac{dx_2}{dx_1}$ is exactly the rate at which the consumer is willing to substitute an infinitesimal amount of good 2 for good 1, that is, the change of x_2 for x_1 which keeps the agent's utility constant. That is,

$$MRS(x_1, x_2) = - \frac{\frac{\partial u(x_1, x_2)}{\partial x_1}}{\frac{\partial u(x_1, x_2)}{\partial x_2}}$$

Consider the Cobb-Douglas utility function $u(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, for $0 < \alpha < 1$, as an example. In this case the marginal rate of substitution can be easily solved for:

$$\begin{aligned} MRS(x_1, x_2) &= - \frac{\frac{\partial u(x_1, x_2)}{\partial x_1}}{\frac{\partial u(x_1, x_2)}{\partial x_2}} = \\ &= - \frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{\beta x_1^\alpha x_2^{-\alpha}} = - \frac{\alpha}{1-\alpha} \frac{x_2}{x_1}. \end{aligned}$$

In fact, in this case we can even solve for the indifference curve:

$$x_1^\alpha x_2^{1-\alpha} = U \iff x_2 = U^{\frac{1}{1-\alpha}} (x_1)^{\frac{1-\alpha}{\alpha}}.$$

Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuous and smooth (at least twice differentiable). Then there exists a continuous differentiable function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that, locally,

$$u(x_1, f(x_1)) = U$$

and

$$\frac{df(x_1)}{dx_1} = - \frac{\frac{\partial u(x_1, x_2)}{\partial x_1}}{\frac{\partial u(x_1, x_2)}{\partial x_2}}$$

At this stage we shall proceed by describing some general assumptions - *monotonicity* and *convexity* - which we will typically make about preferences. We shall also illustrate the implications of these assumptions for the associated indifference curves.

3.2.1 Monotonicity

We typically assume that more is better, that is, we consider *goods*, not *bads*. More precisely, consider (x_1, x_2) as a bundle of goods and let (x'_1, x'_2) be any other bundle with at least as much of both goods and more of one. That is, $x'_1 \geq x_1$ and $x'_2 \geq x_2$ with at least one strict inequality, or, in short, $(x'_1, x'_2) > (x_1, x_2)$.

(Strict) *monotonicity of preferences* requires that

$$\text{if } (x'_1, x'_2) > (x_1, x_2) \text{ then } (x'_1, x'_2) \succ (x_1, x_2).$$

Thus, (strict) monotonic preferences imply that more of (resp. less) of both goods is a better (resp. worse) bundle. Monotonicity of preferences implies that the utility function which represents preferences is monotonic increasing in both its arguments:

$$\frac{\partial u(x_1, x_2)}{\partial x_1}, \frac{\partial u(x_1, x_2)}{\partial x_2} > 0, \text{ for any } (x_1, x_2) \in X.$$

It also implies that indifference curves have a negative slope:

$$-MRS(x_1, x_2) < 0.$$

Obviously, since the bundle (x_1, x_2) was chosen arbitrarily, we can draw an indifference curve through any bundle.

Starting at an arbitrary bundle (x_1, x_2) and moving up and to the right, by monotonicity we must be at a preferred position. Thus, at a higher indifference curve the consumer is strictly better.

3.2.2 Convexity

A set X is *convex* if $\alpha x + (1 - \alpha)y \in X$ whenever $x, y \in X$ and $\alpha \in [0, 1]$. In words, A set X is convex if whenever it contains two elements x, y , it

Family of indifference curves: $u(x_1, x_2) = \ln(x_1) + \ln(x_2) = U$, for various levels of U

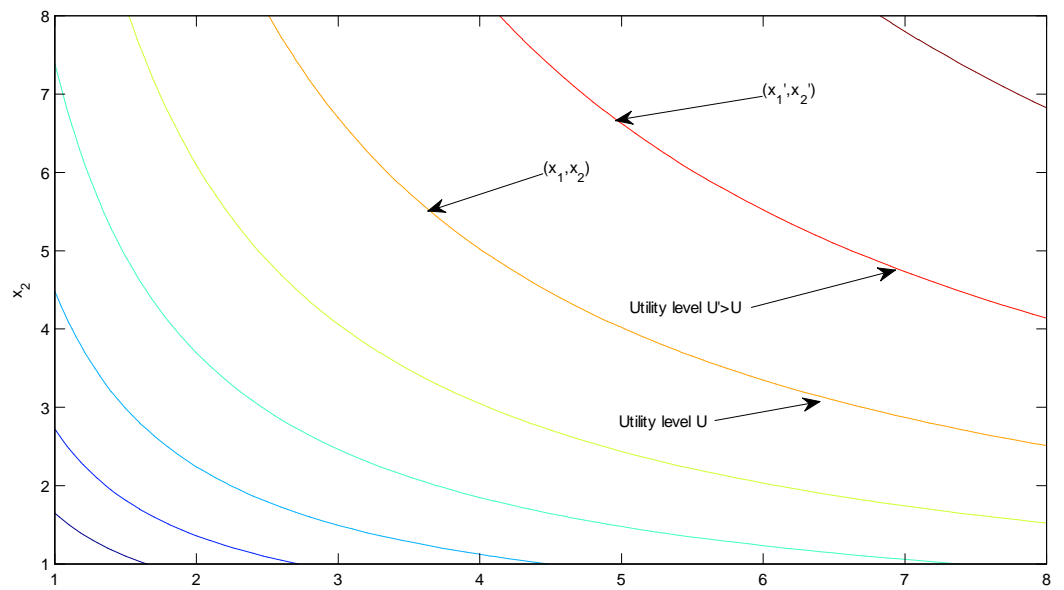
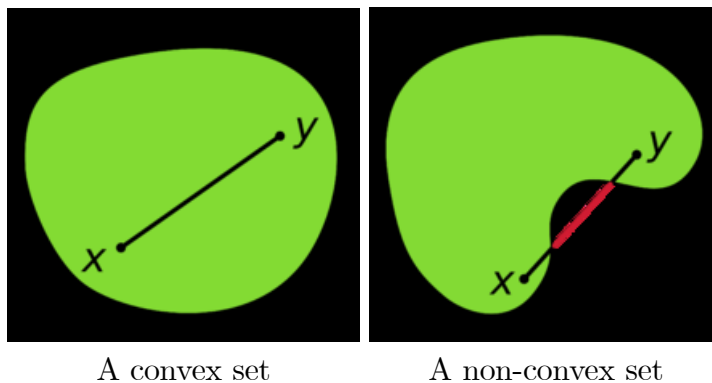


Figure 3.3: Indifference curves and monotonicity

contains the entire segment connecting them.



Convexity of preferences requires that the set of weakly preferred allocations to any given allocation is convex. Formally,

for any $(x_1, x_2) \in X$, the set $\{(y_1, y_2) \in X : (y_1, y_2) \succeq (x_1, x_2)\}$ is convex.

Taking two arbitrary bundles on the same indifference curve, (x_1, x_2) and (x'_1, x'_2) , by convexity the bundle $\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ is preferred to each of the initial bundles. *Convexity of preferences* implies that the utility function representing preferences is concave and that indifference curves are convex. Important examples where convexity does not hold include indivisible goods (e.g., shoes) or addictive goods. Can you see why ?

Often, *monotonicity* and *convexity* of preferences are taken as the defining features for *well-behaved indifference curves*. We shall always assume so in this book.

3.3 Types of preferences

Some examples to get more used with the indifference curve representation of preferences.

Family of indifference curves: $u(x_1, x_2) = \ln(x_1) + \ln(x_2) = U$, for various levels of U

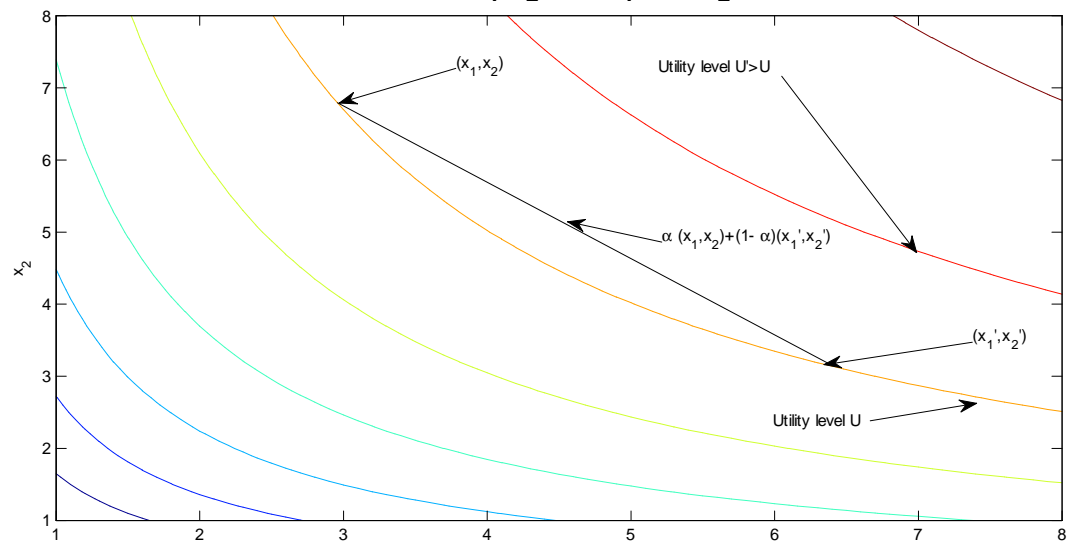


Figure 3.4: Indifference curves and convexity

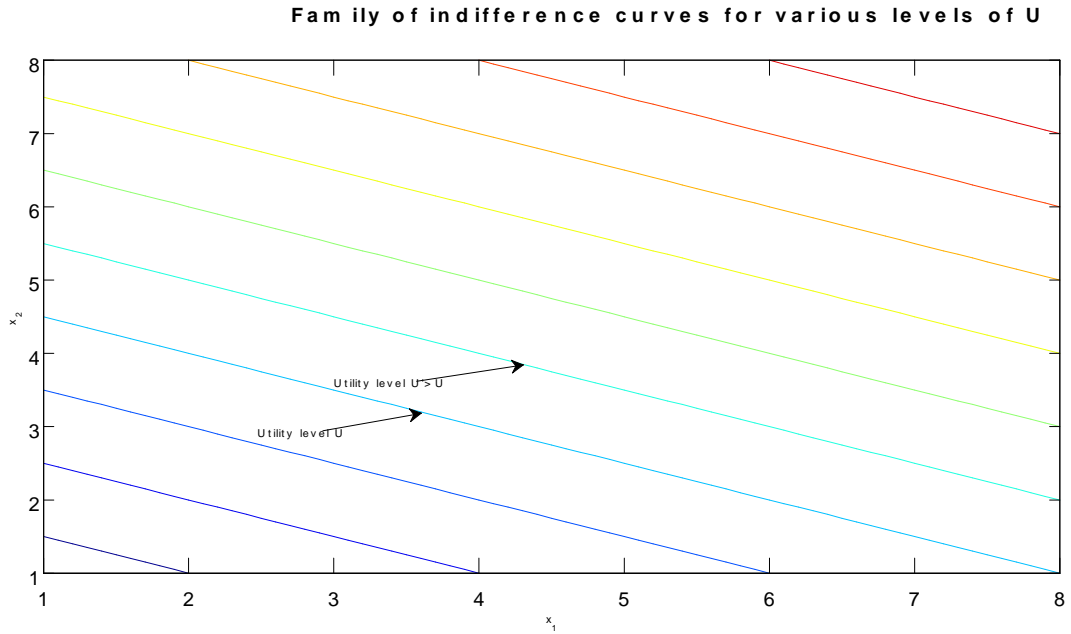


Figure 3.5: Indifference curves for perfect substitutes

3.3.1 Perfect substitutes

Two goods are perfect substitutes if the consumer is willing to substitute one good for the other at a constant rate. Thus, the indifference curves have a constant slope.

The utility function in the case x_1 and x_2 are perfect substitutes is of the (linear) form

$$u(x_1, x_2) = \alpha x_1 + \beta x_2,$$

and indifference curves are also linear.

3.3.2 Perfect complements

Two goods are perfect complements if they are always consumed together in fixed proportions. For example, right shoes and left shoes. The utility function in the case x_1 and x_2 are perfect complements is called Leontief,

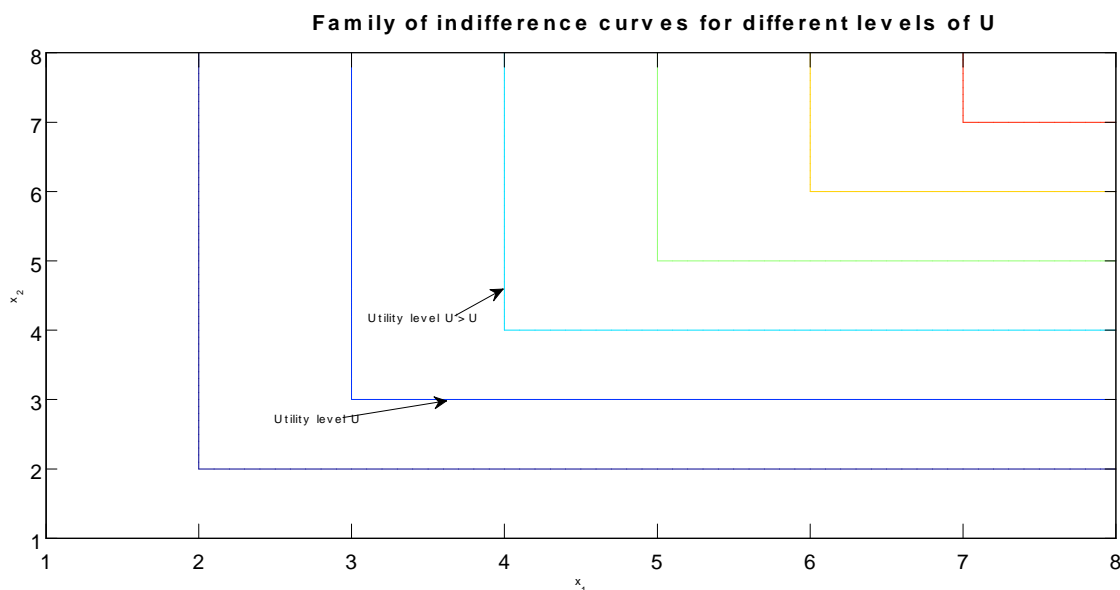


Figure 3.6: Indifference curves for perfect complements

after Wassily Leontief. It is of the form

$$u(x_1, x_2) = \min \{ \alpha x_1, \beta x_2 \}$$

and the associated indifference curves are *L-shaped*.

3.4 Consumer problem

We are ready to solve now the consumer problem:

$$\max_{x \in X} u(x_1, x_2) \text{ s.t. } p_1 x_1 + p_2 x_2 = m;$$

under the assumption that the preferences represented by $u(x_1, x_2)$ are well-behaved: *monotonic* and *convex*. In fact we consider three different ways to study the problem of consumer. We restrict the analysis to the example of Cobb-Douglas preferences to simplify the algebra.

The analysis in this section requires an understanding of the mathematical theory of maximization. The reader is strongly suggested to study the Mathematical Appendix on this topic (or any calculus book). We shall take a shortcut in the analysis, however, in the interest of simplicity: we shall only derive the *first order conditions* (*foc's*) for the maximum in the consumer problem (or an equivalent appropriately manipulated version of it). This is a shortcut because in principle, the *foc's* do not guarantee a maximum - formally, they are necessary but not sufficient conditions. Essentially three problems may arise: (i) the maximum is at a corner (not an interior solution), (ii) the *foc's* characterize a local - not a global - maximum, (iii) the *foc's* characterize a minimum - not a maximum. Possible corner solutions need to be checked individually. The possibility of the *foc's* not representing a maximum is avoided when the maximization problem is convex, that is, a set of *second order conditions* are satisfied. We refer to the Math Appendix. The reader should however check that in fact, in the consumer problem as we formulated, the solution we obtain by the *foc's* is a (global) maximum.

1. Geometric analysis of indifference curves.

A consumption bundle (x_1^*, x_2^*) is an optimal choice for the consumer if the set of bundles that the consumer prefers to (x_1^*, x_2^*) - the set of bundles above the indifference curve through (x_1^*, x_2^*) - has an empty intersection with the bundles she can afford - the bundles beneath her budget line.

It follows that at (x_1^*, x_2^*) the indifference curve is tangent to the budget line:

$$MRS(x_1^*, x_2^*) = -\frac{\alpha x_2^*}{\beta x_1^*} = -\frac{p_1}{p_2};^3$$

and, naturally, at (x_1^*, x_2^*) the budget constraint must be satisfied:

$$p_1 x_1^* + p_2 x_2^* = m$$

We have now two equations in two unknowns that can be solved for the optimal bundle. Thus, substituting:

³Note that, clearly, the above holds for an interior optimum and not for a boundary optimum. Can you see why a *Cobb-Douglas* preferences induce an interior solution? How would you deal with preferences implying perfect substitutes ?

$$p_1 x_1^* + p_2 \frac{\beta p_1 x_1^*}{\alpha p_2} = m$$

and hence

$$x_1^* = \frac{\alpha}{\alpha + \beta} \frac{m}{p_1}$$

This is the demand function for good 1. Similarly, the demand function for good 2:

$$x_2^* = \frac{\beta}{\alpha + \beta} \frac{m}{p_2}$$

2. *Algebraic analysis by substituting the budget constraint in to the objective function.*

After the substitution, the consumer problem is reduced to:

$$\max_{x_1 \geq 0} \alpha \ln x_1 + \beta \ln \left(\frac{m}{p_2} - \frac{p_1}{p_2} x_1 \right)$$

The *foc* for this problem is

$$\frac{\alpha}{x_1^*} - \beta \frac{p_2}{m - p_1 x_1^*} \frac{p_1}{p_2} = 0.$$

By a (very) little algebra,

$$x_1^* = \frac{\alpha}{\alpha + \beta} \frac{m}{p_1}$$

$$x_2^* = \frac{\beta}{\alpha + \beta} \frac{m}{p_2}.$$

3. *General algebraic analysis of constrained maximization.*

Define the *Lagrangian*:

$$L = \alpha \ln x_1 + \beta \ln x_2 - \lambda(p_1 x_1 + p_2 x_2 - m)$$

Differentiate it to obtain the *foc*'s for a maximum of the Lagrangian with respect to (x_1, x_2) and a minimum with respect to λ :

$$\frac{\partial L}{\partial x_1} \Big|_{x_1=x_1^*, x_2=x_2^*, \lambda=\lambda^*} = \frac{\alpha}{x_1^*} - \lambda^* p_1 = 0$$

$$\frac{\partial L}{\partial x_2} \Big|_{x_1=x_1^*, x_2=x_2^*, \lambda=\lambda^*} = \frac{\beta}{x_2^*} - \lambda^* p_2 = 0$$

$$\frac{\partial L}{\partial \lambda} \Big|_{x_1=x_1^*, x_2=x_2^*, \lambda=\lambda^*} = m - p_1 x_1^* - p_2 x_2^* = 0$$

These are three equations with three unknowns. The best way to proceed (which you should do!) is to first solve for λ^* . You will see that you get back to the algebraic condition obtained in the first method, the geometric analysis of indifference curves.

3.5 Demand Functions

Demand functions give the optimal bundle (amounts of each of the goods) as a function of the prices and income faced by the consumer

$$x_l^*(p, m), \quad l = 1, 2.$$

Note that when two goods are *perfect substitutes*, the consumer is willing to substitute the goods on a one to one basis, and $p_2 > p_1$ ($p_2 < p_1$) then the slope of the budget line is flatter (steeper) than of the indifference curve. Thus, the demand function for good 1 will be

$$x_1^* = \begin{cases} m/p_1 & \text{if } p_2 > p_1 \\ [0, m/p_1] & \text{if } p_2 = p_1 \\ 0 & \text{if } p_2 < p_1 \end{cases}$$

How is the demand function for good 1 if the goods are *perfect complements*? Does the demand of a good l depend necessarily negatively on its own price?

A few definitions which exploit the demand function construction are useful to remember.

A *Normal good* is a good l for which the quantity demanded increases with income,

$$\frac{\partial x_l^*(p, m)}{\partial m} > 0.$$

An *Inferior good* is a good for which the quantity demanded decreases with income,

$$\frac{\partial x_l^*(p, m)}{\partial m} < 0.$$

An *Engle curve* is a graph of the demand for a good as a function income, with all prices held constant, that is, a graph with quantity of good demanded on the horizontal axis and income on the vertical axis.

An *Income expansion path* is a graph of the bundles of goods 1 and 2 that are demanded at different levels of income, keeping prices held constant.

3.6 Problems

Problem 3.6.1 Let p_1, p_2, m be the set of prices and income. The utility function is given by:

$$u(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}.$$

(i) Determine the demanded bundle as a function of the prices, p_1 and p_2 , the income, m , and the parameter α . (ii) What fraction of his income will a consumer with this utility function spend on good 1? Does this fraction depend on income m ? Does it depend on the price ratio, $\frac{p_1}{p_2}$? (iii) Is good 2 a normal good? Explain.

Answer. (i) To derive the consumer's demand for good 1, x_1^* , and for good 2, x_2^* , given prices p_1 and p_2 and income m , we need to solve the consumer maximization problem:

$$\begin{aligned} & \max_{(x_1, x_2) \in \mathbb{X}} x_1^\alpha x_2^{1-\alpha} \\ \text{s.t. } & p_1 x_1 + p_2 x_2 \leq m \end{aligned}$$

Note that $X = \mathbb{R}_+^2$ and hence the restrictions on (x_1, x_2) are equivalent to the following three inequality constraints

$$\begin{aligned} x_1 & \geq 0 \\ x_2 & \geq 0 \\ p_1 x_1 + p_2 x_2 & \leq m \end{aligned}$$

But we can simplify the problem. By the nature of the utility function, it is never optimal to have either $x_1^* = 0$ or $x_2^* = 0$ (implicitly assuming finite

prices and $m > 0$). Thus the first two inequality constraints will not bind and can be dropped from the maximization problem.

Since the utility function is strictly increasing in both arguments and "savings" do not derive any utility, the last inequality will bind for sure, i.e. $p_1x_1 + p_2x_2 = m$ (to prove this, suppose that you had an optimal choice that did not use all of the consumer's income, then find another bundle that gives higher utility and is still in the budget set - this is *proof by contradiction*).

The maximisation problem is now:

$$\begin{aligned} & \max_{(x_1, x_2) \in \mathbb{R}_+^2} x_1^\alpha x_2^{1-\alpha} \\ \text{s.t. } & p_1x_1 + p_2x_2 = m \end{aligned}$$

Set up the Lagrangian

$$L = x_1^\alpha x_2^{1-\alpha} + \lambda(p_1x_1 + p_2x_2 - m)$$

where λ is the *Lagrange Multiplier*.

The *foc's* for an (interior) solution to this problem are:

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 0 \\ \frac{\partial L}{\partial x_2} &= 0 \\ \frac{\partial L}{\partial \lambda} &= 0 \end{aligned}$$

where the optimum choices x_1^* and x_2^* solve these three equations *simultaneously* (also note that there are three unknowns since the value of λ is unknown).

This gives the following equations:

$$\begin{aligned} \alpha(x_1^*)^{\alpha-1}(x_2^*)^{1-\alpha} + \lambda p_1 &= 0 \\ (1 - \alpha)(x_1^*)^\alpha(x_2^*)^{-\alpha} + \lambda p_2 &= 0 \\ p_1x_1 + p_2x_2 &= m \end{aligned}$$

The suggested way of solving this system of simultaneous equations is to re-arrange the first two equations so that λ is separated on the right hand side. Then equating these two gives (dropping the "*" notation for the time

being but still understanding that these are the optimal values rather than generic values of x_1 and x_2):

$$\begin{aligned}\frac{\alpha}{p_1} \left(\frac{x_1}{x_2}\right)^{\alpha-1} &= -\lambda = \frac{(1-\alpha)}{p_2} \left(\frac{x_1}{x_2}\right)^{\alpha} \\ \implies \frac{x_1}{x_2} &= \frac{\alpha}{(1-\alpha)} \frac{p_2}{p_1} \\ x_1 &= \frac{\alpha}{(1-\alpha)} \frac{p_2}{p_1} x_2 \\ &\text{and} \\ p_1 x_1 + p_2 x_2 &= m\end{aligned}$$

Putting these two equations (in two unknowns) together gives the solution:

$$\begin{aligned}x_1^* &= \alpha \frac{m}{p_1} \\ x_2^* &= (1-\alpha) \frac{m}{p_2}.\end{aligned}$$

(ii) The solution implies that

$$\begin{aligned}p_1 x_1^* &= \alpha m \\ p_2 x_2^* &= (1-\alpha)m;\end{aligned}$$

so that the fraction of income spent on good 1 is α and on good 2 is $1-\alpha$.

(iii) The solution also implies that

$$\frac{\partial x_2^*}{\partial m} = \frac{(1-\alpha)}{p_2} > 0,$$

and hence good 2 is a *normal good*.

Problem 3.6.2 Solve consumer problems of consumers with the following utility functions. (i) $u(x_1, x_2, x_3) = x_1^\alpha x_2^\beta x_3^{1-\alpha-\beta}$; (ii) $u(x_1, x_2) = 3 \log x_1 + 2 \log x_2$; (iii) $u(x_1, x_2) = \min(x_1, x_2)$; (iv) $u(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}}$; (v) $u(x_1, x_2) = x_1 + x_2^{\frac{3}{4}}$; (vi) $u(x_1, x_2) = x_1 + x_2$. In each case you will have to construct the budget constraint, assuming that any good x_i is available at price p_i , and there is m amount of money available

Answer. We just solve a few of these.

(iv) $u(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}}$. If the solution to the consumer problem is interior⁴, then the following equilibrium condition has to hold

$$\frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}} = \frac{p_1}{p_2} \quad (3.1)$$

In this case this condition is given by

$$\frac{\frac{1}{\rho} (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}-1} \rho x_1^{\rho-1}}{\frac{1}{\rho} (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}-1} \rho x_2^{\rho-1}} = \left(\frac{x_1}{x_2} \right)^{\rho-1} = \frac{p_1}{p_2}$$

implying that

$$x_1 = \left(\frac{p_1}{p_2} \right)^{\frac{1}{\rho-1}} x_2$$

Plugging this expression into the budget constraint, we obtain

$$\begin{aligned} p_1 x_1 + p_2 x_2 &= m \\ p_1 \left(\frac{p_1}{p_2} \right)^{\frac{1}{\rho-1}} x_2 + p_2 x_2 &= m \\ \Rightarrow x_2 &= \frac{m}{p_1 \left(\frac{p_1}{p_2} \right)^{\frac{1}{\rho-1}} + p_2}, \quad x_1 = \frac{m}{p_1 + p_2 \left(\frac{p_1}{p_2} \right)^{\frac{1}{1-\rho}}} \end{aligned}$$

(vi) $u(x_1, x_2) = x_1 + x_2$. Notice that if $\frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}} = \frac{1}{1} > \frac{p_1}{p_2}$, the utility per dollar spent in good 1 is higher than the utility per dollar spent in good 2, irrespective of the values of x_1 and x_2 . Hence, in that case, the individual will only consume good 1. Similarly, in case $1 < \frac{p_1}{p_2}$, the individual only demands good 2. When $1 = \frac{p_1}{p_2}$, any allocation on the budget constraint is optimal.

Having this in mind, we can write the demand function as

$$(x_1, x_2) = \begin{cases} x_1 = \frac{m}{p_1} & x_2 = 0 & \text{if } 1 > \frac{p_1}{p_2} \\ x_1 = 0 & x_2 = \frac{m}{p_2} & \text{if } 1 < \frac{p_1}{p_2} \\ \{(x_1, x_2) \in \mathbb{R}_+^2 : p_1 x_1 + p_2 x_2 = m\} & \text{if } 1 = \frac{p_1}{p_2} \end{cases}$$

⁴Notice that:

- (i) when $\rho = 1$ the indifference curves are linear
- (ii) as $\rho \rightarrow 0$, this utility function comes to represent the same preferences as $u(x) = x_1 x_2$
- (iii) as $\rho \rightarrow -\infty$, indifference curves become "right angles" as with $u(x_1, x_2) = \min\{x_1, x_2\}$

Problem 3.6.3 *Let the utility function be:*

$$u(x_1, x_2) = \sqrt{x_1} + x_2$$

(If you think there is something strange here because the utility is linear in x_2 , you are right! But trust me, do not worry and proceed as usual). (i) Write down the budget constraint: call income m and normalize the price p_1 , that is, $p_1 = 1$. Determine the demanded bundle as a function of price p_2 and income, m . (Substitution works; Lagrange too; whatever is easier for you). (ii) Draw the demand for good x_2 and the demand for x_1 as a function of price p_2 . (Now is the moment to think about the fact that the utility function is linear in x_2 . Does your solution generate always non-negative quantities x_1 and x_2 ?).

Answer. (i) The budget constraint is

$$x_1 + p_2 x_2 \leq m$$

To be formal, the consumer chooses $(x_1, x_2) \in \mathbb{R}_+^2$ to

$$\begin{aligned} & \max\{\sqrt{x_1} + x_2\} \text{ s.t.} \\ & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_1 + p_2 x_2 \leq m \end{aligned}$$

where we have written non-negativity constraints explicitly.

Since $u(x_1, x_2) = \sqrt{x_1} + x_2$ is an increasing function in x_1, x_2 the constraint $x_1 + p_2 x_2 \leq m$ will always bind. Thus always have

$$x_1 + p_2 x_2 = m$$

However, note that for this utility function we cannot say anything about the non-negativity constraints.

The procedure is to analyze the problem in two stages. First assume the solution is interior, i.e., solve the problem of choosing $(x_1, x_2) \in \mathbb{R}_+^2$ to

$$\begin{aligned} & \max\{\sqrt{x_1} + x_2\} \text{ s.t.} \\ & x_1 + p_2 x_2 = m \end{aligned}$$

ignoring the non-negativity constraints. Then look for corner solutions and for conditions on the parameters of the problem (p_2, m) that give rise to such problems.

There are two methods for solving this problem: the substitution method and the Lagrangian method.

Substitution Method. Assume $x_1 > 0$, $x_2 > 0$. Since we have the budget constraint holding with equality and only two choice variables, rearrange the budget constraint to solve for x_2 :

$$x_2 = \frac{m}{p_2} - \frac{1}{p_2}x_1$$

Putting this into the objective function gives, choose $x_1 \in \mathbb{R}_+$ to

$$\max \left\{ \sqrt{x_1} + \frac{m}{p_2} - \frac{1}{p_2}x_1 \right\}$$

The *foc* is:

$$\begin{aligned} \frac{1}{2} \frac{1}{\sqrt{x_1}} - \frac{1}{p_2} &= 0 \\ \frac{1}{\sqrt{x_1}} &= \frac{2}{p_2} \\ \sqrt{x_1} &= \frac{p_2}{2} \\ x_1 &= \left(\frac{p_2}{2} \right)^2 \end{aligned}$$

Putting this back into the expression for x_2 :

$$\begin{aligned} x_2 &= \frac{m}{p_2} - \frac{1}{p_2} \left(\frac{p_2}{2} \right)^2 \\ x_2 &= \frac{m}{p_2} - \frac{p_2}{4} \end{aligned}$$

Note that $x_2 \geq 0$ iff $\frac{m}{p_2} - \frac{p_2}{4} \geq 0$, and hence iff $4m \geq (p_2)^2$.

Now let's check for corner solutions. The solution obtained assuming $x_1 > 0$, $x_2 > 0$ suggests that $x_2 = 0$ could be a potential problem, since when $4m \leq (p_2)^2$ we have $x_2 \leq 0$. Postulating $x_2 = 0$ and solving the problem using the budget constraint, we obtain

$$x_1 = m.$$

This gives a corner solution of:

$$\begin{aligned}x_1 &= m \\x_2 &= 0.\end{aligned}$$

Lagrangian Method. Assume $x_1 > 0$, $x_2 > 0$. Set up the Lagrangian

$$L = \sqrt{x_1} + x_2 + \lambda(x_1 + p_2x_2 - m),$$

and its *foc*'s:

$$\begin{aligned}\frac{\partial L}{\partial x_1} &= \frac{1}{2\sqrt{x_1}} + \lambda = 0 \\&\implies \frac{1}{\sqrt{x_1}} = -2\lambda \\&\implies \sqrt{x_1} = \frac{1}{2(-\lambda)} \\ \frac{\partial L}{\partial x_2} &= 1 + \lambda p_2 = 0 \\&\implies -\lambda = \frac{1}{p_2} \\ \frac{\partial L}{\partial \lambda} &= x_1 + p_2x_2 - m = 0 \\&\implies x_1 + p_2x_2 = m\end{aligned}$$

Plugging in the solution for $-\lambda$ into the equation for $\sqrt{x_1}$ gives

$$\begin{aligned}\sqrt{x_1} &= \frac{p_2}{2} \\x_1 &= \left(\frac{p_2}{2}\right)^2\end{aligned}$$

Plugging this into the budget constraint gives

$$x_2 = \frac{m}{p_2} - \frac{p_2}{4};$$

and therefore $x_2 \geq 0$ iff $4m \geq (p_2)^2$.

Let's check for corner solutions. Note that, re-arranging the *foc*'s of the Lagrangian, we obtain

$$\frac{\frac{\partial u(x_1, x_2)}{\partial x_1}}{\frac{\partial u(x_1, x_2)}{\partial x_2}} = \frac{p_1}{p_2}.$$

The Lagrange method requires equating the slope of the indifference curves with the slope of the budget constraint. Corner solutions come about when the parameters of the problem (p_2, m) are such that the slope of the budget constraint is never equal to that of an indifference curve.

Note that $\lim_{x_1 \rightarrow 0} \frac{\partial x_2}{\partial x_1} = +\infty$. This means that we do not have a corner solution for $x_1 = 0$. However this is not the case for $x_2 = 0$. In fact

$$\frac{\partial x_2}{\partial x_1} = \frac{1}{2} \sqrt{\frac{1}{m}}.$$

When $p_2 \leq \frac{1}{2} \sqrt{\frac{1}{m}}$ we have a corner solution with $x_2 = 0$ and $x_1 = m$. Note finally that $p_2 \leq \frac{1}{2} \sqrt{\frac{1}{m}}$ can be equivalently written $p_2 \geq 2\sqrt{m}$; or, as before, $4m \leq (p_2)^2$.

(ii) We leave this to the reader. We note though that the demand for either of the goods, as a function of p_2 when m is kept constant, has a point of non-differentiability; but despite this, it is continuous.

Problem 3.6.4 *Let the utility function be:*

$$u(x_1, x_2, x_3) = (x_1)^\alpha (x_2)^\beta (x_3)^{1-\alpha-\beta}$$

Write down the budget constraint: call income m and normalize the price p_1 of good x_1 , that is, $p_1 = 1$. Determine the demanded bundle as a function of prices p_2, p_3 , and income, m .

Answer. We solve the following problem:

$$\begin{aligned} \max_{x_1, x_2, x_3} & (x_1)^\alpha (x_2)^\beta (x_3)^{1-\alpha-\beta} \\ \text{s.t.} & p_1 x_1 + p_2 x_2 + p_3 x_3 = m \end{aligned}$$

Write the Lagrangian and its *foc*'s to obtain:

$$\begin{aligned} \alpha (x_1)^{\alpha-1} (x_2)^\beta (x_3)^{1-\alpha-\beta} - \lambda p_1 &= 0 \\ \beta (x_1)^\alpha (x_2)^{\beta-1} (x_3)^{1-\alpha-\beta} - \lambda p_2 &= 0 \\ (1 - \alpha - \beta) (x_1)^\alpha (x_2)^\beta (x_3)^{-\alpha-\beta} - \lambda p_3 &= 0; \end{aligned}$$

and hence

$$\frac{\frac{\alpha x_1}{\beta x_2}}{\frac{\alpha x_1}{(1 - \alpha - \beta) x_3}} = \frac{\frac{p_1}{p_2}}{\frac{p_1}{p_3}}.$$

After some algebra, we obtain

$$(p_1 + \frac{\beta}{\alpha} p_1 + \frac{(1 - \alpha - \beta)}{\alpha} p_1) x_1 = m$$

Hence,

$$\begin{aligned} x_1 &= \frac{\alpha m}{p_1} \\ x_2 &= \frac{\beta m}{p_2} \\ x_3 &= \frac{(1 - \alpha - \beta) m}{p_3}. \end{aligned}$$

Problem 3.6.5 *Let the utility function be:*

$$u(x_1, x_2) = \sqrt{x_1} + (x_2)^\alpha (x_3)^{1-\alpha}, \quad 0 < \alpha < 1.$$

(i) Determine the demanded bundle as a function of the prices, p_1 , p_2 , p_3 and income, m . (ii) Draw the demand for good x_1 as a function of its price p_1 (taking p_2 , p_3 , and m as given). (iii) Let the utility function be instead:

$$u(x_1, x_2) = \frac{1}{2} \log x_1 + 2 \log x_2.$$

(It does not get any easier than this. But be careful and precise. Write down the budget constraint: call income m and normalize the price p_1 , that is, $p_1 = 1$. Determine the demanded bundle as a function of price p_2 and income, m . (substitution works; Lagrange too; whatever is easier for you). Derive the demand for good x_2 and the demand for x_1 as a function of price p_2 . A 10% increase in income, at prices $p_1 = p_2 = 1$, has a bigger effect on x_1 or on x_2 ? Try and compute the effect on x_1 and on x_2 .

Problem 3.6.6 *An agent chooses consumption bundle according to the following utility function: $U(x_1, x_2) = (x_1)^2 + x_1x_2$. (i) She has income $m = 20$, and prices for goods x_1 and x_2 are $p_1 = 3$ and $p_2 = 1$. How much of good x_1 and x_2 will she consume? (ii, HARD) What happens to the agent's consumption if firm X introduces 50% discount on its product, i.e., $p_1 = 1.5$?*

Answer. (i) Proceed as in the previous problems to obtain:

$$x_1 = \frac{2x_1 + x_2}{3}$$

which implies that $x_1 = x_2$. Substituting into the budget constraint, we have $x_1 = x_2 = 5$.

(ii) Proceed as in (i) to obtain

$$x_1 = \frac{2x_1 + x_2}{1.5},$$

which implies that $-0.5x_1 = x_2$. But since both x_1 and x_2 have to be non-negative, the only possible solution to this condition is $x_1 = x_2 = 0$. However, the budget constraint is not binding at this bundle and hence this bundle cannot be maximizing $u(x_1, x_2)$. Observe for instance that the bundle $x_1 = x_2 = 1$ is within the budget set and delivers a higher value of u . In fact, the second order conditions on the concavity of $u(x_1, x_2)$ are not satisfied at the bundle $x_1 = x_2 = 0$. There consumer problem then has no interior maximum, but rather a corner solution. Suppose the consumer consumes only good x_1 . In this case, $x_1 = 13.33$ and $u(13.33, 0) = 13.33^2$. Suppose alternatively the consumer consumes only good x_2 . In this case, $x_2 = 20$ and $u(0, 20) = 0$. We can conclude then that the solution is $x_1 = 13.33$, $x_2 = 0$.

Problem 3.6.7 *Consider an agent whose preferences over any couple $(x_1, x_2) \in \mathbb{R}_+^2$ of e.g., apples and oranges, is such that she prefers the bundle that is closest to having the same number of apples and oranges. Write a utility function $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ which represent these preferences.*

Answer. Let (x_1, x_2) denote a bundle with x_1 apples and x_2 oranges. The agent prefers the bundle that is closest to having the same number of apples and oranges,

$$(x_1, x_2) \succsim (y_1, y_2) \Leftrightarrow -(x_1 - x_2)^2 \geq -(y_1 - y_2)^2$$

Hence, $u(x_1, x_2) = -(x_1 - x_2)^2$ or $u(x_1, x_2) = k|x_1 - x_2|$, with $k < 0$, represent preferences \succsim .

Chapter 4

Equilibrium and efficiency



Leon Walras

Vilfredo Pareto

Ken Arrow

Gerard Debreu

Lionel McKenzie

We investigate the fundamental economic problem of allocation and price determination in a very simple economy. Our aim is to describe what outcomes might arise by giving individuals the opportunity to voluntarily exchange goods.

Thus, we will follow two simple principles:

- (i) *Rationality* - individuals choose the best patterns of consumption that are affordable for them, and
- (ii) *Equilibrium* - prices adjust such that the amount that people demand of some good is equal to the amount that is supplied.

We will determine equilibrium prices, equating demand and supply. We will show that these prices solve the allocation problem efficiently. Nonetheless the notion of equilibrium price is distinct from the notion of *value*, used

in common parlance; for instance in, "water is enormously valuable to each of us." While this is not commonly made precise, we implicitly interpret the *value* of a good as a measure of how difficult it is to do without it. Read your preferred rendition of King Midas' fable. The king understands the value of simple things when his golden touch makes them completely unavailable to him. There is no reason (it is a logical fallacy) to conclude that *values* and not *prices* are what the noble man/woman (or the social scientist) should care about. There is a long list of social scientists/philosophers which have fallen into a version of this fallacy.

4.1 The Economy

We shall consider a *pure exchange* economy: that is, an economy with no production, in which consumers have fixed endowments of goods. In our economy, markets are *competitive*: that is, consumers take prices as given, independently of their trading choices. We say therefore that consumers are *price takers*. Furthermore, we assume consumers are *rational*: that is, they choose the consumption bundle which maximizes their utility (and, as you know, they have a utility which represents their preferences to maximize if and only if they are rational).

These assumptions are fundamental in our analysis. For notational simplicity instead, and without loss of generality in fact, we assume that the economy is a 2×2 *economy*: that is, it is composed of two (types of) consumers who consume two (types of) goods.

4.1.1 Notation and Definitions

Let $\{1, 2\}$ denote the set of consumption goods. Let instead $\{A, B\}$ denote the set of consumers. Consumer A 's and consumer B 's initial endowments are denoted, respectively, $w^A = (w_1^A, w_2^A)$ and $w^B = (w_1^B, w_2^B)$. Similarly, $x^A = (x_1^A, x_2^A)$ and $x^B = (x_1^B, x_2^B)$ denote consumer A 's and consumer B 's consumption bundles respectively. Finally, Consumer A 's and consumer B 's preferences are represented, respectively, by utility functions $u^A(x_1, x_2)$ and $u^B(x_1, x_2)$.

Definition 4.1.1 *An allocation is a pair of consumption bundles, x^A and*

x^B . An allocation (x^A, x^B) is a feasible allocation if:

$$x_1^A + x_1^B = w_1^A + w_1^B$$

and

$$x_2^A + x_2^B = w_2^A + w_2^B$$

That is, if the total amount consumed of each of the goods is equal to the total amount available.

4.2 Pareto efficiency

It is important to be precise about which definition of efficiency we shall use. Suppose we could aggregate each different (type of) agent's preferences into a preference ordering which i) satisfies completeness and transitivity, and which ii) satisfies some general notion of fairness (e.g., it is in accordance with preferences of the majority). Then we could construct a utility function representing such preference ordering and require that it be maximized at an efficient allocation. If such a utility function would exist, we would refer to it as a *Social welfare function*. But we have learnt in the previous chapter (recall Ken Arrow's *Impossibility theorem*) that such a function does not exist in general.

Tough luck. We are then obliged to adopt a much weaker definition of efficiency. Here it is.

Definition 4.2.1 A feasible allocation (x^A, x^B) is Pareto-efficient¹ if there is no other feasible allocation (y^A, y^B) such that $y^A \succsim x^A$ and $y^B \succsim x^B$, with at least one \succ .

In words, an allocation is *Pareto efficient* if it is feasible and there is no other feasible allocation for which one consumer is at least as well off and the other consumer is strictly better off. This implies that at a Pareto efficient allocation (i) there is no way to make both consumers strictly better off, (ii) all of the gains from trade have been exhausted, that is, there are no mutually advantageous trades to be made.

Is a Pareto efficient allocation *fair*? Define fair and think about the answer.

¹In honor of Vilfredo Pareto (1848-1923)

4.2.1 The Social Planner problem

In this section will characterize the set of Pareto efficient allocations as the set of solutions to a maximization problem called the *Social Planning Problem*, (SP).

$$\max_{x^A, x^B} u^A(x_1^A, x_2^A) \quad (4.1)$$

subject to

$$\begin{aligned} x_1^A + x_1^B &= w_1^A + w_1^B \\ x_2^A + x_2^B &= w_2^A + w_2^B \\ u^B(x_1^B, x_2^B) &\geq \bar{u} \end{aligned}$$

Since both consumers' utility is strictly increasing in both goods 1 and 2, it will never be optimal to allocate to consumer B a higher utility than \bar{u} . Thus the final constraint will be binding (i.e., it can be written as an equality constraint).²

Notice that, for any given economy (defined by an endowment vector and a utility function for each type of agent) the set of all Pareto efficient allocations defined by the *Social planning* problem will be a precisely a *set*, not a single allocation. This is what you get by having a weak definition of efficiency: a lot of things (allocations) satisfy it!

The set of efficient allocations is obtained as the solution to the Social Planning problem by varying \bar{u} .

Convince yourself you understand this. It is as important a point as it is subtle. Once you are convinced, note that the *set* of all Pareto efficient allocations can also be obtained as the solution to the following *Modified Social planning* problem:

$$\max_{x_A, x_B} \theta u(x_1^A, x_2^A) + (1 - \theta) u(x_1^B, x_2^B) \quad (4.2)$$

subject to

$$\begin{aligned} x_1^A + x_1^B &= w_1^A + w_1^B \\ x_2^A + x_2^B &= w_2^A + w_2^B \end{aligned}$$

²To be precise, we should also impose non-negativity constraints on allocations:

$$x_A^1 \geq 0; x_A^2 \geq 0; x_B^1 \geq 0; x_B^2 \geq 0;$$

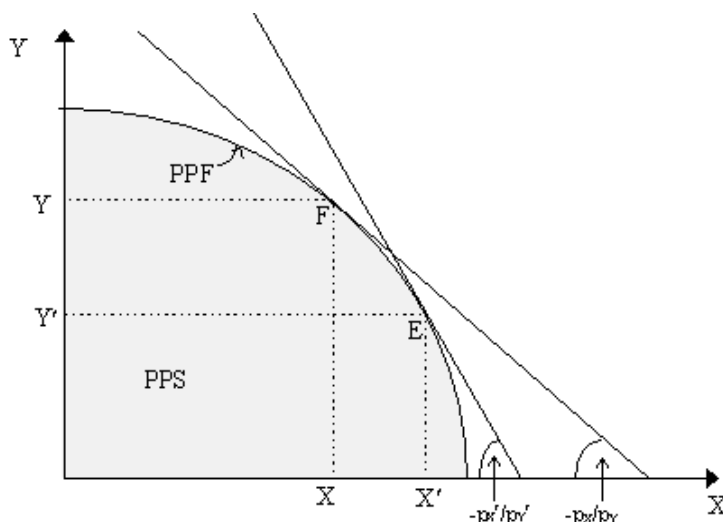


Figure 4.1: Pareto frontier (PPF). Note that, differently from the notation in the text, in this figure the two goods are denoted X and Y

Now the *set* of all Pareto efficient allocations is obtained by varying the relative weight of agent A in the planner's objective, that is, by varying θ between 0 and 1.

Can you prove that Social planning problem is equivalent to the Modified Social planning problem? Can you prove that they are equivalent to the definition of Pareto efficiency? Careful! It is not hard but requires some ability/practice to transform words into precise logical statements. Can a solution of the *Social planning* (resp. *Modified Social planning*) problem not satisfy the definition of *Pareto efficiency*? Can a *Pareto efficient* allocation not arise as a maximum of the *Social Planning* (resp. *Modified Social planning*) problem, that is, for no values of \bar{u} (resp. θ)?

Note that the (solution of) the Social planning problem does not depend on the whole endowment vector $(w_1^A, w_1^B, w_2^A, w_2^B)$, but only on $(w_1^A + w_1^B, w_2^A + w_2^B)$. In other words, the (solution of) the Social planning problem does not depend on the distribution of endowments across agents A and B , but rather only on the aggregate (economy-wide) endowments of the goods. This is a manifestation of the fact that, when the Social planner is choosing, agents have no property rights - the planner chooses how to distribute the endowments to agents for consumption as he/she pleases. This is not the

case in a Competitive equilibrium, which we'll define in turn.

4.3 Competitive equilibrium

Definition 4.3.1 A competitive or Walrasian³ equilibrium in a 2×2 economy is a pair of prices (p_1, p_2) and allocations (x^A, x^B) such that:

(x^A, x^B) are demanded by agents A and B at prices (p_1, p_2) ; and

markets clear:

$$\begin{aligned} x_1^A + x_1^B &= w_1^A + w_1^B \\ x_2^A + x_2^B &= w_2^A + w_2^B \end{aligned}$$

In words, in a competitive market / Walrasian equilibrium the total demand for each good should be equal to the total supply. Put differently, an equilibrium is a set of prices such that each consumer is choosing her most preferred (and affordable) bundle, and both consumers' choices are compatible in the sense that the total demand equals the total supply for each of the goods. Notice the combination of the two principles which guide our analysis: rationality and equilibrium.

Let's start and analyze the notion of *Competitive equilibrium*. What happens to the budget set if both prices, p_1 and p_2 are proportionally increased to λp_1 and λp_2 , for $\lambda > 0$? Note that your answer (well, the correct answer) implies that one price can always be normalized when looking at a competitive equilibrium, that is, $p_1 = 1$ without loss of generality. Note that this is equivalent to saying that only *relative prices*, like $\frac{p_2}{p_1}$, are determined at a competitive equilibrium. Furthermore, at a competitive equilibrium $(x^A, x^B, \frac{p_2}{p_1})$ prices satisfy (check this!):

$$\frac{p_2}{p_1} = \frac{\frac{\partial u^A(x_1^A, x_1^A)}{\partial x_2^A}}{\frac{\partial u^A(x_1^A, x_1^A)}{\partial x_1^A}} = \frac{\frac{\partial u^B(x_1^B, x_1^B)}{\partial x_2^B}}{\frac{\partial u^B(x_1^B, x_1^B)}{\partial x_1^B}} \quad (4.3)$$

That is, at a competitive equilibrium relative prices are equal to both agents' marginal rates of substitution *evaluated at the equilibrium allocation!* This

³In honor of Leon Walras (1834-1910).

implies, clearly, that the competitive equilibrium price of a good does not contain any information about the *value of this good*. And there is no reason why it should.

4.4 Welfare Economics

Is the competitive market mechanism Pareto efficient ? More precisely, Are *Competitive equilibrium* allocations Pareto efficient? Or, in other words, Can the competitive market mechanism really extract all the possible gains from trade? The answer is yes.

Theorem 4.4.1 (First Theorem of Welfare Economics) *Suppose preferences are monotonic. Then, all competitive equilibria are Pareto efficient.*

Proof. We proceed by contradiction. Suppose there exist a feasible allocation (y^A, y^B) that Pareto dominates (is weakly preferred by both agents to, and is strictly preferred by at least one to) the competitive equilibrium (x^A, x^B) . Then, the allocation (y^A, y^B) must be not budget feasible for at least one agent:

$$\begin{aligned} p_1(y_1^A - w_1^A) + p_2(y_2^A - w_2^A) &\geq 0 \\ p_1(y_1^B - w_1^B) + p_2(y_2^B - w_2^B) &\geq 0 \end{aligned}$$

with at least one strict inequality sign. Summing up:

$$p_1(y_1^A - w_1^A) + p_2(y_2^A - w_2^A) + p_1(y_1^B - w_1^B) + p_2(y_2^B - w_2^B) > 0 \quad (4.4)$$

By monotonicity of preferences, prices are positive (convince yourself of this). Then equation (4.4) implies that either

$$(y_1^A + y_1^B - w_1^A - w_1^B) > 0$$

or

$$(y_2^A + y_2^B - w_2^A - w_2^B) > 0$$

which contradicts feasibility of (y^A, y^B) . ■

The First Theorem of Welfare Economics says that all competitive equilibria are Pareto efficient. This is the formal argument behind Adam Smith *invisible hand*, in the 'Wealth of Nations,' (1776).

Is the converse true ? That is, Is any Pareto efficient allocation a competitive equilibrium for some endowments and prices ?

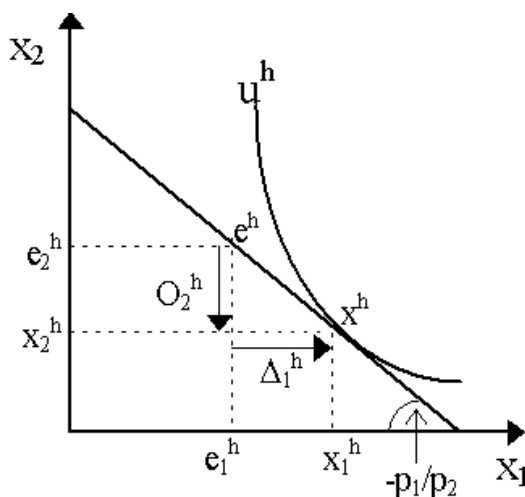


Figure 4.2: Note that, differently from the notation in the text, in this figure an arbitrary agent is denoted h and endowments are denoted with the letter e .

Theorem 4.4.2 (Second Theorem of Welfare Economics) *Suppose preferences are monotonic and have indifference curves which are strictly convex. Then any Pareto efficient allocation is a competitive equilibrium for some prices and endowments (w^A, w^B) such that*

$$\begin{aligned} w_1^A + w_1^B &= w_1^A + w_1^B \\ w_2^A + w_2^B &= w_2^A + w_2^B \end{aligned}$$

The *Second Welfare* theorem implies that all allocations in the *Pareto efficient* set (obtained by varying \bar{u} in the *Social planning* problem or by varying θ in the *Modified Social planning* problem) are obtained also by varying the distribution of aggregate endowments across (types of) agents.

4.4.1 The Edgeworth box

The Edgeworth box⁴ is a convenient graphical way to analyze the aspects of an economy with two consumers and two goods.

⁴In honor of Francis Edgeworth (1845-1926).

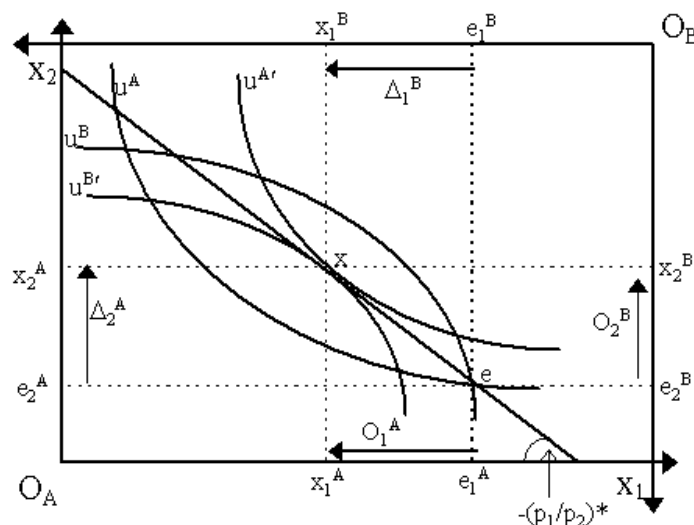


Figure 4.3: Edgeworth Box. Note that, differently from the notation in the text, in this figure endowments are denoted with the letter e .

The *Pareto efficient* set is constructed as the set of points (that is, the set of allocations) in the *Edgeworth box* at which the indifference curves of the two consumers are tangent (in the interior of the box). If indifference curves are not tangent, it must be that there exist some advantageous trade to explore and hence the corresponding allocation is not Pareto efficient.

Definition 4.4.3 *The set of all Pareto efficient allocations is called the contract curve.*

The proof that any competitive equilibrium is Pareto efficient (First Welfare Theorem) has a graphical representation in the *Edgeworth box*. The proof goes something like this: A feasible allocation (in the Edgeworth box) allocation is Pareto efficient if the intersection of consumer A 's strictly preferred set and consumer B 's strictly preferred set is empty (that is, if the indifference curves of the two consumers are tangent in the interior of the box). This is the case for a Competitive equilibrium, as in this case consumer A 's and consumer B 's strictly preferred sets cannot intersect since they lie on different sides of the same budget set line.

4.5 Externalities

The *First Welfare* theorem is broken if the economy we consider is plagued by *externalities*. In this case, the notion of *Competitive equilibrium* (the competitive market mechanism) does not select *Pareto efficient* allocations.

What are externalities? Here's two fundamental examples.

Preference externality. Consider an economy in which the utility of agent A depends also on the consumption choice of agent B ; e.g., it's:

$$u^A(x_1^A, x_2^A, x_1^B).$$

We say that the consumption of good 1 by agent B has an externality on agent A (positive or negative depending on whether $u^A(\cdot)$ is, respectively, increasing or decreasing on x_1^B).

Public good. Consider the case in which the utility of both agents depend on a *public good*, which is bought adding the contributions of both agents, that is:

$$u^A(x_1^A, x_2^A + x_2^B), u^B(x_1^B, x_2^B + x_2^A)$$

Examples of public good are bridges, parks, national defense, etc. They have the property of being *non-exclusive*, that is, an agent's consumption of it does not preclude other agents' consumption. You should immediately see that they introduce a (positive) externality of A on B and viceversa.

4.6 Problems

Problem 4.6.1 Read Chapter 1 of Adam Smith, *The Wealth of Nations*, first published in 1776; Edwin Cannan, ed., London: Methuen & Co., Ltd., 5th edition, 1904. You will find the text at

<http://www.econlib.org/library/Smith/smWN1.html#B.I,Ch.1,OftheDivisionofLabor>

Answer (one page max; typed) the following question: What does labor specialization have to do with the Invisible Hand and market efficiency?

Problem 4.6.2 Consider the following 2×2 economy (two consumers A and B and two goods 1 and 2): $w^A = (w_1^A, w_2^A) = (1, 0)$; $w^B = (w_1^B, w_2^B) = (1, 2)$; and the utility functions u^A and u^B for consumers A and B respectively are:

$$u^A(x_1^A, x_2^A) = x_1^A \cdot x_2^A$$

$$u^B(x_1^A, x_2^A) = x_1^B \cdot x_2^B$$

Find an algebraic formula for the the contract curve (Hint: write down and solve the Social Planner problem)? Consider again the above 2×2 economy. Solve for the competitive equilibrium (prices and allocations). (Hint: Solve for all the demand functions in the economy - how many are there? Impose feasibility and solve for equilibrium prices - how many prices do you have to solve for? Substitute prices in demand functions to find equilibrium allocations.)

Answer. An *allocation* is a list of consumption bundles, one for each consumer:

$$(x_1^A, x_2^A, x_1^B, x_2^B).$$

A *feasible allocation* is an allocation that satisfies

$$\begin{aligned} x_1^A + x_1^B &= w_1^A + w_1^B = 2 \\ x_2^A + x_2^B &= w_2^A + w_2^B = 2 \end{aligned}$$

i.e., satisfies the *resource constraints* of the economy. An allocation is *Pareto Optimal (PO)* if it is feasible and \nexists another allocation that Pareto Dominates it. The *contract curve* is the set of all Pareto Optimal allocations. Rewrite the contract curve as the set of solutions to the following a maximisation problem:

$$\max_{(x_1^A, x_2^A, x_1^B, x_2^B)} x_1^A \cdot x_2^A$$

subject to

$$\begin{aligned} x_1^A &\geq 0 \\ x_2^A &\geq 0 \\ x_1^B &\geq 0 \\ x_2^B &\geq 0 \\ x_1^A + x_1^B &= w_1^A + w_1^B = 2 \\ x_2^A + x_2^B &= w_2^A + w_2^B = 2 \\ x_1^B x_2^B &\geq \bar{u} \end{aligned}$$

Since both consumers' utility is strictly increasing in both goods 1 and 2, it will never be optimal to give consumer *B* more utility than \bar{u} . Thus the final constraint will be binding (i.e. an equality constraint).

Assuming an interior solution ($x_1^A > 0, x_2^A > 0, x_1^B > 0, x_2^B > 0$), construct the Lagrangian

$$L = x_1^A x_2^A + \lambda(x_1^B x_2^B - \bar{u}) + \mu_1(x_1^A + x_1^B - 2) + \mu_2(x_2^A + x_2^B - 2).$$

The f.o.c. of the associated maximization problem are:

$$\begin{aligned}\frac{\partial L}{\partial x_1^A} &= x_2^A + \mu_1 = 0 \implies x_2^A = -\mu_1 \\ \frac{\partial L}{\partial x_2^A} &= x_1^A + \mu_2 = 0 \implies x_1^A = -\mu_2 \\ \frac{\partial L}{\partial x_1^B} &= \mu_1 + \lambda x_2^B = 0 \implies \lambda x_2^B = -\mu_1 \\ \frac{\partial L}{\partial x_2^B} &= \mu_2 + \lambda x_1^B = 0 \implies \lambda x_1^B = -\mu_2\end{aligned}$$

Equating the equations with the same multiplier gives

$$\begin{aligned}x_2^A &= \lambda x_2^B \\ x_1^A &= \lambda x_1^B.\end{aligned}$$

Dividing the first equation by the second gives

$$\begin{aligned}\frac{x_2^A}{x_1^A} &= \frac{\lambda x_2^B}{\lambda x_1^B} \\ \frac{x_2^A}{x_1^A} &= \frac{x_2^B}{x_1^B}\end{aligned}$$

This condition is the marginal rate of substitution condition

$$-\frac{x_2^A}{x_1^A} = MRS_{12}^A(x_1^A, x_2^A) = MRS_{12}^B(x_1^B, x_2^B) = -\frac{x_2^B}{x_1^B}$$

Rewriting the two resource constraints we substitute out x_1^B, x_2^B

$$\begin{aligned}
 x_1^B &= 2 - x_1^A \\
 x_2^B &= 2 - x_2^A \\
 &\implies \\
 \frac{x_2^A}{x_1^A} &= \frac{2 - x_1^A}{2 - x_2^A} \\
 &\implies \\
 x_1^A &= x_2^A \\
 &\implies \\
 x_1^B &= x_2^B
 \end{aligned}$$

Using the last equality constraint

$$\begin{aligned}
 x_1^B x_2^B &= \bar{u} \\
 &\implies (x_1^B)^2 = \bar{u} \\
 &\implies x_1^B = \sqrt{\bar{u}}.
 \end{aligned}$$

This gives us the following allocation

$$\begin{aligned}
 x_1^A &= 2 - \sqrt{\bar{u}} \\
 x_2^A &= 2 - \sqrt{\bar{u}} \\
 x_1^B &= \sqrt{\bar{u}} \\
 x_2^B &= \sqrt{\bar{u}}
 \end{aligned}$$

As long as $\bar{u} \in (0, 4)$, the assumption of an interior solution is valid. If $\bar{u} = 0$

$$\begin{aligned}
 x_1^A &= x_2^A = 2 \\
 x_1^B &= x_2^B = 0
 \end{aligned}$$

and if $\bar{u} = 4$

$$\begin{aligned}
 x_1^A &= x_2^A = 0 \\
 x_1^B &= x_2^B = 2
 \end{aligned}$$

There does not exist a feasible allocation that can make $\bar{u} < 0$ or $\bar{u} > 4$.

At a competitive equilibrium in this economy, given prices (p_1, p_2) , the bundle (x_1^i, x_2^i) clears the markets:

$$\begin{aligned}x_1^A + x_1^B &= w_1^A + w_1^B = 2 \\x_2^A + x_2^B &= w_2^A + w_2^B = 2.\end{aligned}$$

Since both consumers have strictly increasing utility in both goods, the budget constraint will hold with equality. Assuming an interior solution, $(x_1^A > 0, x_2^A > 0, x_1^B > 0, x_2^B > 0)$, the solution to the consumer optimisation problem can be found constructing the Lagrangian

$$L = x_1^i x_2^i + \lambda(p_1 x_1^i + p_2 x_2^i - p_1 w_1^i - p_2 w_2^i).$$

The f.o.c. of the associated maximization problem are

$$\begin{aligned}\frac{\partial L}{\partial x_1^i} &= x_2^i + \lambda p_1 = 0 \implies \frac{x_2^i}{p_1} = -\lambda \\ \frac{\partial L}{\partial x_2^i} &= x_1^i + \lambda p_2 = 0 \implies \frac{x_1^i}{p_2} = -\lambda\end{aligned}$$

which imply:

$$\frac{x_2^i}{x_1^i} = \frac{p_1}{p_2}$$

and the budget constraint

$$p_1 x_1^i + p_2 x_2^i = p_1 w_1^i + p_2 w_2^i.$$

Normalising $p_2 = 1$, we obtain

$$\begin{aligned}x_2^i &= p_1 x_1^i \\ p_1 x_1^i + x_2^i &= p_1 w_1^i + w_2^i\end{aligned}$$

For $i = A$, substituting, we have $x_1^A = \frac{1}{2}$.

For $i = B$, substituting, we have $x_1^B = \frac{1}{2} + \frac{1}{p_1}$. Using market clearing for good 1

$$\frac{1}{2} + \left(\frac{1}{2} + \frac{1}{p_1}\right) = 2,$$

which implies $p_1 = 1$. To check that you have not made any mistakes, make sure the market clearing condition for good 2 holds with $p_1 = 1$:

$$\left(\frac{p_1}{2}\right) + \left(1 + \frac{p_1}{2}\right) = \frac{1}{2} + 1 + \frac{1}{2} = 2.$$

The Competitive Equilibrium price-allocation pair is then

$$p_1 = 1, x_1^A = \frac{1}{2}, x_2^A = \frac{1}{2}, x_1^B = \frac{3}{2}, x_2^B = \frac{3}{2}.$$

Note that our assumption that the solution is interior is valid.

Also note that the First Welfare theorem holds. That is that the Competitive Equilibrium is indeed Pareto efficient.

Problem 4.6.3 Consider the following 2×2 economy (two consumers A and B and two goods 1 and 2. The total endowments of the two goods are $w_1^A + w_1^B = 2$ and $w_2^A + w_2^B = 1$. The utility functions u^A and u^B for consumers A and B respectively are:

$$\begin{aligned} u^A(x_1^A, x_2^A) &= x_1^A \cdot x_2^A \\ u^B(x_1^B, x_2^B) &= x_1^B \cdot x_2^B. \end{aligned}$$

Suppose the weights of the planner for the two agents are $\theta^A = 1/4$ and $\theta^B = 3/4$. i) Find the Pareto efficient allocation for this economy? ii) Find the competitive equilibrium prices of the economy which has the Pareto optimal allocation you have found as a competitive equilibrium? [this might seem hard but it is not; think before jumping in the sea of computation] iii) Find the distribution of the endowments across agents of the economy which has the Pareto efficient allocation you have found as a competitive equilibrium? [this is hard and requires you indeed to jump]

Answer. i) The social planning problem is:

$$\max_{(x_1^A, x_2^A, x_1^B, x_2^B)} \frac{1}{4} x_1^A \cdot x_2^A + \frac{3}{4} x_1^B \cdot x_2^B$$

subject to

$$\begin{aligned} x_1^A + x_1^B &= w_1^A + w_1^B = 2 \\ x_2^A + x_2^B &= w_2^A + w_2^B = 1. \end{aligned}$$

Here we assumed non-negativity constraints on allocations are not binding, so we have an interior solution. (Later on we will see this assumption is not valid, so we will check the boundary solutions. However, for now we will solve the questions as if the solution is interior). The problem can be simplified in

the following way. Using $x_1^B = 2 - x_1^A$ and $x_2^B = 1 - x_2^A$, the social planner solves:

$$\max_{(x_1^A, x_2^A)} \frac{1}{4}x_1^A \cdot x_2^A + \frac{3}{4}(2 - x_1^A)(1 - x_2^A)$$

We now have an unconstrained optimization problem with two variables. The associated f.o.c.'s are

$$\frac{1}{4}x_2^A - \frac{3}{4}(1 - x_2^A) = 0 \quad (4.5)$$

and

$$\frac{1}{4}x_1^A - \frac{3}{4}(2 - x_1^A) = 0. \quad (4.6)$$

Hence, $x_1^B = 2 - x_1^A = \frac{1}{2}$ and $x_2^B = 1 - x_2^A = \frac{1}{4}$. We can plug these number back into the social planning welfare, to obtain $\frac{1}{4}x_1^A \cdot x_2^A + \frac{3}{4}(2 - x_1^A)(1 - x_2^A) = \frac{12}{32}$. It can be easily checked that this solution is not a maximum. In fact, the solution is at the boundary. It is not hard to see that it is optimal to give all endowment to agent B : $x^B = (2, 1)$; and nothing to agent A , $x^A = (0, 0)$.

ii) At the competitive equilibrium, price ratio has to be equal to the MRS of each agent. However, we have to be careful here since one agent has no endowment, and his indifference curve has a kink at $(x_1^A, x_2^A) = (0, 0)$. Therefore any price line would be tangent at this point. The only condition comes from agent B .

$$MRS^B = \frac{x_2^B}{x_1^B} = \frac{p_1}{p_2} \quad (4.7)$$

So price ratio is equal to

$$\frac{p_1}{p_2} = \frac{1}{2} \quad (4.8)$$

iii) Since this is a boundary point, The only endowment across agents of the economy which has this boundary Pareto efficient allocation as a competitive equilibrium is the Pareto optimal allocation itself. So, $x^B = (2, 1)$ and $x^A = (0, 0)$.

Consider a variation of the above question:

Consider the same 2×2 economy, with utility functions :

$$\begin{aligned} u^A(x_1^A, x_2^A) &= \log x_1^A + \log x_2^A \\ u^B(x_1^A, x_2^A) &= \log x_1^B + \log x_2^B \end{aligned}$$

Suppose the weights of the planner for the two agents are $\theta^A = 1/4$ and $\theta^B = 3/4$.

Assuming the non-negativity constraints on allocations are not binding, so we have an interior solution (later on we will see this assumption is valid), the proceed as above to obtain the following f.o.c.'s of the social planning problem:

$$\frac{1}{4x_1^A} - \frac{3}{4(2 - x_1^A)} = 0 \quad (4.9)$$

and

$$\frac{1}{4x_2^A} - \frac{3}{4(1 - x_2^A)} = 0. \quad (4.10)$$

Hence, $x_1^B = 2 - x_1^A = \frac{3}{2}$ and $x_2^B = 1 - x_2^A = \frac{3}{4}$. It can be checked that this allocation is the maximum. At the competitive equilibrium of this economy,

$$MRS^A = MRS^B = \frac{p_1}{p_2} \quad (4.11)$$

At the Pareto optimal allocation, $x^A = (\frac{1}{2}, \frac{1}{4})$ and $x^B = (\frac{3}{2}, \frac{3}{4})$, so that

$$MRS^A = \frac{1}{2}, MRS^B = \frac{1}{2}, \text{ and } \frac{p_1}{p_2} = \frac{1}{2}. \quad (4.12)$$

Note that we implicitly already shown that the Pareto efficient allocation can be supported as a competitive equilibrium: we know one point on the budget line, and we also know the slope of the budget line. So we can write down the equation for the budget line as:

$$(x_2^A - \frac{1}{4}) = -\frac{1}{2}(x_1^A - \frac{1}{2}) \quad (4.13)$$

Any point on this line, which satisfies the feasibility constraints, will give Pareto efficient allocation as the competitive equilibrium allocation at prices, $\frac{p_1}{p_2} = \frac{1}{2}$. Note that agent B 's endowment is given by: $x_1^B = 2 - x_1^A$ and $x_2^B = 1 - x_2^A$.

Problem 4.6.4 *Given a 2x2 economy with preferences and endowments*

$$\begin{aligned} u^A(x_1^A, x_2^A) &= 4 \log x_1^A + 5 \log x_2^A & w^A &= (3, 7) \\ u^B(x_1^B, x_2^B) &= 2 \log x_1^B + \log x_2^B & w^B &= (2, 4), \end{aligned}$$

compute the contract curve and a competitive equilibrium.

Problem 4.6.5 Below are utility functions and allocations for different economies. In each case, give an equation for the contract curve and solve for the competitive equilibrium. In each case show explicitly the three steps for solving for a competitive equilibrium (find demand functions, substitute into feasibility and solve for prices, then substitute back into demand function to solve for allocations:

- i) $u^A(x_1^A, x_2^A) = (x_1^A)^\alpha (x_2^A)^{(1-\alpha)}, \quad w^A = (3, 7),$
 $u^B(x_1^B, x_2^B) = (x_1^B)^\beta (x_2^B)^{(1-\beta)}, \quad w^B = (2, 4);$
- ii) $u^A(x_1^A, x_2^A, x_3^A) = 4 \log x_1^A + 5 \log x_2^A + \log x_3^A, \quad w^A = (3, 7),$
 $u^B(x_1^B, x_2^B) = 2 \log x_1^B + \log x_2^B, \quad w^B = (2, 4);$
- iii) $u^A(x_1^A, x_2^A) = 4 \log x_1^A + 5 \log x_2^A, \quad w^A = (2, 1, 0),$
 $u^B(x_1^B, x_2^B, x_3^B) = 3 \log x_1^B + \log x_2^B + 2 \log x_3^B, \quad w^B = (1, 1, 1);$
- iv) $u^A(x_1^A, x_2^A) = 4 \log x_1^A + 5 \log x_2^A, \quad w^A = (1, 2),$
 $u^B(x_1^B, x_2^B) = \min(x_1^B, 2x_2^B), \quad w^B = (2, 3);$
- iv) $u^A(x_1^A, x_2^A) = 3x_1^A + x_2^A, \quad w^A = (1, 2),$
 $u^B(x_1^B, x_2^B) = \min(x_1^B, 2x_2^B), \quad w^B = (2, 3);$

Problem 4.6.6 For the following economies, calculate the competitive equilibrium prices and allocations

- i) $u^A(x_1^A, x_2^A) = (x_1^A)^\alpha (x_2^A)^{(1-\alpha)}, \quad w^A = (1, 2),$
 $u^B(x_1^B, x_2^B) = (x_1^B)^\beta (x_2^B)^{(1-\beta)}, \quad w^B = (2, 1),$
 $u^C(x_1^C, x_2^C) = (x_1^C)^\gamma (x_2^C)^{(1-\gamma)}, \quad w^C = (2, 2);$
- ii) $u^A(x_1^A, x_2^A) = \log x_1^A + 4 \log x_2^A, \quad w^A = (4, 4),$
 $u^B(x_1^B, x_2^B) = 2 \log x_1^B + 3 \log x_2^B, \quad w^B = (2, 1),$
 $u^C(x_1^C, x_2^C) = 3 \log x_1^C + 2 \log x_2^C, \quad w^C = (2, 2);$

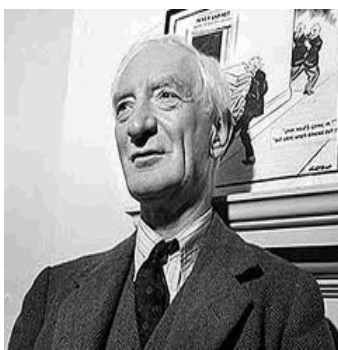
Problem 4.6.7 For the following economy, show that the First Welfare theorem doesn't hold (in other words, calculate the competitive equilibrium, calculate the set of Pareto efficient allocations and show that the competitive equilibrium is not Pareto efficient,

$$u^A(x_1^A, x_2^A) = (x_1^A)^\alpha (x_2^A)^{(1-\alpha)} + x_1^B, \quad w^A = (3, 7),$$

$$u^B(x_1^B, x_2^B) = (x_1^B)^\beta (x_2^B)^{(1-\beta)} + x_1^A, \quad w^B = (2, 1).$$

Chapter 5

Equilibrium in the labor market



William Beveridge



George Stigler



Peter Diamond



Dale Mortensen



Chris Pissaredes

5.1 Competitive labor markets

The consumer problem is:

$$\begin{aligned} \max_{c, l^s} & u(c) + v(1 - l^s) \\ \text{s.t. } & c \leq wl^s \end{aligned}$$

Note that the budget constraint can also be written as

$$c + w(1 - l^s) \leq w,$$

which stresses the interpretation of the wage rate w as the opportunity cost of leisure.

The firm problem is:

$$\max_{l^d} al^d - wl^d.$$

At equilibrium, market clearing requires:

$$\begin{aligned} l^d &= l^s = l \\ c &= al. \end{aligned}$$

A characterization of the competitive equilibrium includes:

$$\begin{aligned} w &= a \\ w &= \frac{\frac{\partial u(c)}{\partial c}}{\frac{\partial V(1-l)}{\partial(1-l)}}. \end{aligned}$$

5.2 Search frictions

Let l be workers searching for a job and let v be vacancies offered by firms. Matches are determined by an aggregate matching function $M(l, v)$ which satisfies

$$\begin{aligned} M(l, v) &\text{ is increasing in both } l \text{ and } v \\ M(l, v) &= p\left(\frac{l}{v}\right)l = q\left(\frac{l}{v}\right)v. \end{aligned}$$

Let $\theta = \frac{l}{v}$ denote the tightness of the labor market. The matching probability that a searching worker finds a job is

$$p(\theta) = \frac{M(l, v)}{l}$$

while the probability that a firm which posted a vacancy fills it is

$$q(\theta) = \frac{M(l, v)}{v}.$$

It is easily shown that the postulated properties of $M(l, v)$ imply that

$$\begin{aligned} p(\theta) &\text{ is decreasing in } \theta, \\ q(\theta) &\text{ is increasing in } \theta. \end{aligned}$$

A unit cost h is required to post a vacancy. A match between a worker and a firm produces a units of the consumption good. The relationship between the firm and the worker after matching is not competitive and hence the wage is the outcome of bargaining:

$$w = \sigma a,$$

where σ is a measure of the bargaining power of the worker.

The consumer problem is:

$$\begin{aligned} &\max_{c, l} p(\theta)u(c) + v(1 - l) \\ \text{s.t. } &c \leq wl \text{ and } w = \sigma a \end{aligned}$$

The firm problem is:

$$\begin{aligned} &\max_v q(\theta)(a - w)v - hv \\ \text{s.t. } &w = \sigma a \end{aligned}$$

At a search equilibrium,

$$\theta = \frac{l}{v}.$$

Note that both consumers and firms take as given the tightness of the market θ , when choosing their labor supply and their vacancies. At equilibrium then, their choices are consistent and $\theta = \frac{l}{v}$. A characterization of the search equilibrium includes the following conditions:

$$\begin{aligned} \sigma a p(\theta) \frac{\partial u(c)}{\partial c} - \frac{\partial v(1 - l)}{\partial (1 - l)} &= 0, \text{ where } c = \sigma a l \\ (1 - \sigma)q(\theta)a &= h \\ \theta &= \frac{l}{v}. \end{aligned}$$

5.2.1 Pareto inefficiency of search equilibrium

The search equilibrium is inefficient. The Social planner problem is:

$$\begin{aligned} & \max_{l,v} p\left(\frac{l}{v}\right)u(c) + v(1-l) \\ \text{s.t. } & c = wl, \\ & w = \sigma a \\ & q\left(\frac{l}{v}\right)(a-w) = h. \end{aligned}$$

The f.o.c.'s include

$$\begin{aligned} \sigma ap\left(\frac{l}{v}\right)\frac{\partial u(c)}{\partial c} + \frac{1}{v}\frac{\partial p\left(\frac{l}{v}\right)}{\partial \frac{l}{v}}u(c) - \frac{\partial v(1-l)}{\partial (1-l)} &= 0, \text{ where } c = \sigma al \\ (1-\sigma)q\left(\frac{l}{v}\right)a &= h. \end{aligned}$$

Compare with the equilibrium conditions, which can be written as:

$$\begin{aligned} \sigma ap\left(\frac{l}{v}\right)\frac{\partial u(c)}{\partial c} - \frac{\partial v(1-l)}{\partial (1-l)} &= 0, \text{ where } c = \sigma al \\ (1-\sigma)q\left(\frac{l}{v}\right)a &= h. \end{aligned}$$

Note that $\frac{1}{v}\frac{\partial p\left(\frac{l}{v}\right)}{\partial \frac{l}{v}}u(c) < 0$ at a competitive equilibrium, as $u(c) > 0$ (recall we normalized $u(0) = 0$) and $p\left(\frac{l}{v}\right)$ is decreasing in $\theta = \frac{l}{v}$. As a consequence, at a competitive equilibrium $\sigma ap\left(\frac{l}{v}\right)\frac{\partial u(c)}{\partial c} + \frac{1}{v}\frac{\partial p\left(\frac{l}{v}\right)}{\partial \frac{l}{v}}u(c) - \frac{\partial v(1-l)}{\partial (1-l)} < 0$ and an l such that $\left[\sigma ap\left(\frac{l}{v}\right)\frac{\partial u(c)}{\partial c} + \frac{\partial v(1-l)}{\partial (1-l)}\right] > 0$ would increase efficiency. By concavity of $p\left(\frac{l}{v}\right)u(c) + v(1-l)$ in l , this requires a smaller l , that is, a decrease in search.

5.3 Empirical implications of search equilibrium

At equilibrium l satisfies

$$(1-\sigma)q\left(\frac{l}{v}\right)a = h,$$

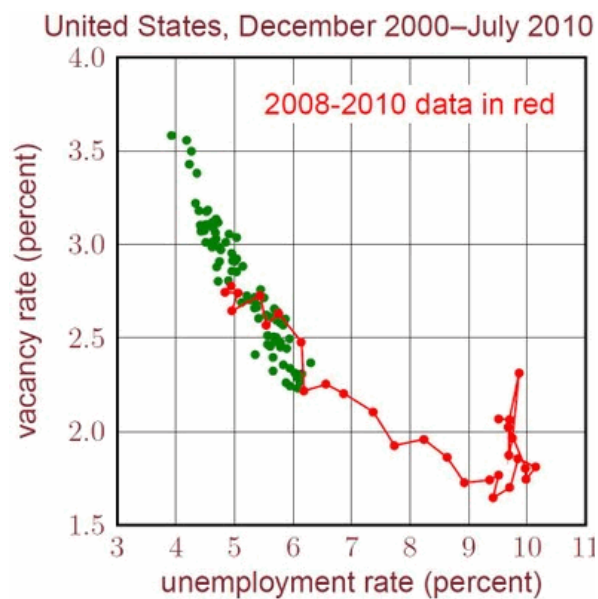


Figure 5.1: Beveridge Curve

and unemployment is

$$u = \left(1 - p\left(\frac{l}{v}\right)\right) l.$$

An equilibrium relationship between u and v can then be obtained, as exogenous variables, e.g., productivity a , vary. It can be shown that it satisfies $\frac{\partial p}{\partial v} < 0$. This relationship is called Beveridge curve, and in the U.S. economy it looks as in the Figure.

5.4 Problems

Problem 5.4.1 *Prove directly that the equilibrium allocations of the competitive labor market economy are Pareto efficient. (Hint: write a social planning problem and compare the f.o.c.'s of this problem with the equilibrium conditions.)*

Problem 5.4.2 *Consider the search friction economy with*

$$\begin{aligned}u(c) &= c, \\v(1-l) &= \log(1-l), \\M(l, v) &= l^\alpha v^{1-\alpha}.\end{aligned}$$

Characterize the search equilibrium as much as possible. Characterize also the solution of the Social planner problem in this economy and compare it with the equilibrium. Is unemployment larger at equilibrium or at the solution of the social planner problem?

Chapter 6

Time and uncertainty

In this section, by using the techniques that we have already learned, we examine two interesting choice problems:

- i) saving over time, and
- ii) allocating wealth to a portfolio of assets with risky returns.

In order to focus on saving and consumption over time and under uncertainty, we shall restrict our attention to economies with a single consumption good: an agent's consumption is then, in fact, appropriately interpreted as her total expenditure on consumption .

In this section, by using the techniques that we have already learned, we examine two interesting choice problems:

- i) saving over time, and
- ii) insuring risky income.

In order to focus on saving and consumption over time and under uncertainty, we shall restrict our attention to economies with a single consumption good: an agent's consumption is then, in fact, appropriately interpreted as her total expenditure on consumption .

6.1 Choice over time; a.k.a. Saving



Milton Friedman



Franco Modigliani

The economy evolves over 2 periods: t and $t + 1$, say the first and the second part of an agent's life. This is just for simplification, but you can think of an economy with any T periods to be interpreted as years, or quarters, even $T = \infty$. We denote the income that the consumer will have in each period by (m_t, m_{t+1}) , and the amount of consumption in each period by (c_t, c_{t+1}) . Income and consumption are measured in the same unit of account, say dollars. (Do not think about inflation yet; we'll discuss inflation later). Assume that the consumer can borrow and save money at some fixed interest rate $r \geq 0$. Don't be bothered by the fact that a single interest rate applies to borrowing and saving (a counterfactual in the real world). This is a model, remember, and the bid and ask spread on active and passive interest rates is a detail of financial markets which is irrelevant for our purposes, that is, to understand how savings are determined in a competitive economy.

6.1.1 Intertemporal preferences for consumption

The consumer preferences for consumption at time t and $t + 1$ are represented by the following utility function:

$$u(c_t) + \beta u(c_{t+1}), \quad 0 < \beta < 1$$

Note that we have assumed that *i*) preferences for consumption are the same in each period, and *ii*) future utilities are discounted at rate $\beta < 1$. Furthermore, we assume that the function $u : \Re \rightarrow \Re$ is differentiable and strictly

concave (that is, $\frac{d^2}{dc^2}u(c) < 0$, for any $c \geq 0$).¹ We shall also assume:

$$\lim_{c \rightarrow 0} u(c) = \infty$$

that is, the marginal utility of consumption for an agent not consuming anything is arbitrarily high (infinity).²

Those of you who are trying to think about how to write the economy when agents are infinitely lived, rather than just two periods as we are assuming, should know that in this case, we typically write the utility function as:

$$\sum_{t=0}^{\infty} \beta^t u(c_t), \quad 0 < \beta < 1$$

6.1.2 The intertemporal budget constraint

Let's construct now the budget constraint of agents. In fact we have two budget constraints, one for each period. Let the price of consumption in periods t and $t+1$ be the same (this amounts to assume there is no inflation; we'll deal with inflation later). We can then without loss of generality assume that $p_t = p_{t+1} = 1$ (this is just a normalization). We say that an agent saves at t if $m_t > c_t$, that is, if he consumes less than she can afford. Otherwise, if $m_t < c_t$, she borrows. Let S denote the amount the consumer takes to or from period 1. If $S > 0$ (resp. $S < 0$) we say the consumer saves (resp. borrows). The budget constraint at time t is then

$$c_t + S = m_t,$$

indicating that the consumer uses his income m_t either to consume or to save. The budget constraint in period $t+1$ is:

$$c_{t+1} = m_{t+1} + (1+r)S,$$

It indicates that the agent in the second period consumes her income plus the proceeds of his saving (minus the repayment of the borrowing, if $S < 0$).

¹Note that *strict concavity* is a more stringent assumption than *convexity of preferences* as required up to now. It'll be clear later on why we need concavity.

²This is for simplicity, and it can be easily relaxed.

We implicitly assumed that the agent cannot (is not allowed to) default on her obligations.³

Fortunately, we can combine the two budget constraints into one, called the *intertemporal budget constraint*, thereby transforming the choice problem of the agent into one which looks like (and is solved as) the ones we have studied in the previous chapter. Solve the system of budget constraints for S_t to have:

$$c_t + \frac{1}{1+r}c_{t+1} = m_t + \frac{1}{1+r}m_{t+1}$$

Consider consumption in period t and consumption in period $t+1$ as two distinct goods. Then $\frac{1}{1+r}$ represents the relative price of consumption at time $t+1$; that is, the amount of period t consumption that the agent has to give up to get an extra unit of consumption in period $t+1$. Thus, the maximum amount of period t consumption that the consumer can buy, if he borrows as much money as she can possibly repay in period $t+1$ is $m_t + \frac{1}{1+r}m_{t+1}$. This is called *permanent income*; it is the whole present and future stream of income of the agent evaluated as of time t .

The intertemporal budget constraint for an infinitely lived agent will be:

$$\sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right)^t c_t = \sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right)^t m_t$$

and the right hand side is the permanent income.

6.1.3 Choice over time, for given r

We are ready now for the agent's saving problem (we have preferences and budget constraints, we do not need anything else):

$$\max_{c_1, c_2} u(c_t) + \beta u(c_{t+1})$$

subject to

$$c_t + \frac{1}{1+r}c_{t+1} = m_t + \frac{1}{1+r}m_{t+1}$$

³Note also that we implicitly assume that the agent does not save nor borrow at time $t+1$. This is without loss of generality because at time $t+1$: i) she will not want to save as she receives no utility from consuming at time $t+2$; and ii) she cannot borrow as she cannot repay at time $t+2$ (nor she can default).

The first order conditions for the maximization problem are:

$$\frac{u'(c_t)}{\beta u'(c_{t+1})} = 1 + r \quad (6.1)$$

$$c_t + \frac{1}{1+r}c_{t+1} = m_t + \frac{1}{1+r}m_{t+1} \quad (6.2)$$

Develop your intuition about what the condition means by drawing the corresponding indifference curve diagram.

We now solve the problem, that is, we provide some implications of the solution. We distinguish three different cases.

1. Suppose $\beta = \frac{1}{1+r}$. Then the first order condition (9.3) implies (convince yourself)

$$c_{t+1} = c_t,$$

and the agent *smooths consumption* over time. Using the budget constraint:

$$c_{t+1} = c_t = \frac{1}{2} \left(m_t + \frac{1}{1+r}m_{t+1} \right) \quad (6.3)$$

that is, consumption at time t (and $t+1$) depends on *permanent income*, not just today's income. The dependence of consumption on *permanent income* rather than *income*, holds generally (not just when $\beta = \frac{1}{1+r}$); it is in fact the fundamental insight of the modern (after the 60's) theory of consumption, due to Milton Friedman and Franco Modigliani.

2. Suppose $\beta > \frac{1}{1+r}$. That is, the interest rate is relatively high (with respect to discounting β). In this case, the first order condition (9.3) implies (convince yourself)

$$c_{t+1} > c_t.$$

For a relatively high interest rate the agent will save more or borrow less (with respect to the case in which $\beta = \frac{1}{1+r}$), choosing a consumption path which is increasing over time.

3. Suppose $\beta < \frac{1}{1+r}$. That is, the interest rate is relatively low (with respect to discounting β). In this case, the first order condition (9.3) implies (convince yourself)

$$c_{t+1} < c_t.$$

For a relatively low interest rate the agent will save less or borrow more (with respect to the case in which $\beta = \frac{1}{1+r}$), choosing a consumption path which is decreasing over time.

6.1.4 Determinants of savings

Recall that we say that an agent saves if $m_t > c_t$ and that she borrows if $m_t < c_{t+1}$. Suppose the interest rate in the economy $1 + r$ increases. What happens to the savings of a saver? And, to the borrowings of a borrower? A figure with the budget constraint and indifference curves might help: consider first the case in which $m_t = 0$, so that the agent is certainly (under our assumptions on preferences) a borrower; and consider then the case in which $m_{t+1} = 0$ and the agent is a saver.

6.1.5 The interest rate r

The interest rate is just any other price. It is determined at a competitive equilibrium by the market clearing (demand equal supply) condition. In this case what is clearing is the demand and the supply of loans, that is, borrowing must equal lending. Suppose there are many agents in our economy, indexed by $i \in I$. Let the income of an arbitrary agent i at t (resp. at $t + 1$) be denoted by m_t^i (resp. m_{t+1}^i), and her consumption, respectively, c_t^i and c_{t+1}^i . The market clearing conditions in this economy are then written formally,

$$\sum_{i \in I} c_t^i - m_t^i = 0, \quad \sum_{i \in I} c_{t+1}^i - m_{t+1}^i = 0$$

Note that the first condition captures borrowing equal lending. What is the role of the second condition then? In fact, none: if $\sum_{i \in I} c_t^i - m_t^i = 0$, then, it is necessarily the case that $\sum_{i \in I} c_{t+1}^i - m_{t+1}^i = 0$ (you can convince yourself by summing all budget constraints over agents $i \in I$).

A classical model of equilibrium interest rate is one where i) the economy at t is populated by a group of *young* agents (with little income) and a group of *middle-aged* agents (with relatively large positive income); and at $t + 1$ the young agents become middle-aged (with relatively large positive income) and the *middle-aged* become *young* (with little income) - ok, if you think Benjamin Button agents are too curious to be realistic, just assume that the

middle-aged became *old* (with little income) and nothing is changed.⁴

6.1.6 Inflation and Monetary Neutrality

It turns out that it is easy to modify our analysis to deal with the case of inflation. To do this, let us suppose that the price of the consumption good increases at time $t + 1$ (that is, the same amount of dollars buys less consumption, there's inflation). Keeping the normalization that $p_t = 1$, without loss of generality, we have $p_{t+1} > p_t = 1$. Given the prices of consumption in the two periods, the rate of inflation, π , is given by

$$\pi = \frac{p_{t+1} - p_t}{p_t} = p_{t+1} - 1$$

Suppose the income of the agent m_t in the first period and m_{t+1} in the second, are *real*, that is in units of goods (not dollars). Let i denote the interest rate (we called it r before, but bear with me for a couple of lines). Then the agent budget constraint is:

$$c_t + \frac{1}{1+i} p_{t+1} c_{t+1} = m_t + \frac{1}{1+i} p_{t+1} m_{t+1}.$$

Substituting our definition of π , the budget constraint can be re-written as:

$$c_t + \frac{1+\pi}{1+i} c_{t+1} = m_t + \frac{1+\pi}{1+i} m_{t+1}$$

Note that if you substitute $1+r = \frac{1+i}{1+\pi}$ in this budget set you get back the budget set of the economy with no inflation. We are ready distinguish between the *nominal interest rate*, i , and the *real interest rate*, r , which we define by

$$1+r = \frac{1+i}{1+\pi}.^5$$

The real interest rate is the interest rate in terms of goods; while the nominal interest rate is in terms of dollars. If dollars lose value over time (because

⁴These economies, called *overlapping generation*, were developed by Maurice Allais and Robert Samuelson in the 50's.

⁵If you really want to see the exact expression for the *real interest rate* in terms of the nominal interest rate and the inflation rate, here it is:

$$r = \frac{1+i}{1+\pi} - 1 = \frac{i-\pi}{1+\pi}.$$

of inflation), they are different. When there is no inflation, however, the real and the nominal interest rates are equal (check this). In the previous section there was no inflation, what we called r was both a real and a nominal interest rate.

Does inflation have any effect on the economy? Are equilibrium quantities changed because of inflation? Are prices? The answer to these questions is straightforward once you noticed that inflation only affects the units of the budget constraint. The equilibrium condition determines the real interest rate; the nominal interest rate is determined by adding inflation. If the preferences and the incomes of the economy are not changed with the advent of inflation, the equilibrium real interest rate will not change because of inflation per se, nor will equilibrium allocations. Make sure you really understand this! Inflation will have no effects in our economy unless something is changed. For instance, the situation is different if the agent's income is nominal rather than real, that is, if m_t and m_{t+1} are in dollar units. In this case, the budget constraint becomes (convince yourself):

$$c_t + \frac{1 + \pi}{1 + i} c_{t+1} = m_t + \frac{1}{1 + i} m_{t+1}$$

and inflation acts as a tax, diminishing the purchasing power of the agent's income at time $t + 1$, and hence diminishing his/her permanent income. We would expect that in the short run *unexpected* inflation has an effect on wages, and hence on agents' income, but not in the long run. See Bob Lucas' paper in my webpage on *monetary neutrality*. *Monetary neutrality* is nothing else than the equilibrium property we have just discussed, that is, inflation has no effects if incomes are real (in the long run). We call it *monetary neutrality* because we usually think of the inflation rate as essentially the same as the growth rate of the quantity of money. Can you see why is this the case?

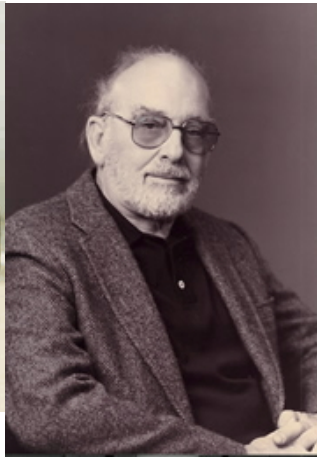
For low inflation rates the real interest rate is approximately

$$r \approx i - \pi.$$

6.2 Choice under Uncertainty; a.k.a. Portfolio choice



John Von Neumann and
Oskar Morgenstern



David Cass



Joeseeph Stiglitz

We first introduce random variables and expected values.

6.2.1 Probability

Let $\Omega = \{\omega_1, \omega_2\}$ be the set of all possible *events*; for instance, $\omega_1 = \textit{it rains}$, $\omega_2 = \textit{it does not rain}$. Consider the following *probability distribution* (on events): event ω_1 will occur with probability p , $0 < p < 1$, and event ω_2 with probability $1 - p$. Stop here and reflect on what does the wording "with probability p " means.

We need now to introduce some simple concepts from statistics, concepts which refer to random variables: a *random variable* y is a map from Ω to R which takes values

$$y_1 = y(\omega_1), \quad y_2 = y(\omega_2).$$

The *expected value* of y is:

$$py_1 + (1 - p)y_2,$$

denoted Ey .

The *variance* of y is:

$$p(y_1 - Ey)^2 + (1 - p)(y_2 - Ey)^2,$$

denoted $\sigma^2(y)$; $\sigma(y) = \sqrt{\sigma^2(y)}$ is the *standard deviation* of y . Both the variance and the standard deviation are measures of the variability of y . Note for instance that if y is constant, then $y = Ey$ and $\sigma^2(y) = \sigma(y) = 0$.

Consider two random variables y and z taking values

$$y_1 = y(\omega_1), \quad y_2 = y(\omega_2), \quad z_1 = z(\omega_1), \quad z_2 = z(\omega_2)$$

The *covariance* of y and z is:

$$p(y_1 - Ey) \cdot (z_1 - Ez) + (1 - p)(y_2 - Ey) \cdot (z_2 - Ez),$$

denoted $cov(y, z)$; $\frac{cov(y, z)}{\sigma(y) \cdot \sigma(z)}$ is the *correlation coefficient* of y and z .

The covariance and the correlation coefficient of y and z are a measure of how the variables y and z move together, that is, in our example the covariance and the correlation coefficient are maximal when either y_1 and z_1 are both greater than their respective means Ey and Ez ; or else y_1 and z_1 are both smaller than their respective means Ey and Ez .

6.2.2 Expected Utility

Consider an agent trying to evaluate the utility associated to consuming the random variable y (defined above). Let the utility function of the agent be $u(c)$. His utility when consuming y is:

$$u(y_1) \text{ with probability } p, \text{ and } u(y_2) \text{ with probability } 1 - p$$

We assume an agent evaluates the utility of a random variable y by *expected utility*; that is,

$$pu(y_1) + (1 - p)u(y_2)$$

In other words, we assume that, when facing uncertainty, agents maximize expected utility. A lot of experiments document failures of this assumption in various circumstances. A lot of theoretical work addresses the failure, postulating different (still optimizing) behavior on the part of agents. Most economic theory still uses the assumption in first approximation.

6.2.3 Risk Aversion

Does an agent with utility function $u(c)$ prefer to consume the random variable y or its expected value Ey ?

Consider an agent with preferences represented by a continuous utility function $u : \Re \rightarrow \Re$. Assume the agent is an expected utility maximizer. Then the agent will prefer the expected value Ey to the (non-degenerate) random variable y iff $u(c)$ is strictly concave.

Proof. Draw a strictly concave utility function, and represent in the figure $u(Ey)$ and $pu(y_1) + (1 - p)u(y_2)$. It will be apparent that

$$u(Ey) > pu(y_1) + (1 - p)u(y_2)$$

iff u is strictly concave. ■

If an agent has strictly concave utility function, we say he is *risk averse*.

6.2.4 State contingent consumption

How does the consumption problem of the agent look like with uncertainty? Suppose the agent's income is a random variable y , that is, it takes value $y_1 = y(\omega_1)$ with probability p and $y_2 = y(\omega_2)$ with probability $1 - p$. Consumption in this case is also a random variable, denoted c . The agent distributes his consumption across states (ω_1, ω_2) by trading state contingent consumption and income, that is, consumption and income to be delivered conditionally on the realized state. Assume, without loss of generality, that $y_1 > y_2$, so that state ω_2 is the unfavorable state for the agent. Convince yourself that allowing trading in state contingent consumption is equivalent to allowing consumers trading insurance contracts: to insure the agent buys consumption in the unfavorable state ω_2 and sells consumption in the favorable state ω_1 . Let the price of consumption delivered in state ω_1 (resp. ω_2) to be q_1 (resp. q_2). The agent's choice problem is then:

$$\begin{aligned} \max_{\{c_1, c_2\}} & pu(c_1) + (1 - p)u(c_2) \\ q_1 c_1 + q_2 c_2 &= q_1 y_1 + q_2 y_2 \end{aligned}$$

The first order conditions of this problem are:

$$\begin{aligned} \frac{1 - p}{p} \frac{u'(c_2)}{u'(c_1)} &= \frac{q_2}{q_1} \\ q_1 c_1 + q_2 c_2 &= q_1 y_1 + q_2 y_2 \end{aligned}$$

We now solve the problem, that is, we provide some implications of the solution. We distinguish three different cases.

1. Suppose $\frac{1-p}{p} = \frac{q_2}{q_1}$. We say that, in this case, the implied *insurance price is fair*. The first order condition (6.4) implies (convince yourself)

$$c_1 = c_2,$$

and the agent *smooths consumption* over states of uncertainty. In other words, the agent fully insures himself, consuming the same amount independently of the realized state of uncertainty. Using the budget constraint:

$$c_1 = c_2 = \frac{q_1 y_1 + q_2 y_2}{q_1 + q_2} \quad (6.4)$$

2. Suppose $\frac{1-p}{p} > \frac{q_2}{q_1}$. We say that, in this case, the implied *insurance price is more than fair* (the relative price of consumption in state insurance ω_2 is low and insurance is cheap). The first order condition (6.4) implies (convince yourself)

$$c_1 < c_2,$$

and the agent *over-insures* himself, consuming more in state ω_1 , when his income is higher.

3. Suppose $\frac{1-p}{p} < \frac{q_2}{q_1}$. We say that, in this case, the implied *insurance price is less than fair* (the relative price of consumption in state insurance ω_2 is high and insurance is expensive). The first order condition (6.4) implies (convince yourself)

$$c_1 < c_2,$$

and the agent *over-insures* himself, consuming more in state ω_1 , when his income is higher.

6.3 Problems

Problem 6.3.1 Consider a two period economy. Agents are all identical, that is, there is one representative agent. The representative agent is alive at time t and $t + 1$, and has preferences:

$$\ln x_t + \beta \ln x_{t+1}, \quad \beta < 1.$$

This agent is endowed with 10 units of the consumption good at time t and at time $t + 1$. There is no inflation in this economy, and hence you can assume throughout that the price for the good at time t is 1. i) Write down the consumer maximization problem of the representative agent (call the real interest rate r in the budget constraint), first order conditions, and demand functions. ii) Write down the market clearing conditions (that is, feasibility conditions) for the whole economy. iii) Solve for the equilibrium interest rate and for the representative agent equilibrium allocation. iv) Suppose the agent cannot borrow and lend, that is, savings are zero and there is no interest rate r : he/she has to consume his/her own endowment in each period. How would you write the budget constraints? [hint: the plural is not a typo; there is a budget constraint for each time period] Once again the price of consumption at t and also at $t + 1$ now can be normalized to 1. Are his/her equilibrium allocations changed? v) Suppose again that the representative agent cannot borrow or lend, but he/she can now invest units of the consumption good at time t (he/she still has his/her endowment as before). Call the amount invested k_{t+1} ; the production function is

$$y_{t+1} = \alpha k_{t+1}, \alpha > 1$$

*that is, the agent can give up k_{t+1} units of consumption at time t to get αk_{t+1} extra units at time $t + 1$ (this is all per-capita; in terms of the notation used in class, $k_t = 0$). Assume $\alpha\beta > 1$ (you will need this later). Write down the budget constraints for the representative agent [be careful here! the plural is still not a typo]. Write down the consumer maximization problem, using the budget constraints you derived, to solve for the optimal choice of investment k_{t+1} . Then solve for the equilibrium allocations. **[Bonus Question; this is hard]** Suppose that borrowing and lending markets are now open, that is, there is an interest rate to be determined. This of course together with the investment and production technology as above. How does the budget constraint [yes, singular] look like? Are equilibrium allocations changed? What is the equilibrium interest rate?*

Answer. i) For the maximisation problem, we need the objective function, which is the utility function above, and a budget constraint. To get the budget constraint we use the definition of savings and solve as we did in

class

$$\begin{aligned}
 s &= m_t - x_t = 10 - x_t \\
 x_{t+1} &= m_{t+1} + (1+r)s = 10 + (1+r)s \\
 \Rightarrow x_t + \frac{1}{1+r}x_{t+1} &= 10 + \frac{1}{1+r}10 = 10\left(\frac{1}{1+r} + 1\right)
 \end{aligned}$$

This gives the Lagrangian

$$L(x_1, x_2, \mu) = \ln x_t + \beta \ln x_{t+1} - \mu \left(x_t + \frac{1}{1+r}x_{t+1} - 10 - \frac{1}{1+r}10 \right)$$

And first order necessary conditions:

$$\begin{aligned}
 \frac{\partial L}{\partial x_t} &= \frac{1}{x_t} - \mu = 0 \\
 \frac{\partial L}{\partial x_{t+1}} &= \frac{\beta}{x_{t+1}} - \mu \frac{1}{1+r} = 0 \\
 \frac{\partial L}{\partial \mu} &= x_t + \frac{1}{1+r}x_{t+1} - 10 - \frac{1}{1+r}10 = 0
 \end{aligned}$$

Combining these gives us

$$\begin{aligned}
 \frac{x_{t+1}}{x_t \beta} &= (1+r) \\
 \Rightarrow x_{t+1} &= \beta(1+r)x_t
 \end{aligned}$$

Substituting into the constraint gives us

$$\begin{aligned}
 x_t + \beta x_t &= 10\left(1 + \frac{1}{1+r}\right) \\
 \Rightarrow x_t &= \frac{10}{1+\beta} \left(\frac{2+r}{1+r}\right) \\
 \Rightarrow x_{t+1} &= \frac{10\beta}{1+\beta} (2+r)
 \end{aligned}$$

ii) As always, the market clearing (feasibility) conditions tell us that the sum of the demand for a particular good across all agents has to equal the total supply (or endowment). Here this is pretty easy, as there is only one agent! So we have

$$\begin{aligned}
 x_t &= m_t = 10 \\
 x_{t+1} &= m_{t+1} = 10
 \end{aligned}$$

Substituting in from above gives

$$\begin{aligned}\frac{10}{1+\beta}\left(\frac{2+r}{1+r}\right) &= 10 \\ \frac{10\beta}{1+\beta}(2+r) &= 10\end{aligned}$$

iii) Due to Walras' Law, we know we only need to use one of these conditions to solve for the equilibrium interest rate. For ease, we'll use the second one. We can rearrange as follows:

$$\begin{aligned}\frac{10\beta}{1+\beta}(2+r) &= 10 \\ 2+r &= \frac{1+\beta}{\beta} \\ r &= \frac{1+\beta-2\beta}{\beta} \\ r &= \frac{1-\beta}{\beta}\end{aligned}$$

Substituting back into the demand functions gives

$$\begin{aligned}x_t = \frac{10}{1+\beta}\left(\frac{2+r}{1+r}\right) &= 10 \frac{1}{1+\beta} \frac{1+\beta}{\beta} \frac{1}{1+\frac{1-\beta}{\beta}} \\ &= 10 \frac{1}{1+\beta} \frac{1+\beta}{\beta} \frac{\beta}{\beta+1-\beta} \\ &= 10\end{aligned}$$

and

$$\begin{aligned}x_{t+1} &= \frac{10\beta}{1+\beta}(2+r) \\ &= \frac{10\beta}{1+\beta} \frac{1+\beta}{\beta} \\ &= 10\end{aligned}$$

Note that we didn't really have to substitute back in at this stage. We already know from feasibility that consumption has to equal endowment in each period.

iv) If the agent cannot borrow or lend, then they have to balance the budget in each period. In other words, savings have to be equal to zero. Using the expressions we wrote down for savings before, we get

$$\begin{aligned}s &= m_t - x_t = 10 - x_t = 0 \\ \Rightarrow x_t &= 10\end{aligned}$$

and

$$x_{t+1} = m_{t+1} + (1+r)s = 10 + (1+r)s = 10$$

As the consumer cannot borrow and save, there is no choice for her to make: she just has to eat her endowment each period. In fact, this is the same as the equilibrium allocation from the previous problem though note that they come about from very different processes. Here, we have constrained the agents to eat their own endowment each period. Previously it was the feasibility constraint that led to the agent eating their own endowment.

v) We know that in period 1, the agent receives his endowment of 10 units. At this point, she can choose to consume these units or save them as capital. This gives us our period 1 budget constraint:

$$x_t + k_{t+1} = m_t = 10$$

In the second period, the agent receives her endowment, plus the output from investment in period 1. She will spend all this income on consumption. This gives the second budget constraint:

$$x_{t+1} = y_{t+1} + m_{t+1} = \alpha k_{t+1} + 10$$

We solve for the equilibrium allocation. There are lots of different ways to do this. The way we will follow here is as follows. First note that the only thing that the agent really gets to choose is the amount she saves k_{t+1} . Once we know this, we know c_t from the first budget constraint, and c_{t+1} from the second budget constraint. We are going to therefore proceed by using the budget constraints above to get an expression for c_t and c_{t+1} and substitute this into the utility function

$$\begin{aligned}x_t &= 10 - k_{t+1} \\ x_{t+1} &= \alpha k_{t+1} + 10 \\ \Rightarrow U &= \ln(10 - k_{t+1}) + \beta \ln(\alpha k_{t+1} + 10)\end{aligned}$$

This now an unconstrained maximisation problem in which the agents have to choose k_{t+1} . Taking first order conditions gives:

$$\begin{aligned}
 \frac{\partial U}{\partial k_{t+1}} &= -\frac{1}{10 - k_{t+1}} + \alpha\beta \frac{1}{\alpha k_{t+1} + 10} = 0 \\
 \Rightarrow \frac{1}{10 - k_{t+1}} &= \alpha\beta \frac{1}{\alpha k_{t+1} + 10} \\
 \alpha k_{t+1} + 10 &= \alpha\beta(10 - k_{t+1}) \\
 \alpha k_{t+1} + \alpha\beta k_{t+1} &= \alpha\beta 10 - 10 \\
 \alpha(1 + \beta)k_{t+1} &= (\alpha\beta - 1)10 \\
 k_{t+1} &= \frac{(\alpha\beta - 1)10}{\alpha(1 + \beta)}
 \end{aligned}$$

Note that this is only positive if $\alpha\beta \geq 1$. Otherwise (as we don't allow the person to invest negative amounts), the optimal choice will be $k_{t+1} = 0$. Assuming that this condition is satisfied, we can calculate x_t and x_{t+1}

$$\begin{aligned}
 x_t &= 10 - k_{t+1} = 10 - \frac{(\alpha\beta - 1)10}{\alpha(1 + \beta)} = 10\left(\frac{\alpha(1 + \beta) - (\alpha\beta - 1)}{\alpha(1 + \beta)}\right) \\
 &= 10\frac{(\alpha + 1)}{\alpha(1 + \beta)} \\
 x_{t+1} &= \alpha k_{t+1} + 10 = \alpha \frac{(\alpha\beta - 1)10}{\alpha(1 + \beta)} + 10 = \frac{(\alpha\beta - 1)10}{(1 + \beta)} + 10 \\
 &= 10\frac{((\alpha\beta - 1) + (1 + \beta))}{(1 + \beta)} = 10\frac{\beta(1 + \alpha)}{1 + \beta}
 \end{aligned}$$

[Bonus Question] First, let's write down the budget constraint for each period. In the first period, the number of bonds that the agent will buy is equal to her income minus her spending on consumption minus her spending on capital:

$$s = m_t - x_t - k_{t+1}$$

In the second period, she will spend all the income she gets, which is equal to her endowment, plus the return she gets from bonds, plus the return she gets from her capital

$$x_{t+1} = m_{t+1} + (1 + r)s + \alpha k_{t+1}$$

We can now combine these two budget constraints in the usual way:

$$\begin{aligned}
 s &= \frac{1}{(1+r)}(x_{t+1} - m_{t+1} - \alpha k_{t+1}) \\
 \Rightarrow m_t - x_t - k_{t+1} &= \frac{1}{(1+r)}(x_{t+1} - m_{t+1} - \alpha k_{t+1}) \\
 \Rightarrow x_t + \frac{1}{(1+r)}x_{t+1} &= m_t + \frac{1}{(1+r)}m_{t+1} + \left(\frac{1}{1+r}\alpha - 1\right)k_{t+1}
 \end{aligned}$$

Now stare hard at the right hand side of the equation, remembering that the left hand side is the amount that the person spends and the right hand side is the amount that they have to spend. What will happen if $\frac{1}{1+r}\alpha - 1 > 0$? Then the agent just gets richer and richer the more capital they buy. In this case, they will just keep on borrowing money and spending it on capital, so the demand for capital will be infinite. This cannot be an equilibrium. Why is this happening? Note that the condition $\frac{1}{1+r}\alpha - 1 > 0 \Rightarrow \alpha > 1+r$ or in other words the return on the bond is lower than the marginal product of capital. This explains our previous result: There is an arbitrage opportunity in this economy. The agent can borrow money on bonds and invest it in capital to make free money.

The only possible equilibrium of this economy is therefore when $\frac{1}{1+r}\alpha - 1 \leq 0$, or the interest rate on the bond is greater than or equal to the marginal product of capital. If $\frac{1}{1+r}\alpha - 1 < 0 \Rightarrow \alpha < 1+r$ then the marginal product of capital is less than the interest rate. In this case, we might think that there is an arbitrage opportunity from the agent selling capital and buying bonds. However, this is not the case because the agent cannot buy less than 0 units of capital. However, we do know that the agent will never buy capital, as she would do better by investing in the bond. The economy therefore looks exactly the same as if there was no capital and investment, in other words the one we solved in part one of this question. In this case, we know that $r = \frac{1-\beta}{\beta}$. So if $1+r = \frac{1}{1+\beta} > \alpha$, this will be an equilibrium with allocations 10 in each period.

There will always be another equilibrium in which $\alpha = 1+r$. Here the rate of return on bonds is exactly the same as the rate of return on capital, so people will be indifferent between the two methods of investing. We can therefore safely assume that the bonds don't exist. In this case, we will be in exactly the same equilibrium we calculated above with no bonds.

Problem 6.3.2 Consider the following economy. Time is $t = 1, 2$. There is

a representative consumer whose production technology is:

$$Y_t = F(k) = Ak_t^\alpha$$

where k_t is the per capita capital stock at time t and A has a "productivity parameter" interpretation. The consumer OWNS the production technology and thus makes the production decisions. The capital stock from one period can be stored and re-used in the next period (with NO depreciation). The consumer is endowed with $k_1 > 0$ units of capital stock in $t = 1$. There is no uncertainty. Write the consumer's capital accumulation problem.

Answer. The "story" is the following: There is one consumer (who is assumed to represent the aggregate preferences of many consumers) who has access to a production technology (capital can be used to produce an output good) and a storage technology (capital is accumulated from period to period).

The output good in each period can be used for one of two purposes - consumption or increasing the capital stock for the next period (investment/savings).

The period by period "resource" constraints are that consumption today plus the change in capital stock must equal production today. The change in capital stock is the capital stock tomorrow minus the capital stock today. This means the resource constraints are

$$c_t + (k_{t+1} - k_t) = y_t$$

Remember that the production is given by the production function, $y_t = Ak_t^\alpha$. Putting this into the resource constraint gives

$$c_t + (k_{t+1} - k_t) = Ak_t^\alpha$$

The final thing to notice is that the consumer only lives for two periods. Thus he/she has no incentive to build up capital stock for the next period. As a consequence he/she will not save/invest in the final period. Thus the second period resource constraint will be

$$c_2 = Ak_2^\alpha$$

Given these constraints, the capital stock in the initial period that the consumer is endowed with, the consumer needs to find the combination of consumption in the two periods and capital stock in the second period that maximises his/her utility. This leads to the capital accumulation problem.

Problem 6.3.3 *The consumer needs to solve the following problem*

$$\begin{aligned} &\text{Given } \{k_1\} \\ &\text{Choose } \{c_1, c_2, k_2\} \text{ to} \\ &\text{Max} \{ \ln c_1 + \beta \ln c_2 \} \text{ s.t.} \\ &\quad c_1 + (k_2 - k_1) = Ak_1^\alpha \\ &\quad c_2 = Ak_2^\alpha \end{aligned}$$

This problem can be vastly simplified by substituting out c_1, c_2 and solving the following problem

$$\begin{aligned} &\text{Given } \{k_1\} \\ &\text{Choose } \{k_2\} \text{ to} \\ &\text{Max} \{ \ln[Ak_1^\alpha - (k_2 - k_1)] + \beta \ln Ak_2^\alpha \} \end{aligned}$$

The f.o.c. for this problem is

$$\frac{1}{[Ak_1^\alpha - (k_2 - k_1)]}(-1) + \frac{\beta}{Ak_2^\alpha}(\alpha Ak_2^{\alpha-1}) = 0$$

To solve for the growth rate of consumption. The thing to notice here is that, by the nature of using the chain rule, we can easily substitute back the expressions for c_1 and c_2 . Remember from our period by period budget constraints

$$\begin{aligned} c_1 &= Ak_1^\alpha - (k_2 - k_1) \\ c_2 &= Ak_2^\alpha \end{aligned}$$

Putting these into the f.o.c. and rearranging

$$\begin{aligned} \frac{1}{[c_1]}(-1) + \frac{\beta}{c_2}(\alpha Ak_2^{\alpha-1}) &= 0 \\ \frac{\beta}{c_2}(\alpha Ak_2^{\alpha-1}) &= \frac{1}{c_1} \\ \beta \alpha Ak_2^{\alpha-1} &= \frac{c_2}{c_1} \end{aligned}$$

as was done in the lecture notes. Do notice that this is not technically a full solution to the problem yet since k_2 is the choice variable of the problem. To

solve for k_2 , rearrange the f.o.c. and massage them:

$$\begin{aligned}
 \frac{1}{[Ak_1^\alpha - (k_2 - k_1)]} &= \frac{\beta}{Ak_2^\alpha}(\alpha Ak_2^{\alpha-1}) \\
 \implies 1 &= [Ak_1^\alpha - (k_2 - k_1)] \frac{\beta(\alpha Ak_2^{\alpha-1})}{Ak_2^\alpha} \\
 \implies 1 &= [Ak_1^\alpha - (k_2 - k_1)] \alpha \beta k_2^{-1} \\
 \implies k_2 &= \alpha \beta Ak_1^\alpha - \alpha \beta k_2 + \alpha \beta k_1 \\
 \implies (1 + \alpha \beta)k_2 &= \alpha \beta (Ak_1^\alpha + k_1) \\
 \implies k_2 &= \frac{\alpha \beta}{(1 + \alpha \beta)} (Ak_1^\alpha + k_1)
 \end{aligned}$$

The economy is exactly the same as in Question 1, except that now we introduce uncertainty over the value of the productivity parameter, A , which is now a random variable. In any time period, t , it takes the value A_1 with probability p and the value A_2 with probability $(1 - p)$. Formulate the capital accumulation problem.

Answer. As a consequence of uncertainty, the consumer does not know what his resource constraints will be in future periods. This depends on which "state" the world is in, i.e. whether productivity is A_1 or A_2 . Thus in future periods there will be one resource constraint for each state. Specifically this will mean in period 2 we have

$$\begin{aligned}
 c_2(A_1) &= A_1 k_2^\alpha \text{ if } A = A_1 \\
 &\text{and} \\
 c_2(A_2) &= A_2 k_2^\alpha \text{ if } A = A_2
 \end{aligned}$$

Thus the consumer is also uncertain as to his/her future consumption.

To describe the knowledge that the consumer has more fully, at any time period, t , the consumer *knows*

k_t

A at time t (this can be either A_1 or A_2 , but we will denote it A)

c_t (once he/she has chosen it)

k_{t+1} (once he/she has chosen it)

The consumer does *not know*

A at time $t+1$

c_{t+1}

The consumer is assumed to be an expected utility maximiser. Thus he chooses $\{c_1, c_2, k_{t+1}\}$ in order to maximise the following objective function

$$\mathbb{E}(u(c_1, c_2)) = \mathbb{E}(\ln c_1 + \beta \ln c_2)$$

where \mathbb{E} is the expectation operator and the expectation is taken over the random variable A . Since the \mathbb{E} is a *linear operator* (essentially this means that $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$) and c_1 is known when the optimisation is done (in period one)

$$\mathbb{E}(u(c_1, c_2)) = \ln c_1 + \beta \mathbb{E}(\ln c_2)$$

Since the random variable A is discrete, the expectation of a function of the random variable is given by

$$\mathbb{E}(f(A)) = pf(A_1) + (1-p)f(A_2)$$

where $f(\cdot)$ is a function that takes the random variable, A , as an argument.

This means that the second term of the objective function can be expressed as

$$\beta \mathbb{E}(\ln c_2) = \beta p \ln[c_2(A_1)] + \beta(1-p) \ln[c_2(A_2)]$$

The consumer needs then to solve the following problem

$$\begin{aligned} & \text{Given } \{k_1, A\} \\ & \text{Choose } \{c_1, c_2(A_1), c_2(A_2), k_2\} \text{ to} \\ & \text{Max } \{\ln c_1 + \beta p \ln[c_2(A_1)] + \beta(1-p) \ln[c_2(A_2)]\} \text{ s.t.} \\ & \quad c_1 + (k_2 - k_1) = Ak_1^\alpha \\ & \quad c_2(A_1) = A_1 k_2^\alpha \\ & \quad c_2(A_2) = A_2 k_2^\alpha \end{aligned}$$

This problem can be vastly simplified by substituting out $c_1, c_2(A_1), c_2(A_2)$ and solving the following problem

$$\begin{aligned} & \text{Given } \{k_1, A\} \\ & \text{Choose } \{k_2\} \text{ to} \\ & \text{Max } \{\ln[Ak_1^\alpha - (k_2 - k_1)] + \beta p \ln[A_1 k_2^\alpha] + \beta(1-p) \ln[A_2 k_2^\alpha]\} \end{aligned}$$

The *f.o.c.* for this problem is

$$\frac{1}{[Ak_1^\alpha - (k_2 - k_1)]}(-1) + \frac{\beta p}{A_1 k_2^\alpha}(\alpha A_1 k_2^{\alpha-1}) + \frac{\beta(1-p)}{A_2 k_2^\alpha}(\alpha A_2 k_2^{\alpha-1}) = 0$$

Solve for the growth rate of consumption. Notice that, by the nature of using the chain rule, we can easily substitute back the expressions for c_1 and $c_2(A_1)$, $c_2(A_2)$. Remember from our period by period budget constraints

$$\begin{aligned} c_1 &= Ak_1^\alpha - (k_2 - k_1) \\ c_2(A_1) &= A_1 k_2^\alpha \\ c_2(A_2) &= A_2 k_2^\alpha \end{aligned}$$

Putting these into the *f.o.c.* and rearranging

$$\begin{aligned} \frac{1}{[c_1]}(-1) + \frac{\beta p}{c_2(A_1)}(\alpha A_1 k_2^{\alpha-1}) + \frac{\beta(1-p)}{c_2(A_2)}(\alpha A_2 k_2^{\alpha-1}) &= 0 \\ \frac{\beta p}{c_2(A_1)}(\alpha A_1 k_2^{\alpha-1}) + \frac{\beta(1-p)}{c_2(A_2)}(\alpha A_2 k_2^{\alpha-1}) &= \frac{1}{c_1} \\ \alpha \beta k_2^{\alpha-1} \left[p \frac{A_1}{c_2(A_1)} + (1-p) \frac{A_2}{c_2(A_2)} \right] &= \frac{1}{c_1} \\ \alpha \beta k_2^{\alpha-1} \mathbb{E} \left[\frac{A}{c_2(A)} \right] &= \frac{1}{c_1} \\ \mathbb{E} \left[\alpha \beta A k_2^{\alpha-1} \frac{c_1}{c_2(A)} \right] &= 1 \\ \mathbb{E} \left[(MPK) \left(\frac{u'(c_2)}{u'(c_1)} \right) \right] &= 1 \end{aligned}$$

Problem 6.3.4 Consider the following economy. There are 3 periods: $t = 1, 2, 3$. Population is constant population equal to L every period. Consider a representative agent. Preferences of the representative agent are

$$u(c_1) + \beta u(c_2) + \beta^2 u(c_3), \quad \text{where } u(c) = \log(c)$$

The technology is as follows. Production is Cobb-Douglas, i.e., $Y_t = AK_t^\alpha L_t^{1-\alpha}$. The capital stock from can be stored and re-used in the next period (with depreciation rate δ). The initial capital's endowment is $K_1 > 0$. There is no uncertainty. Formulate the problem for the representative consumer and solve for consumption and investment per capita every period with full depreciation, i.e. $\delta = 1$.

Answer. To formulate the problem for the representative agent, we first define the per capita variables.

Let y_t and k_t denote output per capita and capital per capita, respectively, i.e. $y_t = \frac{Y_t}{L_t}$ and $k_t = \frac{K_t}{L_t}$. Notice that, by dividing the production function by total labor L_t , we express y_t in terms of k_t ,

$$\begin{aligned}\frac{Y_t}{L_t} &= A \left(\frac{K_t}{L_t} \right)^\alpha \left(\frac{L_t}{L_t} \right)^{1-\alpha} \\ y_t &= A k_t^\alpha\end{aligned}$$

In our setup there is no population growth, i.e. $L_t = L$ for $t = 1, 2, 3$. In case there was, it is necessary to make the appropriate adjustment in the investment process by the population growth rate. For example, consider a population growth percent rate equal to μ every period, i.e. $1 + \mu = \frac{L_{t+1}}{L_t}$. Then, investment per capita in period 2 is given by

$$\begin{aligned}i_2 &= \frac{K_3 - (1 - \delta)K_2}{L_2} = \frac{K_3}{L_2} - (1 - \delta)\frac{K_2}{L_2} \\ &= \frac{K_3}{L_3} \frac{L_3}{L_2} - (1 - \delta)k_2 = k_3(1 + \mu) - (1 - \delta)k_2\end{aligned}$$

Before formulating the problem notice that, since the consumer lives for only three periods, it makes no sense to invest in the third period, that is, increase future capital stock. In the third period, the consumer will optimally choose not to invest and consume all the output instead.

Hence, the capital accumulation problem for the consumer consists of choosing consumption and investment allocations $(c_1, c_2, c_3, i_1, i_2)$. to

$$\begin{aligned}\max \quad & \log(c_1) + \beta \log(c_2) + \beta^2 \log(c_3) \\ \text{subject to } & c_1 + i_1 = A k_1^\alpha \quad \text{with } i_1 = k_2 - (1 - \delta)k_1\end{aligned} \quad (6.5)$$

$$c_2 + i_2 = A k_2^\alpha \quad \text{with } i_2 = k_3 - (1 - \delta)k_2 \quad (6.6)$$

$$c_3 = A k_3^\alpha \quad (6.7)$$

$$\text{given } k_1 = \frac{K_1}{L} > 0$$

With $\delta = 1$ investment collapses to $i_1 = k_2$ and $i_2 = k_3$.

In our model the representative consumer decides to invest in capital stock for two main reasons. First, it is clear from the preferences of the

consumer that he wants to smooth consumption over time, and the only means to do this is by carrying over capital to the next period and using it in the production. Second, the higher the capital stock, the higher the output and, consequently, the higher the consumption.

Using the three resource constraints we obtain expressions for consumption per capita in terms of capital per capita and plug them into the objective function:

$$\max \log(Ak_1^\alpha - k_2) + \beta \log(Ak_2^\alpha - k_3) + \beta^2 \log(Ak_3^\alpha)$$

Taking FOCs with respect to the two choice variables, k_2 and k_3 , we have⁶

$$\begin{aligned} -\frac{1}{c_1} + \beta \frac{1}{c_2} \alpha A k_2^{\alpha-1} &= 0 \\ -\beta \frac{1}{c_2} + \beta^2 \frac{1}{c_3} \alpha A k_3^{\alpha-1} &= 0 \end{aligned}$$

which imply respectively

$$\frac{c_2}{c_1} = \beta \alpha A k_2^{\alpha-1} \quad (6.8)$$

$$\frac{c_3}{c_2} = \beta \alpha A k_3^{\alpha-1} \quad (6.9)$$

Plugging (6.7) into (6.9) we obtain the following expression for k_3

$$k_3 = \beta \alpha c_2 \quad (6.10)$$

Substituting k_3 by expression (6.10) in resource constraint (6.6) leads to the following condition for c_2

$$c_2 = \frac{1}{1 + \beta \alpha} A k_2^\alpha \quad (6.11)$$

Then, combining (6.11) and (6.8) we obtain

$$c_1 = \frac{1}{1 + \alpha \beta} \frac{1}{\alpha \beta} k_2 \quad (6.12)$$

⁶Recall the Chain Rule:

$$\frac{\partial f(g(x))}{\partial x} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$$

We can now pin down optimal k_2^* as a function of k_1 by plugging (6.12) in resource constraint (6.5)

$$k_2^* = \left(\frac{1}{1 + \alpha\beta} \frac{1}{\alpha\beta} \right)^{-1} A k_1^\alpha$$

Given this value for k_2^* , optimal c_1^* , c_2^* , k_3^* and c_3^* are determined by applying equations (6.12), (6.11), (6.10) and (6.7), respectively.

Problem 6.3.5 *Consider a savings problem with budget constraints*

$$\begin{aligned} c_1 + s &= m_1 \\ c_2 &= m_2 + (1 + r)s \end{aligned}$$

and preferences:

$$\ln c_1 + \beta \ln c_2$$

Construct the intertemporal budget constraint (IBC) and solve for savings s . Find how s depends on r . Embed now the economy in a overlapping generations economy with 2 consumers, $i = A, B$ with preferences:

$$\begin{aligned} U^i(c_1^i, c_2^i) &= \ln c_1^i + \beta \ln c_2^i \\ \beta &\in (0, 1); \end{aligned}$$

and endowments:

$$\begin{aligned} (m_1^A, m_2^A) &= (0, m) \\ (m_1^B, m_2^B) &= (m, 0). \end{aligned}$$

Answer. The intertemporal *budget constraint (IBC)* is constructed as follows from the budget constraints:

$$\begin{aligned} s &= \frac{c_2}{(1 + r)} - \frac{m_2}{(1 + r)} \\ \implies c_1 + \frac{c_2}{(1 + r)} - \frac{m_2}{(1 + r)} &= m_1 \\ \implies c_1 + \frac{c_2}{(1 + r)} &= m_1 + \frac{m_2}{(1 + r)} \end{aligned}$$

The intertemporal choice problem is:

$$\begin{aligned} & \text{To choose } \{c_1, c_2\} \text{ to} \\ & \text{Max}\{\ln c_1 + \beta \ln c_2\} \text{ s.t.} \\ & c_1 \geq 0 \\ & c_2 \geq 0 \\ & c_1 + \frac{c_2}{(1+r)} = m_1 + \frac{m_2}{(1+r)} \end{aligned}$$

Use Langrangian technique

$$\mathcal{L} = \ln c_1 + \beta \ln c_2 + \lambda \left(c_1 + \frac{c_2}{(1+r)} - m_1 - \frac{m_2}{(1+r)} \right)$$

The *f.o.c.* are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial c_1} &= \frac{1}{c_1} + \lambda = 0 \\ \implies \frac{1}{c_1} &= -\lambda \\ \frac{\partial \mathcal{L}}{\partial c_2} &= \frac{\beta}{c_2} + \frac{\lambda}{(1+r)} = 0 \\ \implies \frac{(1+r)\beta}{c_2} &= -\lambda \end{aligned}$$

Eliminating λ

$$\begin{aligned} \frac{1}{c_1} &= \frac{(1+r)\beta}{c_2} \\ \implies c_2 &= (1+r)\beta c_1 \end{aligned}$$

Putting this into the IBC

$$\begin{aligned} c_1 + \frac{(1+r)\beta c_1}{(1+r)} &= m_1 + \frac{m_2}{(1+r)} =: W \\ \implies c_1 &= \frac{W}{(1+\beta)} \\ \implies c_2 &= \frac{(1+r)\beta}{(1+\beta)} W \end{aligned}$$

Solving for the first period savings

$$\begin{aligned}
 s &= m_1 - c_1 \\
 &= m_1 - \left(\frac{m_1 + \frac{m_2}{(1+r)}}{(1+\beta)} \right) \\
 \Rightarrow s &= \frac{\beta}{(1+\beta)} m_1 - \frac{m_2}{(1+r)(1+\beta)}
 \end{aligned}$$

Then

$$\begin{aligned}
 \frac{\partial s}{\partial r} &= -\frac{m_2}{(1+\beta)} \frac{d}{dr} ((1+r)^{-1}) \\
 &= \frac{m_2}{(1+\beta)(1+r)^2} \\
 &\geq 0
 \end{aligned}$$

In the overlapping generations economy, an *allocation* is a list of consumption bundles, one for each consumer

$$\begin{aligned}
 &\{c^A, c^B\} \text{ where} \\
 c^A &= (c_1^A, c_2^A) \in X \\
 c^B &= (c_1^B, c_2^B) \in X
 \end{aligned}$$

A *feasible allocation* is an allocation that satisfies

$$\begin{aligned}
 c_1^A + c_1^B &= m_1^A + m_1^B = m \\
 c_2^A + c_2^B &= m_2^A + m_2^B = m
 \end{aligned}$$

i.e. satisfies the *resource constraints* of the economy.

A price-allocation pair, $(\hat{\underline{r}}, \{\hat{c}^A, \hat{c}^B\})$ where $\hat{\underline{r}} = (1, \hat{r})$ and $\hat{c}^i = (\hat{c}_1^i, \hat{c}_2^i)$, is a *competitive equilibrium (CE)* if:

For $i = A, B$ and given prices $\hat{\underline{r}} = (1, \hat{r})$, the bundle $\hat{c}^i = (\hat{c}_1^i, \hat{c}_2^i)$ solves the problem:

$$\begin{aligned}
 &\text{Choose } \{c_1^i, c_2^i\} \text{ to } \max \{\ln c_1^i + \beta \ln c_2^i\} \text{ s.t.} \\
 c_1^i &\geq 0 \\
 c_2^i &\geq 0 \\
 c_1^i + \frac{c_2^i}{(1+\hat{r})} &= m_1^i + \frac{m_2^i}{(1+\hat{r})}
 \end{aligned}$$

Markets clear:

$$\begin{aligned}\widehat{c}_1^A + \widehat{c}_1^B &= m_1^A + m_1^B = m \\ \widehat{c}_2^A + \widehat{c}_2^B &= m_2^A + m_2^B = m\end{aligned}$$

Therefore, for $\{(c_1^A, c_2^A), (c_1^B, c_2^B)\}$ the consumer optimisation condition implies:

$$\begin{aligned}c_1^A &= \frac{W^A}{(1+\beta)} = \frac{m}{(1+\beta)(1+r)} \\ c_2^A &= \frac{(1+r)\beta}{(1+\beta)} W^A = \frac{\beta}{(1+\beta)} m \\ c_1^B &= \frac{W^B}{(1+\beta)} = \frac{m}{(1+\beta)} \\ c_2^B &= \frac{(1+r)\beta}{(1+\beta)} W^B = \frac{(1+r)\beta}{(1+\beta)} m\end{aligned}$$

Using the market clearing for consumption at time 1:

$$\begin{aligned}\frac{m}{(1+\beta)(1+r)} + \frac{m}{(1+\beta)} &= m \\ \implies 1 + (1+r) &= (1+\beta)(1+r) \\ \implies 2+r &= (1+\beta) + r + r\beta \\ \implies 1-\beta &= r\beta \\ \implies 1+r &= 1 + \frac{1-\beta}{\beta} \\ \implies 1+r &= \frac{1}{\beta}\end{aligned}$$

Checking the solution with the market clearing condition for consumption at time 2

$$\begin{aligned}\frac{\beta}{(1+\beta)} m + \frac{\frac{1}{\beta}\beta}{(1+\beta)} m &= (1+\beta) \frac{m}{(1+\beta)} \\ &= m \\ &\text{as required.}\end{aligned}$$

The price-allocation pair that forms a Competitive Equilibrium is

$$\begin{aligned}\hat{r} &= \frac{1}{\beta} \\ \hat{c}_1^A &= \frac{\beta m}{(1 + \beta)} \\ \hat{c}_2^A &= \frac{\beta m}{(1 + \beta)} \\ \hat{c}_1^B &= \frac{m}{(1 + \beta)} \\ \hat{c}_2^B &= \frac{m}{(1 + \beta)}\end{aligned}$$

Notice two important elements of this equilibrium:

Agent B receives more consumption in both periods and thus has a higher utility than agent A .

Both agents perfectly smooth their consumption over the two time periods.

Problem 6.3.6 For the following intertemporal utility functions, solve the consumer's problem (in other words, solve for the demand of each good in each period as a function of prices. To do this, you will have to construct a budget constraint for the agent. Assume that the agent has an income of w_i in each period and that there is an interest rate $1 + r_i$ between period i and $i + 1$). In each case, calculate savings as a function of interest rate(s), and decide how savings will change as the interest rate changes. i) $U = u(c_1) + \beta u(c_2)$ where $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. ii) $U = u(c_1) + \beta u(c_2) + \beta^2 u(c_3)$ where $u(c) = \ln c$. (i.e., there are three periods). iii) $U = x_1 x_2 + \beta x_1 x_2$. Here there are 2 goods available in each period. Assume that, in each time period, the price of x_1 is p_1 and the price of good x_2 is p_2 . iv) $U = c_1 + \beta c_2$ (here you should consider three cases, where β is greater than, the same as, or less than $1 + r$). v) For Question i above, assume that there is inflation is 10% between the two periods (i.e. there is a price p_1, p_2 for the good in the two periods, and that $p_2 = 1.1p_1$). First assume that the agent's income is denominated in real terms. What is the equilibrium nominal interest rate. Now assume that the agent's income is denominated in nominal terms. What is the agent's income? vi) For the following economies, solve for the competitive equilibrium interest rate(s) and allocations 1) $U_A = u(c_A^1) + \beta u(c_A^2)$, $U_B = u(c_B^1) + \beta u(c_B^2)$, where the instantaneous utility functions are logarithmic, and with endowments (w_A^1, w_A^2) , (w_B^1, w_B^2) .

2) As above, but let $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $(w_A^1, w_A^2) = (1, 3)$, $(w_B^1, w_B^2) = (2, 1)$. 3) $U_A = u(c_A^1) + \beta u(c_A^2) + \beta^2 u(c_A^3)$, $U_B = u(c_B^1) + \beta u(c_B^2) + \beta^2 u(c_B^3)$, where the instantaneous utility functions are logarithmic, and with endowments (w_A^1, w_A^2, w_A^3) , (w_B^1, w_B^2, w_B^3) . 4) $U_A = \log c_A^1 + \beta \log c_A^2$, $U_B = c_B^1 + \beta c_B^2$ and endowments $(w_A^1, w_A^2) = (1, 3)$, $(w_B^1, w_B^2) = (2, 1)$.

Answer. i) $U = u(c_1) + \beta u(c_2)$, where $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$

The consumer problem is given by

$$\begin{aligned} \max \quad & u(c_1) + \beta u(c_2) \\ \text{subject to} \quad & c_1 + \frac{c_2}{1+r_1} = w_1 + \frac{w_2}{1+r_1} \end{aligned}$$

Taking FOC with respect to c_1 and c_2 we obtain respectively

$$\begin{aligned} u'(c_1) - \lambda &= 0 \\ \beta u'(c_2) - \frac{\lambda}{1+r_1} &= 0 \end{aligned}$$

where λ is the Lagrange multiplier associated with the budget constraint. Using these two equations we obtain

$$\beta \frac{u'(c_2)}{u'(c_1)} = \frac{1}{1+r_1}$$

Given the utility specification we assumed, this condition collapses to

$$\beta \left(\frac{c_2}{c_1} \right)^{-\sigma} = \frac{1}{1+r_1}$$

implying that

$$\frac{c_2}{c_1} = (\beta(1+r_1))^{\frac{1}{\sigma}}$$

Let $k = (\beta(1+r_1))^{\frac{1}{\sigma}}$. Then, $c_2 = kc_1$. Substituting c_2 by kc_1 in the budget constraint leads to

$$c_1 \left(1 + \frac{k}{1+r_1} \right) = w_1 + \frac{w_2}{1+r_1}$$

Then we can express c_1 as a function of the endowments, β , r_1 and σ ,

$$c_1 = \frac{w_1(1+r_1) + w_2}{1+r_1 + (\beta(1+r_1))^{\frac{1}{\sigma}}} \quad (6.13)$$

We can now pin down the saving function as

$$s_1 = w_1 - c_1 = w_1 - \frac{w_1(1 + r_1) + w_2}{1 + r_1 + (\beta(1 + r_1))^{\frac{1}{\sigma}}} \quad (6.14)$$

ii) $U = u(c_1) + \beta u(c_2) + \beta^2 u(c_3)$, where $u(c) = \log(c)$

The consumer problem is given by

$$\begin{aligned} & \max \quad u(c_1) + \beta u(c_2) + \beta^2 u(c_3) \\ \text{subject to} \quad & c_1 + \frac{c_2}{1 + r_1} + \frac{c_3}{(1 + r_1)(1 + r_2)} = w_1 + \frac{w_2}{1 + r_1} + \frac{w_3}{(1 + r_1)(1 + r_2)} \end{aligned}$$

Taking FOC with respect to c_1 , c_2 and c_3 we obtain respectively

$$\begin{aligned} u'(c_1) - \lambda &= 0 \\ \beta u'(c_2) - \frac{\lambda}{1 + r_1} &= 0 \\ \beta^2 u'(c_3) - \frac{\lambda}{(1 + r_1)(1 + r_2)} &= 0 \end{aligned}$$

where λ is the Lagrange multiplier associated with the budget constraint. Using the first two equations we obtain

$$\beta \frac{u'(c_2)}{u'(c_1)} = \frac{1}{1 + r_1} \quad (6.15)$$

while using the first and third equation leads to

$$\beta^2 \frac{u'(c_3)}{u'(c_1)} = \frac{1}{(1 + r_1)(1 + r_2)} \quad (6.16)$$

Given that $u(c) = \log(c)$, condition 6.8 collapses to

$$\beta \left(\frac{c_2}{c_1} \right) = \frac{1}{1 + r_1}$$

implying that

$$\frac{c_2}{c_1} = \beta(1 + r_1)$$

or, equivalently,

$$c_2 = \beta(1 + r_1)c_1 \quad (6.17)$$

Condition 6.9 turns to be

$$\beta \left(\frac{c_3}{c_1} \right) = \frac{1}{(1+r_1)(1+r_2)}$$

so,

$$c_3 = \beta(1+r_1)(1+r_2)c_1 \quad (6.18)$$

We can now plug equations 6.17 and 6.18 into the budget constraint and derive an expression for c_1 as a function of the endowments, interest rates and β .

$$c_1^* = \frac{1}{(1+\beta+\beta^2)} \left(w_1 + \frac{w_2}{1+r_1} + \frac{w_3}{(1+r_1)(1+r_2)} \right)$$

We now pin down c_2 and c_3 by substituting this expression into 6.17 and 6.18

$$\begin{aligned} c_2^* &= \frac{\beta(1+r_1)}{(1+\beta+\beta^2)} \left(w_1 + \frac{w_2}{1+r_1} + \frac{w_3}{(1+r_1)(1+r_2)} \right) \\ c_1^* &= \frac{\beta(1+r_1)(1+r_2)}{(1+\beta+\beta^2)} \left(w_1 + \frac{w_2}{1+r_1} + \frac{w_3}{(1+r_1)(1+r_2)} \right) \end{aligned}$$

It follows that the saving functions in period 1 and 2 are

$$\begin{aligned} s_1^* &= w_1 - c_1^* = w_1 - \frac{1}{(1+\beta+\beta^2)} \left(w_1 + \frac{w_2}{1+r_1} + \frac{w_3}{(1+r_1)(1+r_2)} \right) \\ s_2^* &= w_2 + s_1^*(1+r_1) - c_2^* \end{aligned}$$

iii) $U = x_1^1 x_2^1 + \beta x_1^2 x_2^2$, where x_i^t is the consumption of good i in period t .

We can formulate this consumer problem as follows

$$\begin{aligned} \max \quad & x_1^1 x_2^1 + \beta x_1^2 x_2^2 \\ \text{subject to} \quad & x_1^1 + p x_2^1 + \frac{x_2^2 + p' x_2^2}{1+r_1} = w_1^1 + p w_2^1 + \frac{w_1^2 + p' w_2^2}{1+r_1} \end{aligned}$$

where $p = \frac{p_2^1}{p_1^1}$ is the relative price in the first period and $p' = \frac{p_2^2}{p_1^2}$ is the relative price in the second period.

This utility function is convex and we will be having corner solutions. In particular either all the consumption today is zero or all future consumption is zero⁷.

⁷We face the same problem as in the Planner problem solved in the lab.

Let $W = w_1^1 + pw_2^1 + \frac{w_1^2 + p'w_2^2}{1+r_1}$. Now, if we assume that consumption today is positive, we can take FOC for x_1^1 and x_2^1 ,

$$x_1^1 : x_2^1 - \lambda = 0$$

$$x_2^1 : x_1^1 - p\lambda = 0$$

where λ is the Lagrange multiplier for the budget constraint. It follows that

$$px_2^1 = x_1^1$$

Hence, if consumption today is positive, the individual will be spending half of W in good 1 and the other half in good 2. In that case the utility is

$$U = \frac{W}{2} \frac{W}{2p} + \beta \cdot 0 \cdot 0 = \frac{W^2}{4p}$$

Now let us assume that consumption in period 1 is zero and consumption in period 2 is positive. Taking FOC with respect to x_1^2 and x_2^2 we have,

$$x_1^2 : \beta x_2^2 - \frac{\lambda}{1+r_1} = 0$$

$$x_2^2 : \beta x_1^2 - \frac{p'\lambda}{1+r_1} = 0$$

implying that

$$p'x_2^2 = x_1^2$$

Now the individual is spending half of $W(1+r_1)$ in good 1 and the other half in good 2,⁸ getting a utility value of

$$U = 0 \cdot 0 + \beta \frac{W(1+r_1)}{2} \frac{W(1+r_1)}{2p'} = \frac{\beta}{p'} \left(\frac{W(1+r_1)}{2} \right)^2$$

We have to decide whether the individual has zero consumption in period 1 or 2. As long as $\frac{W^2}{4p} \geq \frac{\beta}{p'} \left(\frac{W(1+r_1)}{2} \right)^2$, or equivalently, $\frac{p'}{p} \geq \beta(1+r_1)^2$, the individual will only be consuming in period 1. In case $\frac{p'}{p} \geq \beta(1+r_1)^2$ he will only consume in period 2. It is left to you to derive the saving function.

⁸Notice that now he consumes $W(1+r_1)$ rather than just W because he is consuming in period 2.

1) $U_A = u(c_A^1) + \beta u(c_A^2)$, $U_B = u(c_B^1) + \beta u(c_B^2)$, where $u(c) = \log(c)$, and (w_A^1, w_A^2) , (w_B^1, w_B^2) .

In the previous question we showed that when $u(c) = \log(c)$ condition 6.17 holds. This means that in this case we have

$$\begin{aligned} c_A^2 &= \beta(1 + r_1)c_A^1 \\ c_B^2 &= \beta(1 + r_1)c_B^1 \end{aligned}$$

Summing up these two equations we obtain

$$C^2 = \beta(1 + r_1)C^1$$

where C^i is the total consumption in period i . Using the resource constraints, we know that $C^1 = w_A^1 + w_B^1$ and $C^2 = w_A^2 + w_B^2$. Hence, the equilibrium interest rate is given by

$$(1 + r_1) = \frac{1}{\beta} \frac{w_A^2 + w_B^2}{w_A^1 + w_B^1}$$

The allocations can be pinned down by substituting this value for $1 + r_1$ in the saving functions:

$$\begin{aligned} s_A^* &= w_A^1 - \frac{1}{1 + \beta} \left(w_A^1 + \frac{w_A^2}{1 + r_1} \right) \\ s_B^* &= w_B^1 - \frac{1}{1 + \beta} \left(w_B^1 + \frac{w_B^2}{1 + r_1} \right) \end{aligned}$$

2) $U_A = u(c_A^1) + \beta u(c_A^2)$, $U_B = u(c_B^1) + \beta u(c_B^2)$, where $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, and $(w_A^1, w_A^2) = (1, 3)$, $(w_B^1, w_B^2) = (2, 1)$.

Now we are considering an economy populated by two individuals, A and B , where they can borrow and lend from each other, and the interest rate r_1 will adjust endogenously to clear the market.

Since we know that the two individuals have same preferences as before, it follows that

$$\begin{aligned} c_A^2 &= k c_A^1 \\ c_B^2 &= k c_B^1 \end{aligned}$$

where $k = (\beta(1 + r_1))^{\frac{1}{\sigma}}$. If we aggregate consumption we obtain that $C_2 = k C_1$, where C_1 and C_2 are total consumption in period 1 and 2, respectively.

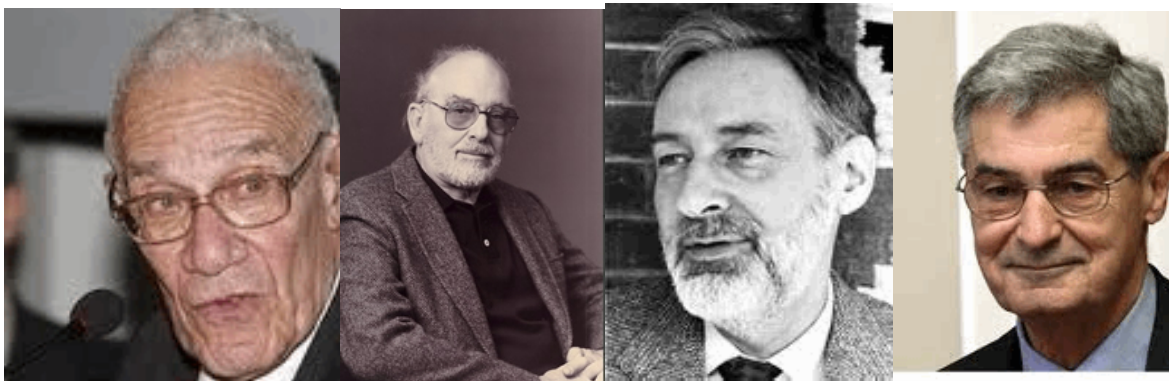
Moreover from the resource constraints we have $C_1 = w_A^1 + w_B^1 = 3$ and $C_2 = w_A^2 + w_B^2 = 4$. Thus, $k = \frac{4}{3}$ which implies that

$$1 + r_1 = \frac{\left(\frac{4}{3}\right)^{\frac{1}{\sigma}}}{\beta}$$

The saving functions will look the same as in the previous exercise; we just need to plug in this expression for r_1 .

Chapter 7

Growth



Robert Solow

David Cass

Tjalling Koopmans

Robert Lucas

What determines the rate of growth of consumption or income (as measured e.g., by Gross Domestic Product) ? What determines the rate of growth of *per-capita* consumption or *per-capita* income ?

Let variables in capital letters denote aggregate quantities, and lower case variables denote per-capita quantities; and let L denote the number of agents in the economy:

$$y = \frac{Y}{L}$$

Let an index t denote time: y_t is y at time t . Aggregate growth rates depend on the growth rate of the population, $\frac{L_{t+1}}{L_t}$:

$$\frac{Y_{t+1}}{Y_t} = \frac{y_{t+1}L_{t+1}}{y_tL_t} = \frac{y_{t+1}}{y_t} \frac{L_{t+1}}{L_t}$$

As a consequence per-capita growth are a better indicator of economic development. As an example, consider the average rate of growth of income in Mexico and in Italy in the period 1950-2000 (get them from the Penn World Dataset of Summers-Heston, at

http://pwt.econ.upenn.edu/php_site/pwt61_form.php)

While the growth rate of the population is in general related to the growth rate of per-capita consumption and income of an economy (and in interesting ways, when you endogeneize fertility and allow the agents choose how many children to have), in first approximation we consider it instead exogenous and we study the growth rate of per-capita consumption and income per-se.

7.0.1 Production

Aggregate production is described by a production function

$$Y_t = F(K_t, L_t)$$

where Y_t is aggregate income at time t , K_t is aggregate capital (in the same units as income) at t , and L_t is the population of (the total number of agents in) the economy at t ; assuming that each agent is the same (no productivity differences, e.g., related to education) and works the same number of hours, we also refer to L_t as the aggregate labor supply in the economy.

We assume the production function is Cobb-Douglas (this works well with data):

$$Y_t = A (K_t)^\alpha (L_t)^{1-\alpha}, \quad 0 < \alpha < 1$$

The marginal product of capital in the economy is:

$$MP_K = \frac{\partial Y_t}{\partial K_t} = \alpha A (K_t)^{\alpha-1} (L_t)^{1-\alpha}$$

Note it is decreasing in capital, K_t , and increasing in labor L_t . How do you see this? Compute the derivative of MP_K with respect to K_t and L_t , respectively.

The marginal product of labor in the economy is:

$$MP_L = \frac{\partial Y_t}{\partial L_t} = (1 - \alpha) A (K_t)^\alpha (L_t)^{-\alpha}$$

Note it is increasing in capital, K_t , and decreasing in labor L_t . How do you see this ? Compute the derivative of MP_L with respect to K_t and L_t , respectively.

The per-capita production function can be computed as follows (the Cobb-Douglas specification makes it very easy):

Pass to per capita income by dividing by L_t :

$$y_t = \frac{Y_t}{L_t} = A \frac{(K_t)^\alpha (L_t)^{1-\alpha}}{L_t} = A \frac{(K_t)^\alpha}{(L_t)^\alpha} \frac{(L_t)^{1-\alpha}}{(L_t)^{1-\alpha}}$$

that is

$$y_t = Ak^\alpha$$

The marginal product of capital in per-capita terms is:

$$\frac{\partial y_t}{\partial k_t} = \alpha Ak^{\alpha-1}$$

it is decreasing in per-capita capital k_t ; that is: *rich economies have lower marginal product of capital*; where of course "rich" means with high per-capita capital, and the marginal product of capital is in per-capita terms.

draw figure

7.0.2 Growth and Savings

Consider an economy in which all agents are identical (we are not interested in distributional issues, but only in per-capita variables); and live for 2 periods, t and $t + 1$.

The budget constraint of the representative agent at time t is:

$$c_t + (k_{t+1} - k_t) = A(k_t)^\alpha$$

Notice that $(k_{t+1} - k_t)$ is his savings, and also his investment in capital (he has no other ways of saving, and cannot or will not borrow). The agent's income is due to production at time t : $A(k_t)^\alpha$.

The budget constraint of the representative agent at time $t + 1$ is:

$$c_{t+1} = A(k_{t+1})^\alpha$$

and the agent consume all his production, since he will not be around next period.

The representative agent's preferences are as usual:

$$u(c_t) + \beta u(c_{t+1})$$

where we assume $0 < \beta < 1$; and also $u(c) = \ln c$ for simplicity.

The representative agent's maximization problem (after substitution his budget constraints) into the specification of preferences) is:

$$\max_{k_{t+1}} \ln(A(k_t)^\alpha - k_{t+1} + k_t) + \beta \ln(A(k_{t+1})^\alpha) \quad (7.1)$$

The first order condition (derive it yourself) can be written as:

$$\frac{c_{t+1}}{c_t} = \alpha A(k_{t+1})^{\alpha-1}$$

Notice¹ that the right-hand-side is the marginal product of capital (in per-capita terms) at time $t + 1$. Notice also that this is the same equation we derived when studying savings, only that the marginal product of capital takes the place of $(1+r)$. In fact this is the same problem; the agent is saving by investing in capital and receives a return equal to the marginal product of capital.

Finally, notice that the first order condition says that: *rich economies have lower rates of growth of consumption*; where, again, "rich" means with high per-capita capital, and the growth rate of consumption is in per-capita terms.

Is this last implication true in the data ? Look at the the Penn World Dataset of Summers-Heston, at

[http : //pwt.econ.upenn.edu/php_site/pwt61_form.php](http://pwt.econ.upenn.edu/php_site/pwt61_form.php)).

It is true in a subset of countries, Asian New Industrialized Countries (South Korea, Hong Kong, Thailand, ..), China which is "poor" in per-capita capital and has been growing at 8% recently, while Germany (and the rest of Europe), the U.S. have been growing at less than 3%. It is not true however for African and Latin American countries, who are both "poor" in per capita income (and capital) and have been growing very slowly if at all.

¹By substituting the budget constraint equations into the first order condition you can solve for the growth rate of capital in the economy, $\frac{k_{t+1}}{k_t}$. Try doing it (when you have some free time, it's easy but algebraically involved. The solution I got is:

$$\frac{k_{t+1}}{k_t} = \frac{\beta\alpha}{1 - \beta\alpha} \left(1 + A(k_t)^{\alpha-1}\right)$$

7.0.3 Other Determinants of Growth

Write the first order conditions of the maximization problem in (7.1) in terms of k_{t+1} and k_t . You will have an equation of the form:

$$k_{t+1} = \text{constant}_1 : (k_t)^\alpha + \text{constant}_2 : k_t, \quad \text{constant}_2 < 1 \quad (7.2)$$

Study this equation mathematically as a *difference equation* (but be careful here: the equation is derived as the first order condition of the maximization problem of an agent who lives only 2 periods; by considering it a difference equation we are implicitly assuming that the same equation would hold if the agents lived an infinite number of periods; in fact this is true, in the sense that an equation of the form (7.2) holds in this case, even though the expression for the constants is different). The equation has two steady states: $k = 0$ and $k^* > 0$. The steady state k^* is globally stable: all paths of k_t starting from any $k_0 > 0$ converge to it. We conclude that for $\alpha < 1$ the economy does not grow indefinitely!

There are two possible modifications in our basic model which are such that the economy grows indefinitely. We can then ask ourselves in these economy what are the determinants of the growth rate and hence attempt some explanations which help account for the low growth rates of Africa and Latin America. Both modifications require changing the production function.

Human Capital.

Suppose the production function is:

$$Y_t = (AK_t)^\alpha (h_t L_t)^{1-\alpha}$$

where h_t is an index of quality of labor, called *human capital*, and

$$h_{t+1} = \gamma h_t, \quad \gamma > 1$$

In per capita terms, the production function becomes:

$$y_t = (k_t)^\alpha (h_t)^{1-\alpha}$$

$$h_{t+1} = \gamma h_t$$

In this case, by proceeding exactly as in the previous section (do it as an exercise), we can compute the growth rate of consumption:

$$\frac{c_{t+1}}{c_t} = \beta \alpha A \left(\frac{h_{t+1}}{k_{t+1}} \right)^{1-\alpha}$$

Convince yourself that in this case the economy grows. The argument is as follows. Suppose not. Then $\frac{k_{t+1}}{k_t} \rightarrow 1$ for k_t large enough. $\frac{h_{t+1}}{h_t} \rightarrow \gamma > 1$ instead, independently of k_t . Therefore, for large k_t , $\frac{h_{t+1}}{k_{t+1}}$ increases over time, and the economy grows. The important question then is: *What does γ depend on ?* Partial answer for discussion:

The structure of the family

Fertility and its determinants

Schooling system

Urban Development

Technological Innovations and Knowledge.

Suppose the parameter A in the production function, which measures *total productivity* (sometimes we refer to it as the *general knowledge* of an economy), grows over time,

$$A_{t+1} = \delta A_t \quad (7.3)$$

Once again you can show that (think about how you would do this) in this case the economy will grow. Once again, the important question then is: *What does δ depend on ?* To develop a listing of partial answers for discussion a comment is very useful: *General knowledge* can be used by the whole economy without diminishing returns: think of a blueprint to produce a medicine; it can be freely copied and used by many firms. We say general knowledge is *non-rival*. *General knowledge* can be private or public depending on the institutions: think of the blueprint; we can protect it or not with a system of patents. We say that general knowledge is *excludable*. (Non-excludable goods exist: fishing in the sea has proved very hard to exclude over the years). What is the difference between institutions which guarantee *exclusion* and those which do not ? Think of the following model of total productivity or knowledge:

$$A_t = A(j_t)^{1-\alpha}$$

where j_t is knowledge and

$$j_t = \begin{cases} k_t & \text{under excludability} \\ \bar{k}_t & \text{under non-excludability} \end{cases}$$

and \bar{k}_t is a measure of average capital that the representative agent in the economy cannot affect with his capital investment. At equilibrium though all agents choose the same capital (there is a single representative agent), and hence $\bar{k}_t = k_t$. Compute now for both models the marginal product of capital in per capita terms:

$$\frac{\partial}{\partial k_t} A(\bar{k}_t)^{1-\alpha} (k_t)^\alpha = \alpha A(\bar{k}_t)^{1-\alpha} (k_t)^{\alpha-1} = \alpha A \text{ at equilibrium} \quad [excludable]$$

$$\frac{\partial}{\partial k_t} A(k_t)^{1-\alpha} (k_t)^\alpha = A \quad [non - excludable]$$

In the model then δ is endogenous and depends on the existence of institutions which guarantee excludability of general knowledge. If you are thinking that it might not be that A_{t+1} satisfies (7.3) in equilibrium, you are right. But this is not so important, and we can rig the model so that it does.

Partial answers to the *What does δ depend on ?* question for discussion (also look at the papers by Murphy-Shleifer-Vishny (1991) and Acemoglu-Robinson (2004) posted on my website):

The structure of the legal system

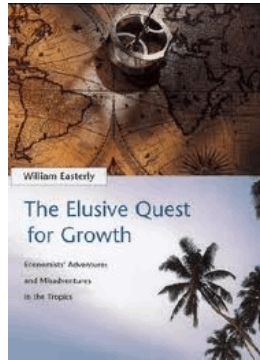
Property rights

Industrial organization and the patent system

Urban Development

A great (and simple to read) book to understand all this and more is:

W. Easterly, *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*, MIT Press, 2001.



Generosity of spirit, good intentions, and great accomplishments in entertainment



Ali Hewson and Bono



Angelina Jolie

are not always enough:

Chapter 8

Finance



Ken Arrow



Robert Merton



Myron Scholes and Fisher Black



Robert Lucas

In this chapter we shall see how a model-driven definition of risk is useful and counterintuitive. We shall also see that, an implication of this definition is that,

there are no free lunches, not even in finance!

8.1 Asset Pricing and the CAPM

Consider an agent living two periods, t and $t + 1$. His utility in terms of consumption c_t and c_{t+1} is:

$$u(c_t) + \beta u(c_{t+1})$$

where $\beta < 1$ is the discount rate. We assume $u(c)$ is smooth, strictly monotonic, and strictly concave, and the agent is an expected utility maximizer.

8.1.1 Portfolio Choice

The agent faces a wealth process w_t, w_{t+1} . He can trade J assets; asset $j = 1, \dots, J$ has payoff x_{t+1}^j at $t + 1$.

The agent's problem is the following:

$$\max_{\theta^j, j=1, \dots, J} u(c_t) + \beta E_t u(c_{t+1}) \quad (8.1)$$

subject to:

$$\begin{aligned} c_t &= w_t - \sum_{j=1}^J q_t^j \theta^j \\ c_{t+1} &= w_{t+1} + \sum_{j=1}^J x_{t+1}^j \theta^j \end{aligned}$$

where q_t^j is the price of asset j , and θ^j is the amount of asset j in the agent portfolio, the choice variable. The first order conditions for the maximization problem imply the fundamental asset pricing equation:

$$q_t^j = E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}^j \right) \quad (8.2)$$

If the asset is long lived, e.g., a stock, its payoff at $t + 1$ is the sum of its resale price, q_{t+1}^j and its cash flow at time $t + 1$, e.g., its dividend; we write $x_{t+1}^j = q_{t+1}^j + d_{t+1}^j$. In this case, the pricing equation becomes:

$$q_t^j = E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} (q_{t+1}^j + d_{t+1}^j) \right) \quad (8.3)$$

8.1.2 Time and Risk Correction

Let's examine the asset pricing equation (8.2). If there is no uncertainty, that is if the agent is able to perfectly insure, so that $c_t = c_{t+1}$, or if he is risk neutral, so that $u'(c)$ is a constant,

$$q_t^j = \beta E_t x_{t+1}^j$$

and the price of asset j is the net present value of its expected payoff at $t + 1$ discounted at the pure discount rate β .

If instead the agent is risk averse and his consumption is a stochastic process, then by (8.2), the discount factor is stochastic: $\beta \frac{u'(c_{t+1})}{u'(c_t)}$. In this case, the discount contains a risk correction. To see this, note that (8.2) can be written:¹

$$q_t^j = E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right) E(x_{t+1}^j) + \text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, x_{t+1}^j \right) \quad (8.4)$$

and the covariance of the asset payoff with the marginal rate of substitution of the agent is the relevant component of the risk of the asset which enters in the price.

Because of the concavity of the utility function $u(c)$ (that is, $u''(c) < 0$), if $\text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, x_{t+1}^j \right) > 0$, then $\text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, x_{t+1}^j \right) < 0$. Therefore, equation (8.4) implies that assets whose payoff is negatively correlated with the agent's consumption are more valued by the agent and hence have a higher price. Why are they more valued, in intuitive terms? Because they allow the agent to insure, that is to reduce the risk (roughly, variance) of his consumption process. Remember:

Agents care about the variance of their consumption and hence about the covariance of the asset's payoff with his consumption.

Note that the consumption which enters the stochastic discount factor is each single agent's consumption in the economy. If financial markets are developed enough (they are complete, in the economist's parlance), and utility functions are well behaved, then $\beta \frac{u'(c_{t+1})}{u'(c_t)}$ is equalized for any agent in equilibrium and we can think without loss of generality of c_{t+1} as of the economy's consumption.²

¹For any two random variables x and y ,

$$E(x \cdot y) = E x \cdot E y + \text{cov}(x, y).$$

²This is a fundamental result in economic theory. If you want to know more about it, look at J. Cochrane, *Asset Pricing*, Princeton University Press, 2001; this is not part of the course though.

8.1.3 Beta Representation

Often the pricing equation is written as an equation for excess returns. The return of asset j is written $R_{t+1}^j = \frac{x_{t+1}^j}{q_t^j}$. We can write then the pricing equation (8.2) as:

$$1 = E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} R_{t+1}^j \right) \quad (8.5)$$

Consider a risk free return R_{t+1}^f , that is, a return which is known at t . In this case equation (8.5) can be written:

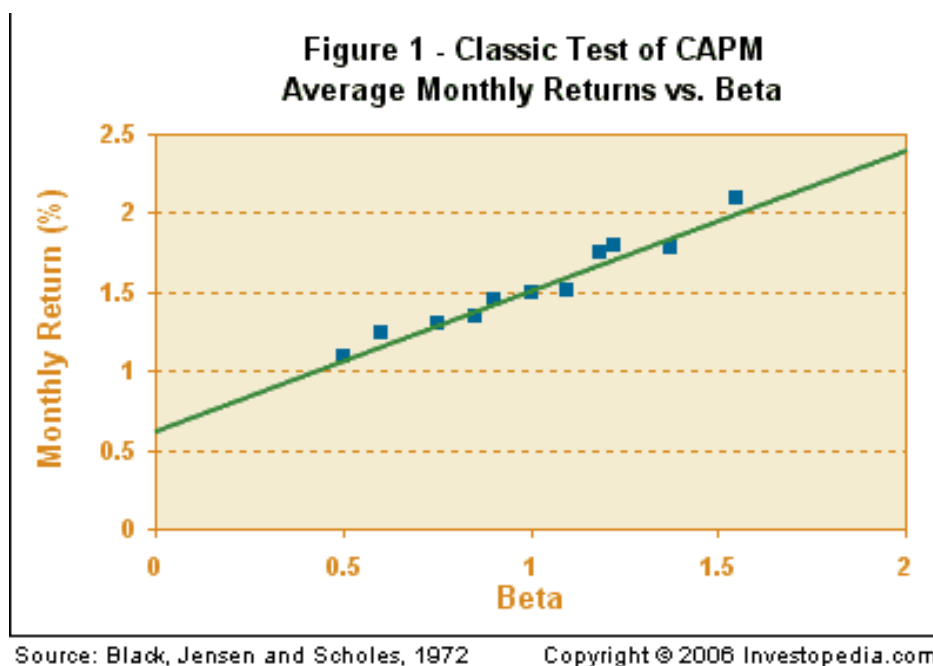
$$R_{t+1}^f = \frac{1}{E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \quad (8.6)$$

Tedious algebra (but nothing more than algebra) will let us now rewrite (8.4) as:

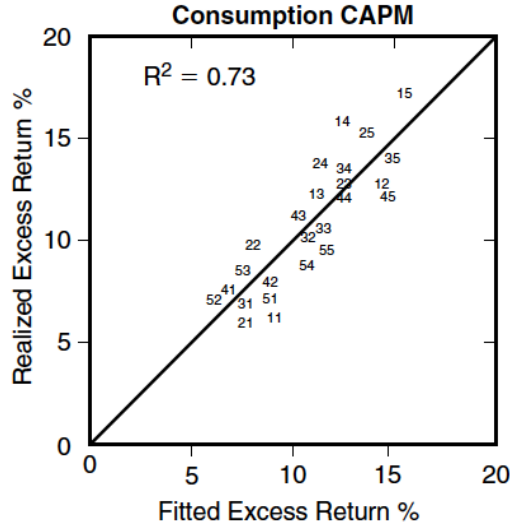
$$E \left(R_{t+1}^j - R_{t+1}^f \right) = \left(- \frac{\text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, R_{t+1}^j \right)}{\text{var} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \right) \left(\frac{\text{var} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)}{E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \right) \quad (8.7)$$

Think of $\beta^j = \left(- \frac{\text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, R_{t+1}^j \right)}{\text{var} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \right)$ as of a measure of the riskiness of return j , its conditional beta (not to be confused with the pure discount factor). Think of $\left(\frac{\text{var} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)}{E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \right)$ instead as a measure of the price of risk. Naturally, the price of risk is 0 if the variance of consumption is 0; and in this case, $R_{t+1}^j = R_{t+1}^f$. Moreover, note that if $\text{cov} \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, R_{t+1}^j \right) < 0$, then $R_t^j > R_t^f$, and the risk premium is positive.

In their 1972 study on *The Capital Asset Pricing Model: Some Empirical Tests*, Fischer Black, Michael C. Jensen and Myron Scholes documented the linear relationship between the financial returns of stock portfolios and their betas, with return data from New York Stock Exchange stocks between 1931 and 1965. Here's the figure:



A more modern figure, from John Cochrane, *Financial markets and the Real Economy*, mimeo (2007). Fitted returns are the returns implied by the Consumption CAPM model, that is, implied by the betas. If the model is correct the returns should be on the 45° line.



The 25 data points are each an asset portfolio which is known not to work very well with the theory (these are called Fama-French 25 portfolios).

8.1.4 The Term Structure of Interest Rates

The risk free rate is usually directly available in the market; it is the rate of return of a 1 period T-bill (as risk-free as possible a bond). Let's define a period to be one year, for simplicity. Risk free rates for many different maturities are traded in financial markets, a quarter, a year, 5 years, 10 and 20 years. Typically, a longer maturity is compensated by a higher yearly return (all our returns are yearly, for comparison). This smell of risk! But is there risk in a 2-year risk free bond? This seems meaningless, at first reading (if it's risk free it's risk free), but it is not. The risk is implicit in the fact that there are two strategies to transfer income from time t to time $t + 2$:

1. Buying a 2 year risk free bond, which pays $R_{t,t+2}^f$ with certainty at time $t + 2$ for a price of 1 at t .
2. Buying a 1 year risk free bond, which pays $R_{t,t+1}^f$ with certainty at time $t + 1$ for a price of 1 at t ; then, at time $t + 1$ investing $R_{t,t+1}^f$ in a 1 year risk free bond, which pays $R_{t,t+1}^f \cdot R_{t+1,t+2}^f$ with certainty at time $t + 2$.

Note that the risk free rate of a 1 year bond at $t + 1$ will be known at $t + 1$, but it is not known at time t . To convince yourself of this, write the asset pricing equation for $R_{t+1,t+2}^f$:

$$\left(R_{t+1,t+2}^f\right)^{-1} = E\left(\beta \frac{u'(c_{t+2})}{u'(c_{t+1})}\right),$$

and note that at time t , c_{t+1} is not known. Hence the uncertainty. (Why is strategy 1 also not riskless?)

We can then compute the value of a 2 year risk free bond:

$$\left(R_{t,t+2}^f\right)^{-1} = E\left(\beta^2 \frac{u'(c_{t+2})}{u'(c_t)}\right). \quad (8.8)$$

From (??) and (8.8) one gets an expression for the term structure of interest rates:

$$\left(R_{t,t+2}^f\right)^{-1} = \left(R_{t,t+1}^f\right)^{-1} E\left(R_{t+1,t+2}^f\right)^{-1} + cov\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, \left(R_{t+1,t+2}^f\right)^{-1}\right). \quad (8.9)$$

If $cov\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, \left(R_{t+1,t+2}^f\right)^{-1}\right) = 0$ it is either because there is no risk in the economy, $\beta \frac{u'(c_{t+1})}{u'(c_t)}$ is constant, or because $R_{t+1,t+2}^f$ is known at time t : $R_{t+1,t+2}^f = E\left(R_{t+1,t+2}^f\right)$. In this case then we have

$$R_{t,t+2}^f = R_{t,t+1}^f \cdot R_{t+1,t+2}^f$$

That is, the yearly interest rates are the same at different maturities and strategies 1 and 2 are equivalent. Otherwise, a risk correction appears.

8.1.5 The Market Rate of Return and the betas

If we knew $u(c)$ and we could measure precisely c_{t+1} , c_t (aggregate consumption), we would have no problem; and we could compute

$$R_t^m = \beta \frac{u'(c_{t+1})}{u'(c_t)}$$

and

$$\beta^j = \left(\frac{\text{cov}_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, R_{t+1}^j \right)}{\text{var}_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right)} \right)$$

In fact we do not know what $u(c)$ and aggregate consumption data is quite bad. Therefore the standard procedure is to estimate a multi-factor beta equation of the form:

$$R_t^j = R_t^f + \beta_{f^1}^j f_t^1 + \beta_{f^2}^j f_t^2 + \dots + \epsilon_t^j \quad (8.10)$$

for any asset j ; where f^i are factors, i.e., proxies for the intertemporal marginal rate of substitution, like several indices of stock returns, GNP, inflation, and so on. (Note that only unconditional betas can be estimated; hence under the assumption that betas are constant over time).

8.1.6 Equivalent Risk Neutral Representation

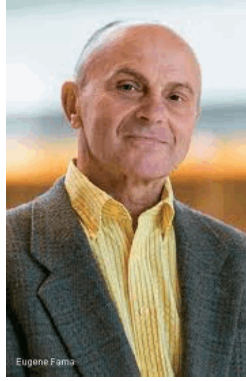
Consider the random variable x_{t+1}^j ; $E(x_{t+1}^j)$ is its expected value with respect to our basic probability distribution. Let m_{t+1} be a non-negative random variable, and let $E^*(x_{t+1}^j) = E\left(\frac{m_{t+1}}{E_t m_{t+1}} x_{t+1}^j\right)$. Notice that we have implicitly defined a different probability distribution with respect to the basic one.³ Recalling that $R_{t+1}^f = \frac{1}{E\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}\right)}$, and letting $m_{t+1} = \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}\right)$ we now have:

$$q_t^j = \frac{E^* x_{t+1}^j}{R_{t+1}^f} \quad (8.11)$$

Notice that (8.11) is the pricing equation of an economy with risk neutral agents with the different probability distribution over uncertainty we constructed; we call therefore this probability distribution "risk neutral". The *risk neutral probability distribution* corrects for risk. In fact, with respect to the basic probability distribution, the risk neutral probability distribution weights more the states with lower consumption. This is to say: a risk averse agent is like a risk neutral agent who believes that the "bad" states of uncertainty are more probable.

³That is, $E_t^*(x) \geq 0$ if $x \geq 0$ and $E_t^*(x) = x$ if x is constant.

8.2 Efficient Market Hypothesis



Eugene Fama

The analysis of asset pricing and valuation in the previous section is often referred to as the *Efficient Market Hypothesis*. It has several implications that have been tested over the years. For a critical review, see A. Schleifer, *Inefficient Markets*, Oxford University Press, 2000.

8.2.1 Idiosyncratic Risk Does Not Affect Prices

An asset j is one of idiosyncratic risk if its payoff is independent of the economy's aggregate consumption. You can check immediately that *idiosyncratic risk is not priced* as its covariance with $\beta \frac{u'(c_{t+1})}{u'(c_t)}$ is zero (see equation 8.4). Note that this is true no matter how large is $\text{var}(x_{t+1}^j)$. What is idiosyncratic risk? Typically most accident risks as well as some health risk are idiosyncratic, that is, independent across agents. (But, Is hurricane risk idiosyncratic?) Idiosyncratic risk is not priced if the cost of an insurance contract on it is the expected present discounted (at the risk-free interest rate) of its net payments.

8.2.2 Prices Adjust Immediately to All Available Information

In the fundamental pricing equation (8.2), the information included in the price at time t is the whole information available at time t , on which the

expectation operator $E(\cdot)$ is conditional. The Efficient Market Hypothesis does not allow for differential information across agents in the market. This is a major (and somewhat unrealistic) assumption.

8.2.3 Risk Adjusted Prices Are Not Predictable

Suppose agents are risk neutral ($u(c)$ is linear) and do not discount the future (β close to 1). Consider a stock j which pays no dividends, $x_{t+1}^j = q_{t+1}^j$; then, using (8.2):

$$q_t^j = E q_{t+1}^j \quad (8.12)$$

and its price is a *random walk* (or, more precisely, a *martingale*). This is the typical example of a non-predictable stochastic process. It relies on extreme implausible assumptions.

In general prices are predictable, as are returns and excess returns. For instance, from (8.7), one sees that even if the price of risk $\left(-\frac{\text{var}\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}\right)}{E\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}\right)} \right)$ is constant over time, the predictability of excess returns follows from the predictability of the conditional beta. So it is, for instance, that stocks pay a higher return than bonds, consistently over time, as stocks are more risky than bonds and hence pay a higher excess return.

Let *risk adjusted prices* be defined as prices minus the risk premium. The fact that *risk adjusted stock prices are unpredictable* is an immediate consequence of the asset pricing equation for stocks, 8.3, with no dividends, when written as:

$$u'(c_t)q_t^j = E\left(\beta u'(c_{t+1})q_{t+1}^j\right).$$

Another way to say the same thing is the following: the price of a stock which pays no dividends is always a martingale in the risk neutral probability measure (which adjusts for risk), see equation (8.11), but not for the basic probability measure (which drives the uncertainty in the economy).

8.2.4 Arbitrage Opportunities Are Not Present in Financial Markets

An *arbitrage* opportunity exists in financial markets when a portfolio can be traded whose price is either 0 or negative and its payoff is always non-negative and strictly positive with positive probability. At equilibrium, however, asset

prices must be such that arbitrage opportunities cannot exist. Otherwise, any agent, exploiting the opportunities would have infinite wealth and would demand an infinite consumption allocation in some states, which contradicts market clearing. In fact, the discount factor, $\beta \frac{u'(c_{t+1})}{u'(c_t)}$ is strictly positive (utility is strictly monotonic, and hence $u'(c) > 0$). Any asset with a payoff which is non-negative and strictly positive with some probability has then a strictly positive price, by (8.2).

To better understand the implications of *No-arbitrage* as a property of equilibrium, consider an economy with two states of nature $\{s_1, s_2\}$ and two primary assets, $\{1, 2\}$. The first asset is a stock with price q_1 and returns $(x_{s_1}^1, x_{s_2}^1)$, $x_{s_1}^1 \neq x_{s_2}^1$. There is also a bond with price q_2 and returns (x^2, x^2) . Taking this data as parameters, consider the problem of pricing a derivative, asset d , with returns $(x_{s_1}^d, x_{s_2}^d)$. Denote the price of the derivative asset as q_d . Construct a portfolio with the two primary assets which replicates the payoff vector of the new asset. This requires solving the following system

$$\begin{pmatrix} x_{s_1}^d \\ x_{s_2}^d \end{pmatrix} = \theta_1 \begin{pmatrix} x_{s_1}^1 \\ x_{s_2}^1 \end{pmatrix} + \theta_2 \begin{pmatrix} x^2 \\ x^2 \end{pmatrix}.$$

Note that the solution to this system exists and is unique.

Now, *No-arbitrage* implies that the *replicating portfolio* and the derivative asset must have the same price since they have the same returns vector:

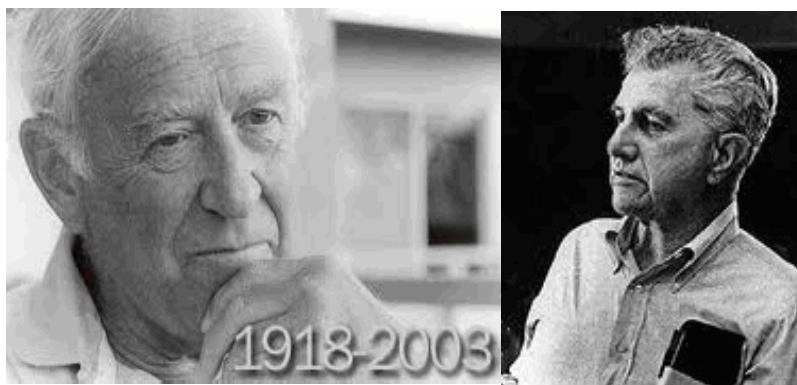
$$q_d = \theta_1 q_1 + \theta_2 q_2.$$

This is called the *Law of one price*. Prove that the *Law of one price* is a consequence of No-arbitrage, that is,

if either $q_d > \theta_1 q_1 + \theta_2 q_2$ or $q_d < \theta_1 q_1 + \theta_2 q_2$ an agent could construct a traded portfolio whose price is negative and its payoff strictly positive with probability 1.

Pricing derivatives by constructing a replicating portfolio and then using the *Law of one price* is one of the main activities of Wall Street. Though the derivatives are typically quite complex, conceptually this activity is a straightforward as exposed in this notes.

8.3 Modigliani-Miller Theorem



Franco Modigliani

Merton Miller

Consider an economy for which the Efficient Market Hypothesis holds and in which all financial streams are possibly traded (abusing words we call this *Complete Markets* assumption). Then

MM1 The value of the firm is independent of its capital structure, that is, of the proportion of debt and equity used to finance the firm's operations.

MM2 The value of the firm is independent of the firm's dividend policy.

To understand the irrelevance propositions, MM1 and MM2, and the assumptions required for them to hold, it is useful to pass to more formal statements.

8.3.1 Modigliani-Miller More Formally

Let's consider MM1 first. Consider a firm whose cash flow at $t + 1$ is a random variable x_{t+1}^j . The cash flow is composed of the dividend plus the re-sale price,

$$x_{t+1}^j = d_{t+1}^j + q_{t+1}^j$$

and let $d_{t+1}^j \geq \bar{d} > 0$.

Suppose the firm finances its operations by issuing *equity only*. The *Efficient Market Hypothesis* implies that any cash flow x is priced by our basic valuation equation, (8.2).

Then the firm's value is

$$q_t^j = E \left(\beta \frac{u'(c_{t+1})}{u'(c_{t+1})} x_{t+1}^j \right).$$

Suppose instead the firm issues α units of debt (assume, but only for simplicity, that $\alpha < \bar{d}$, so that debt is risk free) and the residual is equity. What is the value of the firm? It is the value of the debt and equity issued. The cash flow of the debt at $t + 1$ is α ; while the cash flow of the equity is $x_{t+1}^j - \alpha$. The value of the firm in this case is therefore:

$$E \left(\beta \frac{u'(c_{t+1})}{u'(c_{t+1})} (\alpha + (x_{t+1}^j - \alpha)) \right) = E \left(\beta \frac{u'(c_{t+1})}{u'(c_{t+1})} x_{t+1}^j \right) = q_t^j.$$

In other words, we have proved MM1. Let's try with an intuition. Consider the change in the capital structure of the firm, from all equity to equity and α bonds. In equilibrium, the agents must hold the new debt and the equity, after the change in capital structure. Will they need to be compensated for this change? No, because the agents will not have to change their portfolios in equilibrium: they can reproduce whatever it was their portfolio with no change; as long as the aggregate supply of the cash flows is unchanged and all portfolios can be reproduced (this is the assumption that all financial streams are possibly traded in our economy). The supply of bonds in the economy is increased, as the firm is issuing α extra units. But the demand of bonds is increased exactly to compensate the change in the supply, as those agents holding equity before the change will have to buy the α bonds after the change to keep their portfolio unchanged.

The original proof of the result by Modigliani-Miller in *American Economic Review* 1958 was a No-arbitrage proof. It helps develop intuition for this result. Consider two firms, identical in terms of cash flow, but financed differently: one with only equity, the other with α units of a bond security and the residual as equity. Suppose they have a different price. How could an agent trading in bonds and in the original equity security (with payoff x_{t+1}^j) make a positive arbitrage?

Consider the arguments we just made to prove MM1. Do they depend on the assumption that $d_{t+1}^j > \bar{d} > \alpha$, so that the debt issued by the firm is risk free? Convince yourself that the answer is no.

Let's consider MM2 now.

Consider now a change in the firm's dividend policy. For instance, suppose that the firm's dividends at time $t + 1$ go from d_{t+1}^j to $d_{t+1}^j + e_{t+1}$, for some

random variable e_{t+1} . How does the firm finance its new dividend policy without changing its investment plan? The firm has to add a security to its portfolio which pays e_{t+1} at time $t+1$. Such a security exists by the complete markets' assumption. The price at time t of a security which pays e_{t+1} at time $t+1$ is

$$q_t^e = E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} e_{t+1} \right) \quad (8.13)$$

What is then the value of the firm with the new dividend policy? It is the value of the firm's payoff at time $t+1$, appropriately discounted, minus the cost of the security that the firm is buying at time t :

$$E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} (d_{t+1}^j + e_{t+1}) \right) - q_t^e$$

Combining this equation with (8.13), we have that

$$q_t^j = E \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}^j \right)$$

and we have proved MM2.

Consider the cases in which the firm finances the new issue of dividends with cash that it holds at time t or with cash generated by its project at time $t+1$. Show that the same MM2 result applies in this case.

8.3.2 Other Implicit Assumptions

Several implicit assumptions of practical (but not conceptual) relevance required for MM1 and MM2:

No transaction costs.

*No differential taxation of debt and equity.*⁴

No costs of financial distress (in particular, no bankruptcy costs).

Managers maximize shareholders' wealth, that is, the value of the firm.

⁴In fact, in the U.S., interest rates are not taxed at the corporate level, but dividends and capital gains are. This gives an advantage to debt as a form of financing, contrary to the MM results. On the other hand, the personal tax on equity is higher than the personal tax on debt. This gives an advantage to equity financing. The advantages roughly cancel out.

8.4 Problems

Problem 8.4.1 Consider an economy with two financial assets: A bond, which pays a riskless return R ; and a stock, which pays a return S , a random variable taking values S_1 with probability p and S_2 with probability $1 - p$. Consider an agent with an arbitrary sum of money, w , choosing the share α of bond and $(1 - \alpha)$ of stock in his portfolio to maximize the expected utility of his consumption after returns are paid. i) Write down the agent's maximization problem, by choice of α . ii) Write down the first order condition for the maximization problem. (Careful to corners). iii) Examine the condition (being careful about corners once again). Under which conditions on the parameters R, S_1, S_2 is $\alpha = 1$? Under which conditions is it < 1 ?

Problem 8.4.2 Consider an economy with 2 states of the world $s_1 = H$, $s_2 = L$. 3 assets are traded in this economy: asset 1 has payoff $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (that is, payoff equal to 1 in state H and payoff equal to 0 in state L); asset 2 is a riskfree bond and has payoff $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; asset 3 has payoff $x_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. The prices of the 3 assets are q_1, q_2, q_3 . Suppose now you want to price by no-arbitrage a derivative asset whose payoff is $x^d = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$. Do this as a function of prices q_1, q_2 . Do this again as a function of prices q_1, q_3 . Does no arbitrage impose any restriction on prices q_2 , and q_3 ?

Problem 8.4.3 Consider an agent with the following preferences:

$$u(c) = \frac{1}{1 - \sigma} c^{1 - \sigma}, \quad \sigma > 0$$

In an economy with 2 time periods (and utility as above for each period and discount rate $\beta = .9$) and 4 states of the world with associated probabilities $\pi_1, \pi_2, \pi_3, 1 - \pi_1 - \pi_2 - \pi_3$, what is the expression of the price of an asset

whose payoff is $x = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 4 \end{bmatrix}$?

Problem 8.4.4 Consider an agent living 3 periods, with logarithmic preferences, and with endowments w_t in period $t = 1, 2, 3$. He/she does not discount the future and faces an interest rate $r_1 > 0$ in period 1 and an interest rate $r_2 = 0$ in period 2. What is the effect of a change (e.g., a marginal increase) in the interest rate in period 1 on the agent's consumption in period 2 and in period 3.

Problem 8.4.5 Consider an agent with the following preferences:

$$u(c) = \frac{1}{1-\sigma} c^{1-\sigma}, \quad \sigma > 0$$

How does risk aversion depend on σ ? (Define first risk aversion, there are several definitions, make up one that makes sense to you.) Can σ be negative?

Problem 8.4.6 Consider a risk averse agent. He faces a health risk of D (he/she has to go to the hospital and pay D , and he/she'll be fine) with probability π . He/she can buy insurance at price q (that is he/she can buy at cost q a contract that pays 1 dollar if he/she has to go to the Hospital). How many units of the contract will the agent buy if the price is $q = \pi$?

Problem 8.4.7 On tradesports.com a contract is traded which pays 100 dollars if Bush is elected president in the next presidential election. The cost of the contract is 56 dollars. Does this tell you something about the expected value of Bush winning? (Argue very carefully.) And about the variance? What is the covariance of Bush winning and Kerry winning? (Does Nader have something to do with your last answer?)

Answer. First, consider the case where all the people buying and selling the contract are risk neutral. This means that the only thing they are interested in is the expected value of the contract. If the expected value of buying the contract is greater than 0, the demand for the contract will be $+\infty$. For each contract that an agent buys, their expected wealth goes up.

As they are risk neutral, they want to maximize their expected wealth, so however many contracts an agent owns, she will always want more! Similarly, if the expected value of the contract is less than zero, the supply of the contract will be $+\infty$. For every contract an agent sells (remember, an agent can have a negative number of contracts - this just means that they are acting as the bookmaker - they accept \$56, and pay out \$100 if Bush wins), their expected wealth goes up. Again, however many contracts an agent has sold, they will always want to sell one more. It follows that, for the market to be in equilibrium (i.e., for supply to equal demand), the expected value of the contract must equal zero. At this point, agents are indifferent between buying and not buying the contract

Let p be the probability of Bush winning. The expected value of the contract is therefore:

$$\begin{aligned} p(100 - 56) + (1 - p)(-56) &= 0 \\ \Rightarrow p &= \frac{56}{100} = .56 \end{aligned}$$

What if the agents are not risk neutral? We would still argue that for the market to be in equilibrium, the agent would have to be indifferent between buying and not buying the contract. But if the agent is risk averse, then the expected *value* of a bet has to be greater than zero to make the expected utility of the bet zero (remember, risk averse people don't like gambling - to make them take a gamble the odds have to be better than fair). We therefore know that

$$\begin{aligned} p(100 - 56) + (1 - p)(-56) &> 0 \\ \Rightarrow p &> \frac{56}{100} = .56 \end{aligned}$$

However, without a specific utility function, we cannot say more than this.

Now, to answer questions about the expected value, variance and covariance of Bush winning, we need to define a random variable, X , which is equal to 1 if bush wins and zero otherwise:

State of the World	Probability	Value of X
--------------------	-------------	------------

Bush wins	P	1
-----------	-----	---

Bush Does not win	$1 - p$	0
-------------------	---------	---

Again let p be the probability of bush winning. Employing the formula for expected value, we see that

$$E(X) = p1 + (1 - p)0 = p$$

Or, if we are in the risk neutral world, $E(X) = 0.56$

Employing the formula for variance, we see that

$$\begin{aligned} Var(X) &= E[(x - E(X))^2] \\ &= p(1 - p)^2 + (1 - p)(-p)^2 \\ &= p(1 - p)(1 - p + p) \\ &= p(1 - p) \end{aligned}$$

Or, in the risk neutral world, $Var(X) = 0.2464$

To talk about the covariance between Bush winning and Kerry winning, we need to define another random variable, Y , equal to 1 if Kerry wins and zero otherwise. First consider the case where there is no Nader, or the probability of Nader winning is zero. In this case, there are still only two states of the world, one in which Bush wins and Kerry does not, and another in which Kerry wins and Bush does not:

State of the World	Probability	Value of X	Value of Y
Bush wins (Kerry does not)	p	1	0
Bush Does not win (Kerry does)	$1 - p$	0	1

Employing the formula for covariance, and noting that $E(Y) = 1 - p$, we see that

$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= p(1 - p)(-(1 - p)) + (1 - p)(-p)(1 - (1 - p)) \\ &= -p(1 - p)^2 - p^2(1 - p) \\ &= -p(1 - p) \end{aligned}$$

In the risk neutral case, $cov(X, Y) = -.2464$

Now consider the case where there is a Nader. Here, we have to add a third state of the world, in which Nader wins. We also need to introduce a new probability, q of Nader winning

State of the World	Probability	Value of X	Value of Y
Bush wins	p	1	0
Kerry wins	$1 - p - q$	0	1
Nader wins	q	0	0

Again, employing the covariance formula, and noting that $E(Y) = 1 - p - q$

,

$$\begin{aligned} \text{cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= p(1 - p)(-(1 - p - q)) + (1 - p - q)(-p)(1 - (1 - p - q)) + q(-p)(1 - p - q) \end{aligned}$$

Note that even in the risk neutral case, we cannot solve explicitly for $\text{cov}(x, y)$ in this case, because we do not know q .

Problem 8.4.8 Consider an agent living 3 periods, with logarithmic preferences, and with endowments w_t in period $t = 1, 2, 3$. He/she does not discount the future and faces an interest rate $r_1 > 0$ in period 1 and an interest rate $r_2 = 0$ in period 2. What is the effect of a change (e.g., a marginal increase) in the interest rate in period 1 on the agent's consumption in period 2 and in period 3.

Answer. This is just a standard intertemporal optimization problem. In general, the utility function that the agent is trying to maximize would be of the form

$$U(c_1, c_2, c_3) = u(c_1) + \beta u(c_2) + \beta^2 u(c_3)$$

The question tells us that the utility function in this case is

$$U(c_1, c_2, c_3) = \ln c_1 + \ln c_2 + \ln c_3$$

We can construct the budget constraint in the usual manner. Define savings between period one and two and between two and three respectively as

$$\begin{aligned} s_1 &= w_1 - c_1 \\ s_2 &= w_2 + s_1(1 + r_1) - c_2 \end{aligned}$$

Note that savings at the end of period 3 will be zero, as the agent 'dies' at this point, so:

$$0 = w_3 + s_2(1 + r_2) - c_2$$

By using the last equation to get an expression for s_2 , substituting this into the second equation, getting an expression for s_1 and substituting this into the first equation, we get the present value budget constraint:

$$c_1 + \frac{c_2}{(1 + r_1)} + \frac{c_3}{(1 + r_1)(1 + r_2)} = w_1 + \frac{w_2}{(1 + r_1)} + \frac{w_3}{(1 + r_1)(1 + r_2)}$$

Though for this problem we are told that r_2 is equal to 0.

The problem therefore becomes

$$\begin{aligned} \text{maximize } U(c_1, c_2, c_3) &= \ln c_1 + \ln c_2 + \ln c_3 \\ \text{subject to: } c_1 + \frac{c_2}{(1 + r_1)} + \frac{c_3}{(1 + r_1)} &= w_1 + \frac{w_2}{(1 + r_1)} + \frac{w_3}{(1 + r_1)} \end{aligned}$$

Which gives the Lagrangian:

$$\begin{aligned} L(c_1, c_2, c_3, \mu) &= \ln c_1 + \ln c_2 + \ln c_3 - \mu \left(c_1 + \frac{c_2}{(1 + r_1)} + \frac{c_3}{(1 + r_1)} - w_1 - \frac{w_2}{(1 + r_1)} \right. \\ &\quad \left. - \frac{w_3}{(1 + r_1)} \right) \end{aligned}$$

Which gives first order conditions:

$$\begin{aligned} \frac{\partial L}{\partial c_1} &= \frac{1}{c_1} - \mu = 0 \\ \frac{\partial L}{\partial c_2} &= \frac{1}{c_2} - \frac{\mu}{(1 + r_1)} = 0 \\ \frac{\partial L}{\partial c_3} &= \frac{1}{c_3} - \frac{\mu}{(1 + r_1)} = 0 \end{aligned}$$

The last two expressions combine to give $c_2 = c_3$, while the first and second combine to give

$$c_2 = (1 + r_1)c_1$$

substituting this into the budget constraint gives

$$\begin{aligned}
 c_1 + \frac{c_2}{(1+r_1)} + \frac{c_3}{(1+r_1)} &= \\
 c_1 + \frac{(1+r_1)c_1}{(1+r_1)} + \frac{(1+r_1)c_1}{(1+r_1)} &= \\
 c_1 + c_1 + c_1 &= 3c_1 = w_1 + \frac{w_2}{(1+r_1)} + \frac{w_3}{(1+r_1)} \\
 c_1 &= \frac{1}{3} \left(w_1 + \frac{w_2}{(1+r_1)} + \frac{w_3}{(1+r_1)} \right)
 \end{aligned}$$

And using the expression we have for c_2 (and therefore c_3) as a function of c_1 gives

$$\begin{aligned}
 c_2 &= (1+r_1)c_1 \\
 &= (1+r_1) \frac{1}{3} \left(w_1 + \frac{w_2}{(1+r_1)} + \frac{w_3}{(1+r_1)} \right) \\
 &= \frac{1}{3} ((1+r_1)w_1 + w_2 + w_3) \\
 &= c_3
 \end{aligned}$$

The marginal change of consumption of period two and three with respect to the interest rate in period one is therefore

$$\frac{\partial c_2}{\partial r_1} = \frac{\partial c_3}{\partial r_1} = \frac{w_1}{3}.$$

Problem 8.4.9 Consider an agent with the following preferences:

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

How does risk aversion depend on σ ? (Define first risk aversion, there are several definitions, make up one that makes sense to you.) Can σ be negative?

Answer. There are many possible ways of measuring risk aversion, some of which we will talk about in class. One popular one is the coefficient of relative risk aversion, defined as

$$r = \frac{-cu''(c)}{u'(c)}$$

which is positive for risk averse people, zero for risk neutral people and negative for risk lovers. Plugging into this formula for the utility function above gives:

$$r = \sigma$$

or in other words, the coefficient of relative risk aversion is constant. σ can be negative. This just implies that the agent is a risk lover.

Problem 8.4.10 *Consider a risk averse agent. He faces a risk of incurring a cost D_1 with probability .3 and of D_2 with probability .4. He/she can buy insurance at price q , that is he/she can buy at cost q a contract that pays 1 dollar if he/she either cost D_1 or D_2 are realized. How many units of the contract will the agent buy is the price is $q = 3$.*

Answer. Think about an agent who buys x units of insurance. If they get ill, the amount of money that they will have is $W - D - qx + x$, where W is initial wealth. If they do not get ill, their wealth will be $W - qx$. The expected utility of such an agent is:

$$U = \pi u(W - D - qx + x) + (1 - \pi)u(W - qx)$$

where $u(\cdot)$ is some arbitrary decreasing marginal returns utility function. The agent wishes to choose x to maximize their expected utility. We therefore take the derivative of the above function with respect to x and set it equal to zero.

$$\begin{aligned} \frac{\partial U}{\partial x} &= \pi(1 - q)u'(W - D - qx + x) - (1 - \pi)qu'(W - qx) = 0 \\ \Rightarrow \pi(1 - q)u'(W - D - qx + x) &= (1 - \pi)qu'(W - qx) \end{aligned}$$

The question tells us that $q = \pi$, so we have:

$$\begin{aligned} \pi(1 - \pi)u'(W - D - \pi x + x) &= (1 - \pi)\pi u'(W - \pi x) \\ u'(W - D - \pi x + x) &= u'(W - \pi x) \end{aligned}$$

We would like to conclude from this that $W - D - \pi x + x = W - \pi x$, and we will do. But note that we can only do that under some assumptions. We are assuming that if the slope of the function is equal, then the argument of the function is also equal. This is only true if the slope of the function is not the same for any two points. One way to ensure this is to assume that the second derivative of the function is always less than 0.

Bearing that in mind, we get:

$$\begin{aligned} W - D - \pi x + x &= W - \pi x \\ -D + x &= 0 \\ D &= x \end{aligned}$$

This makes sense. In this case, the expected value of buying insurance is zero, as the cost of one unit of insurance is $q = \pi$, and the expected return on the insurance is $1\pi + 0(1 - \pi) = \pi$. Buying insurance therefore doesn't reduce one's expected wealth, but does reduce risk. A risk averse agent will therefore buy enough insurance to completely cover their risk. This is what happens in the case above: The agent keeps buying insurance until her income is the same in either state of the world.

Problem 8.4.11 *Consider an economy under uncertainty: there are three states of the world, s_1, s_2, s_3 , with probability $p, q, 1 - p - q$, respectively. All agents are identical and live for 2 periods, $t, t + 1$. Agents in this economy at time t can trade at price Q an asset which pays A_1, A_2, A_3 in the three states s_1, s_2, s_3 at time $t + 1$. i) What is the expected value of the payoff of the asset at time $t + 1$? ii) Consider the case of risk neutral agents - more specifically, agents with utility*

$$u(x_t) + \beta u(x_{t+1}) = C + \alpha x_t + \beta(C + \alpha x_{t+1})$$

for some constant C . What is the price of the asset at time t ? (What I mean here is: write down the equation of Q as a function of all exogenous variables in the economy - $p, q, A_1, A_2, A_3, C, \beta, \alpha$). iii) Consider the case of risk averse agents - more specifically, agents with utility

$$u(x_t) + \beta u(x_{t+1}) = \log x_t + \beta \log x_{t+1}$$

Assume also that $x_t = 1$, and x_{t+1} is x_1, x_2, x_3 in the three states s_1, s_2, s_3 , with $x_1 = 2, x_2 = 3, x_3 = 1$. What is the price of the asset at time t ? (Once again, what I mean here is: write down the equation of Q as a function of all exogenous variables in the economy - $p, q, A_1, A_2, A_3, \beta$).

Problem 8.4.12 *Consider an economy with 2 states of the world, either "a tax cut is voted" or "a tax cut is not voted." The tax cut is voted with probability .3. Consider a financial market with 2 basic assets. The first pays*

2 dollars if the tax cut is not voted and 3 if it is. The second asset pays nothing if the tax cut is not voted and 1 dollar if it is. The prices of the two assets are, respectively 3 and 1 dollars. There is also another asset, a risk free bond that pays 10 dollars no matter if the tax cut is voted or not. This last asset sells for 9 dollars. i) Is there an arbitrage opportunity in this economy? ii) Does your answer depend on the probability of the tax cut? Explain why. iii) Just for the fun of it, compute the variance of the two basic assets, and their covariance.

Problem 8.4.13 Consider an economy with two states of nature $\{s_1, s_2\}$ and with two primary assets, $\{1, 2\}$. The first asset is a stock with price q_1 and returns $(x_{s_1}^1, x_{s_2}^1)$, $x_{s_1}^1 \neq x_{s_2}^1$. There is also a bond with price q_2 and returns $(x_{s_1}^2, x_{s_2}^2)$. Which general form do prices q_1 and q_2 have for the financial market to display no arbitrage opportunities? Taking this data as parameters, consider the problem of pricing a derivative, asset d , with returns $(x_{s_1}^d, x_{s_2}^d)$. Denote the price of the derivative asset as q_d . Construct a formula for this price as a function of the parameters of the economy.

Answer. No Arbitrage implies

$$\begin{aligned} q_1 &= m_1 x_{s_1}^1 + m_2 x_{s_2}^1 \\ q_2 &= m_1 x_{s_1}^2 + m_2 x_{s_2}^2 \end{aligned}$$

with m_1 and m_2 both > 0 .

First solve for θ_1 and θ_2 the following linear equation system:

$$\begin{pmatrix} x_{s_1}^d \\ x_{s_2}^d \end{pmatrix} = \theta_1 \begin{pmatrix} x_{s_1}^1 \\ x_{s_2}^1 \end{pmatrix} + \theta_2 \begin{pmatrix} x_{s_1}^2 \\ x_{s_2}^2 \end{pmatrix}$$

then

$$q_d = \theta_1 q_1 + \theta_2 q_2$$

Can you always (for all values of the parameters) do this? Only if there exists a solution for (θ_1, θ_2) , that is if the underlying assets have independent payoffs and hence (markets are complete and) the derivative can be generated by trading the underlying assets.

Problem 8.4.14 Idiosyncratic risk does not affect prices. Can you give an explanation in terms of no arbitrage? In other words, if the price of a stock

of a firm had a component representing the firm's idiosyncratic cash flow risk, which arbitrage opportunity would be left open in the financial markets?

Answer. Suppose the price of the idiosyncratic risk is negative; that is a firm with extra idiosyncratic risk is sold at a discount with respect to another firm with identical systematic risk component but with no idiosyncratic risk. Then, buy the firm with idiosyncratic risk, diversify this risk away by trading an infinite number of other firms (assets) with identical payoff but independent idiosyncratic risk (they exist by complete markets; otherwise the risk would not be idiosyncratic, as explained in class; by another no arbitrage argument these firms must all trade at the same price as our original firm and hence the diversification strategy has cost zero). This portfolio has generated the same payoff as the firm with no idiosyncratic risk. Selling now the more expensive firm without idiosyncratic risk gives you the arbitrage.

Problem 8.4.15 Derive the expression for the two-period ahead risk free rate in terms of marginal rates of substitution. (This is also known as the term structure of interest rates). You might want to know the law of iterated expectations: $E_t(E_{t+1}(x_{t+2})) = E_t(x_{t+2})$.

Answer. From the fundamental equation of asset prices:

$$(R_{t+2}^f)^{-1} = E_t \left(\beta \frac{u'(c_{t+2})}{u'(c_t)} \beta \frac{u'(c_{t+1})}{u'(c_{t+1})} \right)$$

which, using the Law of Iterated Expectations, becomes:

$$= E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} E_{t+1} \beta \frac{u'(c_{t+2})}{u'(c_{t+1})} \right)$$

In turn, using the fundamental equation of asset prices on the one period ahead risk free rate at time $t + 1$, $R_{t+1,t+2}^f$:

$$= E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} (R_{t+1,t+2}^f)^{-1} \right)$$

Finally, using the decomposition

$$E(xy) = ExEy + cov(x, y)$$

one gets the term structure equation:

$$(R_{t+2}^f)^{-1} = E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} \right) E_t \left(R_{t+1,t+2}^f \right)^{-1} + cov_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)}, (R_{t+1,t+2}^f)^{-1} \right)$$

Under which conditions would you expect the term structure to slope up; that is: $(R_{t+3}^f)^{-3} > (R_{t+2}^f)^{-2} > R_{t+1}^f$? (Explain the argument and discuss its plausibility.)

The simplest way is to assume that $E_t(R_{t+1,t+2}^f)^{-1}$ is roughly equal to $(R_{t+1}^f)^{-1}$. Then the sign of the covariance term drives the answer. In the phase of the cycle in which with high probability consumption is increasing at $t + 2$ if it has grown at $t + 1$, the covariance is positive and the term structure slopes down. Note though there is no correct answer here; only reasoned and unreasoned ones.

Problem 8.4.16 *Imagine that there are 3 states of the world: A, B and C which occur with probability p, q and (1 - p - q) respectively. What is the expected value of the following financial instruments? What is the variance of the expected payoff of these financial instruments? What is the covariance between the first financial instrument and each of the others? If the agent was risk neutral, how much would they pay for each of these financial instruments? i) A share that pays off 1 unit in state A. ii) A bond which pays off 1 unit in states A, B and C. iii) A share that pays off 1 unit in state A and C. iv) Insurance that pays off 1 unit in state B and -1 units in any other state. v) A share that pays off $\frac{1}{3p}$ in state A, $\frac{1}{3q}$ in state B and $\frac{1}{3(1-p-q)}$ in state C. For each of the above assets, write down the expected utility of holding such an asset. If the agent has a logarithmic utility function, what is the expected utility of each asset?*

Problem 8.4.17 *Prove (analytically, i.e., without graphs), that an agent with a concave utility function must be risk averse.*

Problem 8.4.18 *Consider an economy with two states of the world and two agents, A and B. In state 1, agent A gets 10 units of good and agent B gets 0 units. In state 2, A gets 0 units and B gets 15 units. Say that there exists a financial instruments which each agent can buy or sell which promises to pay the bearer 1 unit in state 1 at the cost of p units in state 2. Each agent*

has a utility function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. i) Work out how many units of the financial instrument each agent will demand as a function of p . (remember this can be positive or negative as the agents are allowed to buy and sell the asset). ii) Solve for the equilibrium value of p . (note that feasibility states that the total demand for the financial asset must be zero, as if one agent buys a unit of the asset, the other must sell it to him). iii) Work out the allocations of each agent in equilibrium. iv) Repeat steps 1-3, but this time assume that the utility function of agent B is $u(c) = c$.

Problem 8.4.19 Consider an agent with the following preferences:

$$u(c) = \frac{1}{1-\sigma} c^{1-\sigma}, \quad \sigma > 0$$

How does risk aversion depend on σ ? (Define first risk aversion, there are several definitions, make up one that makes sense to you.) Can σ be negative?

Problem 8.4.20 Consider an economy with 2 time periods, t and $t+1$, a representative agent with utility

$$\frac{1}{1-\sigma} (c_t)^{1-\sigma} + .9 \frac{1}{1-\sigma} E((c_{t+1})^{1-\sigma}), \quad \sigma > 0$$

and 3 states of the world with associated probabilities $\pi_1, \pi_2, 1 - \pi_1 - \pi_2$. What

1

is the expression of the price of an asset whose payoff is $x = 0$? (Hint: the

4

expression is in terms of c_t , c_{t+1} , and parameters π_1, π_2, σ).

Problem 8.4.21 Imagine that there are 3 states of the world: A, B and C which occur with probability p , q and $(1 - p - q)$ respectively. What is the expected value of the following financial instruments? What is the variance of the expected payoff of these financial instruments? What is the covariance between the first financial instrument and each of the others? If the agent was risk neutral, how much would he/she pay for each of these financial instruments? i) A share that pays off 1 unit in state A. ii) A bond which pays off 1 unit in states A, B and C. iii) A share that pays off 1 unit in state A and C. iv) Insurance that pays off 1 unit in state B and -1 units in any

other state. v) A share that pays off $\frac{1}{3p}$ in state A, $\frac{1}{3q}$ in state B and $\frac{1}{3(1-p-q)}$ in state C. For each of the above assets, write down the expected utility of holding such an asset. If the agent has a logarithmic utility function, what is the expected utility of each asset.

Problem 8.4.22 Consider a risk averse agent with income W . He faces a risk of incurring a cost D_1 with probability .3 and of D_2 with probability .4. He/she can buy insurance at price q , that is he/she can buy at cost q a contract that pays 1 dollar if he/she either cost D_1 or D_2 are realized? (Consider $D_1 > D_2$). Will the agent be able to perfectly insure? (Need to define perfect insurance; take a stab at this). How many units of the contract will the agent buy if the price is $q = .7$? And will the agent be able to perfectly insure in this case (according to your definition of perfect insurance)?

Chapter 9

Incentives



George Akerlof

Joseph Stiglitz

Michael Spence

Economists are all about incentives. Facing any inefficiency, economists look for misaligned incentives as culprits.

The financial markets crisis? Look at incentive compensation of managers and at Fannie and Freddie's political incentives.

The poor educational system? Look at the incentives implicit in the contract of unionized teachers.

The excessive cost of health insurance? Look at how doctors are compensated for services and at the political incentives which govern Medicare and Medicaid (and at the market power of insurance companies, but this does not have to do with incentives as much as with competition - another obsession of economists).

To discuss incentives properly, we need to introduce *private information* (or, sinonimously, *asymmetric information*)

9.1 Principal-Agent Problems

Consider an economy with $s = 1, \dots, S$ states of uncertainty, and a project (e.g., but not necessarily, a firm) with cash flow $R_s > 0$ in state s . Without loss of generality rank R_s so that $R_1 < R_2 < \dots < R_S$.

We consider two distinct (but related) cases.

Making the agent work. A principal owns the project ex-ante, but an agent necessarily manages it, by spending effort a . By increasing the effort he/she employs into managing the project, effort a , the agent increases the probability that the the project is successful, that is that it has a high cash flow R_s . Let $p_s(a)$ denote the probability that state s occurs (and hence that the firm's cash flow is R_s) when the agent's effort is a . We assume that the expected cash flow $\sum_{s=1}^S p_s(a)R_s$ is increasing in a .¹ The agent's utility function is increasing in consumption c and is decreasing in effort a :

$$u(c) - \psi(a)$$

The principal's utility function is increasing in consumption:

$$v(c).$$

Paying the agent according to his productivity. A principal owns the project ex-ante, but an agent necessarily manages it. The principal hires a random agent, but agents are different in their abilities. Better agents have a higher probability that the the project they manage is successful, that is that it has a high cash flow R_s . Let $p_s(q)$ denote the probability that state s occurs (and hence that the firm's cash flow is R_s) when the agent's quality is $q = 1, \dots, Q$. We order quality q so that the expected cash flow $\sum_{s=1}^S p_s(q)R_s$ is increasing in q .² The agent's

¹Formally, we assume that

$$\text{for any effort } a < b, \text{ the ratio } \frac{p_s(b)}{p_s(a)} \text{ is increasing in } s.$$

This assumption is reasonable, but it is restrictive and far from innocuous. It is called *single-crossing-property* in the literature.

²Once again this requires that the *single-crossing-property* be satisfied.

utility functions are increasing in consumption c

$$u(c),$$

and so is the principal's:

$$v(c).$$

9.1.1 The symmetric information benchmarks

Suppose that the principal observes everything about the agent.

Making the agent work

In this economy the principal observes and measures, in particular, the managerial effort of the agent, a . In other words, the principal can dictate how much the agent works to insure the success of the project.

We can now write the Pareto planning problem for this economy (recall we defined planning problem in the chapter on Equilibrium). The solution of the Pareto planning problem is an optimal (Pareto efficient) contract between the principal and the agent. It is convenient to write the Pareto planning problem in the form of the utility maximization for the principal under the constraint that the agent derives a minimal given parametric utility level, \bar{u} .³

The principal must then choose, before the uncertainty is realized, the best wage schedule, w_s , to pay the agent. Also, the principal chooses which managerial effort a to dictate to the agent. The principal's choices maximize his/her utility, and are such that the agent will work for him/her; that is:

$$\max_{a, \{w_s\}_{s=1}^S} \sum_{s=1}^S p_s(a) v(R_s - w_s) \quad (9.1)$$

$$\text{s.t. } \sum_{s=1}^S p_s(a) u(w_s) - \psi(a) \geq \bar{u} \quad (9.2)$$

The first order conditions of problem (9.1-??) include:

$$\frac{v'(R_s - w_s)}{u'(w_s)} = \lambda \quad (9.3)$$

³The utility level \bar{u} can also be interpreted as the utility the agent obtains if he does not work for the principal (and does something else in the market).

where λ is the Lagrange multiplier associated to (??).

Consider the important case in which the principal is risk neutral, that is, without loss of generality,

$$v(c) = c.$$

In this case, (9.3) implies that

$$w_s = w,$$

that is, the wage is constant, and the principal fully insures the agent. It's natural that it'd be so, as the principal is risk neutral and the agent risk averse.⁴

Paying the agent according to his productivity

In this economy the principal observes and measures, in particular, the managerial skills of the agent, his ability $q = 1, \dots, Q$. In other words, the principal can pay the agents she hires according to his productivity, his ability to insure the success of the project. Assume that a fraction ξ_q of the agents are born with ability q , so that $\sum_{q=1}^Q \xi_q = 1$. The principal, therefore hires an agent with ability q , randomly, with probability ξ_q .

Let's now write the Pareto planning (optimal) problem. It is convenient to write the problem in the form of a weighted maximization of the utility of the agents of different ability types q , with weights γ_q such that $\sum_{q=1}^Q \gamma_q = 1$, under the constraint that the principal derives a minimal given parametric utility level, \bar{v} . (Why is this convenient? Try and write the dual case in which we maximize the utility of the principal under the constraint on the minimum utility of the agents.) The planner chooses wages (contracts) (w_1^q, w_2^q) , respectively, for agents of type q :

$$\max_{(w_s^q)_{q=1, s=1}^{Q, S}} \sum_{q=1}^Q \gamma_q \sum_{s=1}^S p_s(q) u(w_s^q) \quad (9.4)$$

subject to

$$\sum_{q=1}^Q \xi_q \sum_{s=1}^S p_s(q) v(R_s - w_s^q) \geq \bar{v}.$$

⁴Suppose instead that the agent is risk neutral and the principal risk averse. In this case (9.3) implies that $R_s - w_s = r$. That is, the return to the principal is constant, the agent holds all the risk and insures the principal. It is as if the principal would sell the project to the agent.

For simplicity let's only consider the case in which the principal is risk neutral. In this case, without loss of generality, $v(c) = c$ and the planner chooses $(w_s^q)_{q=1, s=1}^{Q, S}$ which satisfy

$$w_s^q = w^q, \text{ for any } q.$$

Once again, the risk neutral principal is fully insuring the agent, no matter what the type of the agent is. But the agent's wage will in general depend on his productivity/ability/skills q .

9.1.2 Moral hazard

Consider now the first economy, where the incentive problem is to make the agent work. Suppose however now that the effort a is not observed, nor can be monitored, by the principal (we say this economy is subject to *moral hazard*, as the agent can hide his/her effort to the principal). The agent, given a schedule $\{w_s\}_{s=1, \dots, S}$ chosen for her/him by the principal, chooses a freely.

Are the solutions of the principal-agent problem we determined in the previous section still optimal? Can the optimal contract we obtained still be achieved if the effort a is not observed, nor can be monitored, by the principal? The answer is obviously not, in general. Consider a few examples of what can go wrong.

In an economy with moral hazard, that is, in an economy where the effort choice a of the agent is not observable, the social planner cannot choose any possible level of effort a on the part of the agent. Instead, the social planner must choose the wage $\{w_s\}_{s=1}^S$ so as to incentivate the agent to choose an optimal effort a .

Formally, the social planning problem needs to satisfy an additional constraint, called the *incentive compatibility constraint* restricting his effort choice to the ones the agent himself would choose under the wage optimal contract:

$$a \in \arg \max \sum_{s=1}^S p_s(a) u(w_s) - \psi(a). \quad (9.5)$$

Consider the case in which the principal is risk neutral. With symmetric information, we have shown that we have $w_s = w$, and the principal fully insures the agent. But with private information, in the moral hazard case, if

the agent is offered a contract such that $w_s = w$, he would choose the minimal effort $a^* = \arg \min \psi(a)$. This in general is not the optimal choice of the principal (the principal will want to maximize a , in fact, under the constraint that the agent will accept working for him). A different wage contract need to be chosen by the social planner.

Rather than attempting a general solution (which is hard), consider a simple the economy in which $s = 1, 2$, and $a = L, H$. Suppose $R_2 > R_1$. This simply means that state $s = 2$ is the good state. Naturally, $p_2(H) = 1 - p_1(H) > p_2(L) = 1 - p_1(L)$, that is, $p_1(H) < p_1(L)$; and furthermore, $\psi(H) > \psi(L)$. This simply means that H is the high effort, which increases the probability of the good state and has a higher utility cost.

Assume the principal is risk neutral. Rather than studying (9.1-??), we study the equivalent problem in which the agent maximizes his/her utility guaranteeing a minimal utility to the principal (yes, just because we can!). By varying the minimal utility to the principal, we obtain the whole set of optimal contracts. We choose to characterize the one in which the minimal utility for the principal is 0. This is because, being the principal risk averse, his/her utility effectively coincides with his/her profits, and in an economy with perfect competition profits will in fact be 0 in equilibrium. If e.g., we interpret the principal as an insurance company (resp. a firm), we are studying then a perfectly competitive insurance market (resp. a perfectly competitive labor market for managerial services). S

If it is optimal for the principal to induce the effort $a = H$ (the only interesting case), under asymmetric information, the social planning (optimal contracting) problem is:

$$\max_{w_1, w_2} p_1(H)u(w_1) + (1 - p_1(H))u(w_2) - \psi(H)$$

subject to

$$p_1(H)u(w_1) + (1 - p_1(H))u(w_2) - \psi(H) \geq \begin{pmatrix} p_1(L)u(w_1) + \\ (1 - p_1(L))u(w_2) - \psi(L) \end{pmatrix} \quad (9.6)$$

$$p_1(H)(R_1 - w_1) + (1 - p_1(H))(R_2 - w_2) \geq 0 \quad (9.7)$$

Notice that (9.6), the *incentive compatibility constraint*, must be binding for $w_1 = w_2$, as in this case the agent prefers $a = L$. But since $w_1 = w_2$ is the solution when the incentive constraint is not imposed (see before), it

follows that (9.6) is in fact binding at the solution. We conclude that w_1, w_2 are chosen so that the agent is indifferent between $a = H$ and $a = L$. The constraint (9.7) is also binding (this is a bit harder to see; if it wasn't then you could transfer resources to the agent, and this could be done without significant effects on (9.6)).

We characterize now the optimal contract (w_1^*, w_2^*) which solves the social planning (optimal contracting) problem under the assumption that $u(c) = \ln(c)$. The wage schedule (contract) (w_1, w_2) is chosen to maximize the utility of the agent under

$$p_1(H)w_1 + (1 - p_1(H))w_2 = p_1(H)R_1 + (1 - p_1(H))R_2,$$

and the incentive constraint (9.7). Note that $p_1(H)w_1 + (1 - p_1(H))w_2 = p_1(H)R_1 + (1 - p_1(H))R_2$ can be interpreted as a budget constraint (R_1 and R_2 are the endowments, ..., what are the prices?). Furthermore, under log preferences, the incentive constraint (9.7), holding with equality, can be written as (do it yourselves; it requires some algebra combined with knowledge of the properties of logs and exponential):

$$\frac{w_2}{w_1} = k$$

for some constant k . If you've done it yourselves, as I asked you to, you have solved for k ; now check that $k > 1$.

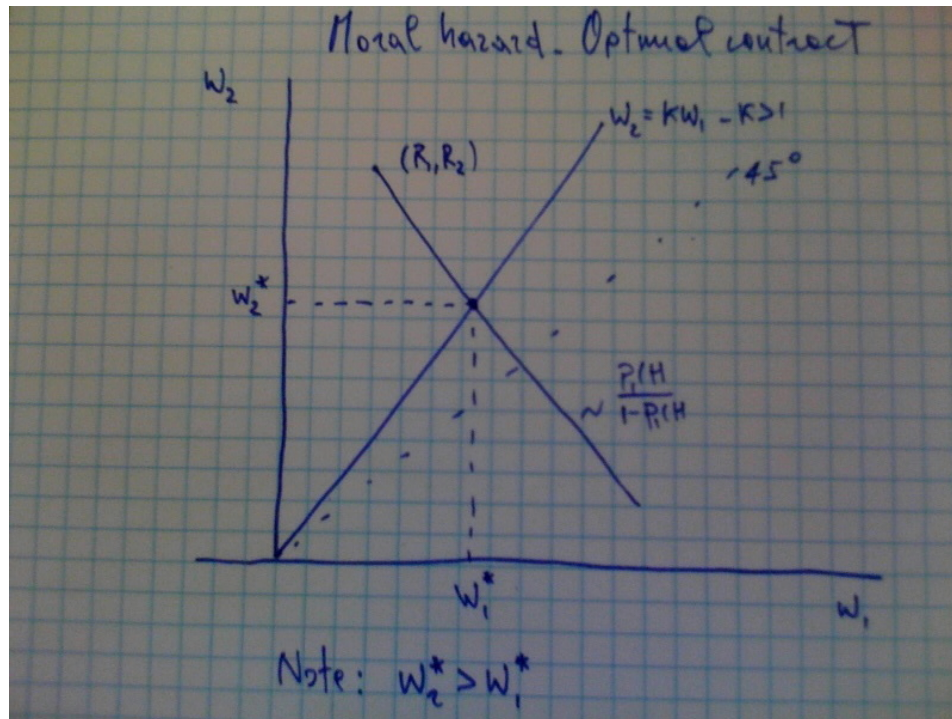
Therefore, the solution of the social planning (optimal contracting) corresponds to the (w_1^*, w_2^*) which solves (10.1-??). It is easily checked that the solution is unique. Since the solution satisfies $\frac{w_2}{w_1} = k$, with $k > 1$, it has the property that $w_2^* > w_1^*$. But w_1^* is the payment to the agent in the low cash flow state, $s = 1$, and $w_1^* - w_2^*$ is a measure of the distance from full insurance in the contract ($w_1 = w_2$ corresponds to full insurance).

You can now draw a picture (I admit in this case a picture is worth a million algebraic equations). Do this in steps:

1. Put w_2 on the vertical axis and w_1 on the horizontal axis.
2. Draw the line corresponding to (budget) constraint (10.1), passing through the (endowment) point (R_1, R_2) .
3. Draw the 45° line through the origin which corresponds to the locus of points with full insurance (that is, such that $w_1 = w_2$; obvious, right?).

4. Draw the incentive constraint (??).
5. Find the unique point (w_1^*, w_2^*) which satisfies (10.1-??).

The result should be something like this:



Moral hazard example

The moral hazard we have just studied is rather abstract and has very many interpretations and applications in economics. It is worth listing a few.

Effort and incentive compensation. Managers might shirk if their compensation is not linked to the success of the project they manage; incentive compensations in the form of wages which depend on the realized cash-flow R_s , that is, on s , are necessary to at least partly avoid this. Can you see this is the rationale for the equity (or stock) compensation of managers?

Risk and insurance. The case in which the principal is risk neutral also captures insurance markets. The principal is the insurer and the agent is the

insuree. The effort a is interpreted as any measure to limit the probability of an accident, e.g., installing fire alarm devices (or at least avoiding to get rid of their batteries).

An analogous moral to the one obtained in the previous case can be extracted: If the agent is fully insured he/she will not invest any effort in limiting the probability of the accident he/she is insured against. Insurance contracts therefore need to steer away from full insurance to avoid this. Can you see this is the rationale for co-payments in insurance contracts?⁵

9.1.3 Adverse selection

Consider now the economy where the incentive problem is to pay the agent according to his productivity. We restrict ourselves to a simple example not unlike the one we studied in the moral hazard economy, in which $s = 1, 2$, and $q = L, H$. As before, suppose $R_2 > R_1$, and $p_2(H) = 1 - p_1(H) > p_2(L) = 1 - p_1(L)$, that is, $p_1(H) < p_1(L)$. Type H is then the high quality type, which increases the probability of the good state. Let ξ_H be the fraction of the agents with $q = H$ ($1 - \xi_H$ with $q = L$).

Let's now write the Pareto planning (optimal) problem in the form of a weighted maximization of the utilities of the two types ($q = H, L$) of agents (see the chapter on Equilibrium once again), where the weights are $(\gamma_H, 1 - \gamma_H)$, not necessarily the population weights of the two types. The planner chooses wages (contracts) (w_1^H, w_2^H) , (w_1^L, w_2^L) , respectively, for agents of type $q = H$ and $q = L$:

$$\begin{aligned} \max_{w_1^H, w_2^H, w_1^L, w_2^L} \quad & \gamma_H [p_1(H)u(w_1^H) + (1 - p_1(H))u(w_2^H)] + \\ & (1 - \gamma_H) [p_1(L)u(w_1^L) + (1 - p_1(L))u(w_2^L)] \end{aligned}$$

⁵A different interesting example can be constructed for the case (which we did not study, but we could have) in which it is the agent to be risk neutral. With symmetric information (and limited liability) the principal payoff is $R_s - w_s = \min\{R_s, r\}$. Suppose that the effort a is now interpreted as a form of managing effort with the effect of reducing the risk of the firm (for a constant mean). If the agent is risk neutral, it is easy to construct examples in which the agent will lose in term of expected utility by reducing the risk of the firm. The agent hence will not do so, charging the principal with excessive risk. Can you see this implies that the manager of a leveraged firm might be induced to take up too much risky projects?

subject to

$$p_1(H)u(w_1^H) + (1 - p_1(H))u(w_2^H) \geq \begin{pmatrix} p_1(H)u(w_1^L) + \\ + (1 - p_1(H))u(w_2^L) \end{pmatrix} \quad (9.8)$$

$$p_1(L)u(w_1^L) + (1 - p_1(L))u(w_2^L) \geq \begin{pmatrix} p_1(L)u(w_1^H) + \\ + (1 - p_1(L))u(w_2^H) \end{pmatrix} \quad (9.9)$$

$$\begin{aligned} & \xi_H [p_1(H)(R_1 - w_1^H) + (1 - p_1(H))(R_2 - w_2^H)] + \\ & + (1 - \xi_H) [p_1(L)(R_1 - w_1^L) + (1 - p_1(L))(R_2 - w_2^L)] \geq 0. \end{aligned} \quad (9.10)$$

Constraint (9.8) are the incentive compatibility constraints. They require that, given the wages (contracts) (w_1^H, w_2^H) , (w_1^L, w_2^L) , selected by the planner, respectively, for agents of type $q = H$ and $q = L$:

- i) agents of type $q = H$ prefer (w_1^H, w_2^H) to (w_1^L, w_2^L) and
- ii) agents of type $q = L$ prefer (w_1^L, w_2^L) to (w_1^H, w_2^H) .

These constraints are necessary because the principal, while offering different contracts to the two different types of agents, in fact does not observe which type is which. The optimal contract therefore must provide the agents with the incentive to report their true type.

Constraint (9.10) is instead the condition that the principal's profits are weakly greater than 0. (Note the principal now obtains profits, in principle, from the two different types of contracts).

Two kind of optimal contracts arise in this economy, a pooling contract or a separating contract (this is hard to prove from the statement of the Pareto optimal problem; I encourage you to try, and then, if necessary to believe me). We examine them in turn.

Pooling contract

In this case, the optimal contracts offered to the two different types of agents are the same:

$$(w_1^L, w_2^L) = (w_1^H, w_2^H) = (w_1, w_2)$$

As a consequence, the incentive compatibility constraints (9.8) are trivially satisfied. The Pareto planning (optimal contracting) problem, restricted to pooling contracts, reduces to:

$$\max_{w_1, w_2} \theta_1 u(w_1) + \theta_2 u(w_2) \quad (9.11)$$

subject to

$$\theta_1 (R_1 - w_1) + \theta_2 (R_2 - w_2) \geq 0 \quad (9.12)$$

where

$$\begin{aligned} \theta_1 &= [\xi_H p_1(H) + (1 - \xi_H) p_1(L)] \text{ and} \\ \theta_2 &= [\xi_H (1 - p_1(H)) + (1 - \xi_H) (1 - p_1(L))] . \end{aligned}$$

In fact (9.12) can be written as a budget constraint:

$$\theta_1 w_1 + \theta_2 w_2 \geq \theta_1 R_1 + \theta_2 R_2$$

and at the solution,

$$w_1 = w_2.$$

Both types of agents are perfectly insured. Note that each L type receives as much as each H type. But the H types are "more productive" (have higher skills, $q = H$) in this economy. A planner with a relatively large weight γ_H will not want to choose this allocation. In this case, the solution will be to select a separating contract.

Separating contract

In this case, the optimal contracts offered to the two different types of agents are not the same:

$$(w_1^L, w_2^L) \neq (w^H, w_2^H).$$

The characterization of these contracts is not easy, and we'll provide it here without proof:

1. The optimal contract for agents of type $q = L$, (w_1^L, w_2^L) , satisfies

$$w_1^L = w_2^L = w^L = \frac{1}{2} (p_1(L) R_1 + (1 - p_1(L)) R_2)$$

so that they are fully insured at *fair* odds: $\frac{p_1(L)}{1-p_1(L)}$.

2. The optimal contract for agents of type $q = H$, (w_1^H, w_2^H) , satisfies

$$\max_{w_1^H, w_2^H} p_1(H)u(w_1^H) + (1 - p_1(H))u(w_2^H) \quad (9.13)$$

subject to

$$p_1(L)u(w_1^L) + (1 - p_1(L))u(w_2^L) = \left(\begin{array}{c} p_1(L)u(w_1^H) + \\ + (1 - p_1(L))u(w_2^H) \end{array} \right) \quad (9.14)$$

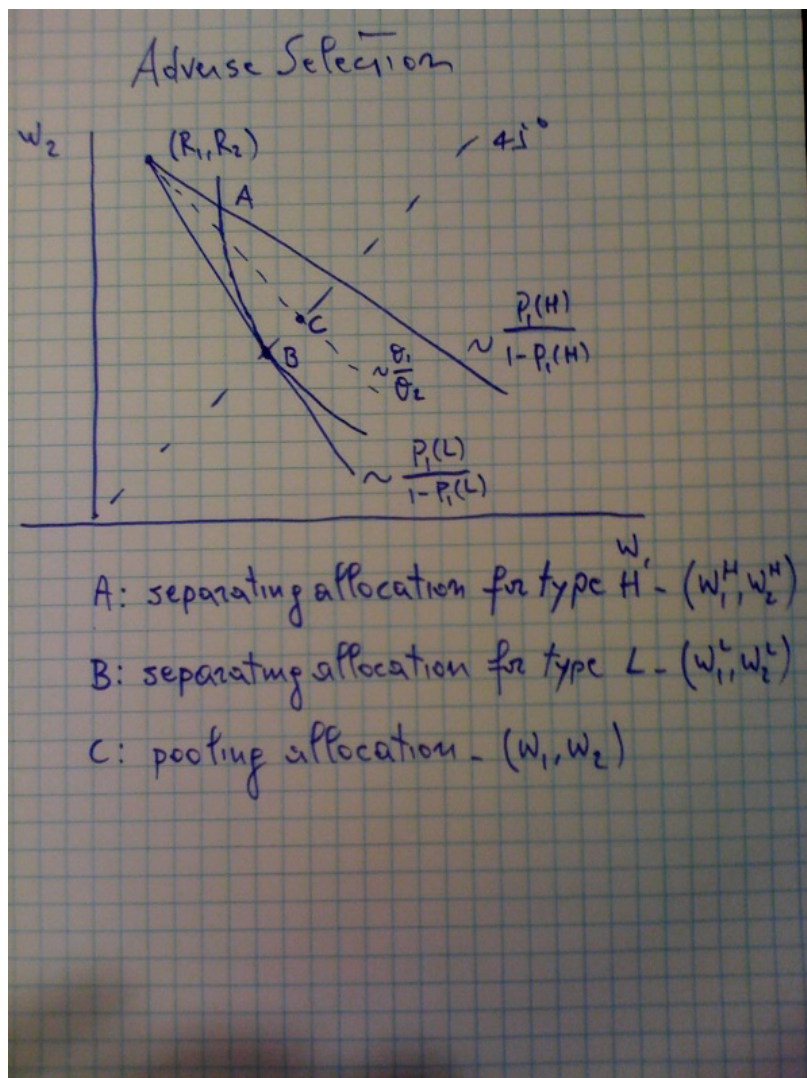
$$p_1(H)(R_1 - w_1^H) + (1 - p_1(H))(R_2 - w_2^H) = 0 \quad (9.15)$$

Agents of type H are not fully insured:

$$w_2^H > w_1^H.$$

They also face *fair* insurance odds, $\frac{p_1(H)}{1-p_1(H)}$. At this price the planner would like to insure them fully. But in this case the planner could not separate the L types, who would claim to be H types to obtain full insurance at better odds: $\frac{p_1(H)}{1-p_1(H)} < \frac{p_1(L)}{1-p_1(L)}$. The planner therefore chooses the best contract for agents of type H which agents of type L do not have incentive to acquire (this is exactly the meaning, in words, of requiring that the planner will be restricted by (??) when maximizing utility 9.13).

Here's the picture worth a million words (and equations).



Adverse selection: Pooling and separating contracts

9.1.4 Social security

Here's another example where asymmetric information plays a role. Consider an economy in which agents derive utility from consumption $u(c)$ and disutility from labor $v(l)$. We assume $u(c)$ is strictly increasing and strictly

concave and $v(l)$ is strictly increasing and strictly convex (so that $-v(l)$ is strictly concave). Agents have stochastic, idiosyncratic, and permanent labor productivity $\theta(s)$, where $s \in S$ and s has probability $\pi(s)$. An agent with productivity $\theta(s)$ can generate income $y = \theta(s)l$. The environment is similar to the one considered by Mirrlees (1971). Individual income y is observable, but neither $\theta(s)$ nor l are. Here's the asymmetric information: a relatively small amount y produced could be due to low productivity or low labor effort.

Symmetric information \rightarrow "Communism"

If the planner observed each agent's productivity and labor, that is, if the planner faced a full information problem, he would choose an allocation which solved:

$$\max_{\{c(s), l(s)\}_{s \in S}} \sum_{s \in S} \pi(s) (u(c(s)) - v(l(s))) \quad (9.16)$$

subject to

$$\sum_{s \in S} \pi(s) \theta(s) l_t(s) \geq \sum_{s \in S} \pi(s) c_t(s). \quad (9.17)$$

At a solution, $c(s) = c(s') \equiv c$, and $v'(l(s))/v'(l(s')) = \theta(s)/\theta(s')$, $\forall s, s' \in S$. Prove this!

All agents consume the same amount and the more productive agents work more. This is the "*communist*" solution where each agent receives consumption according to his needs and contributes labor according to his abilities:

*In a higher phase of communist society, after the enslaving subordination of the individual to the division of labor, and therewith also the antithesis between mental and physical labor, has vanished; after labor has become not only a means of life but life's prime want; after the productive forces have also increased with the all-around development of the individual, and all the springs of co-operative wealth flow more abundantly—only then can the narrow horizon of bourgeois right be crossed in its entirety and society inscribe on its banners: **From each according to his ability, to each according to his needs!** Karl Marx, 1875, Critique of the Gotha Program.*

The phrase *From each according to his ability, to each according to his needs!* summarizes the principles that, under a communist system, every person should contribute to society to the best of his ability and consume from society in proportion to his needs, regardless of how much he has contributed.⁶

Asymmetric information → compensate productivity

Suppose now that the planner observes the production y of each agent, but not $\theta(s)$ nor l . How is the planning problem changed? Which other constraint needs to be added? Can you write it?

⁶The following is taken from Wikipedia on ??

Although Marx is popularly thought of as the originator of the phrase, the slogan was common to the socialist movement and was first used by Louis Blanc in 1840, in "The organization of work", as a revision of a quote by the utopian socialist Henri de Saint Simon, who claimed that each should be rewarded according to how much he works. The origin of this phrasing has also been attributed to the French communist Morelly, who in his *Code of Nature: Sacred and Fundamental Laws that would tear out the roots of vice and of all the evils of a society*, 1755, proposed:

I. Nothing in society will belong to anyone, either as a personal possession or as capital goods, except the things for which the person has immediate use, for either his needs, his pleasures, or his daily work.

II. Every citizen will be a public man, sustained by, supported by, and occupied at the public expense.

III. Every citizen will make his particular contribution to the activities of the community according to his capacity, his talent and his age; it is on this basis that his duties will be determined, in conformity with the distributive laws.

The phrase may also find an earlier origin in the New Testament. In Acts 4:32-35, the Apostles lifestyle is described as communal (without individual possession), and uses the phrase

distribution was made unto every man according as he had need.

Finally, the United Nations' *Universal Declaration of Human Rights*, Article 22, states:

Everyone, as a member of society, has the right to social security.

9.1.5 Other examples [to be added]

1. Akerlof's lemons
2. Spence's signaling in education
3. Myerson-Satterwaite theorems on gains from trade
4. Milgrom-Stokey's no-trade theorem.

9.2 Problems

Problem 9.2.1 *Consider a moral hazard economy with no asymmetric information, that is, the planner can observe and hence choose the agent's effort. Suppose however that the social planner is restricted by limited liability, that is, he has to guarantee that neither the agent nor the principal can receive negative amounts from the firm:*

$$R_s \geq w_s \geq 0$$

Characterize the optimal contract in this case, when either the principal or the agent is risk neutral.

Answer. Formally, the constraint

$$R_s \geq w_s \geq 0$$

need to be added to (9.1-??) in the social planning problem. The optimal contract requires, in this case, that:

If the principal is risk neutral, $w_s = \min\{R_s, w\}$ and $R_s - w_s = \max\{R_s - w, 0\}$. The principal receives the residual cash flow of the project, after having compensated the agent.

If the agent is risk neutral, $w_s = \max\{R_s - r, 0\}$ and $R_s - w_s = \min\{R_s, r\}$. It is the the agent who receives the residual.

Problem 9.2.2 *Consider a two period economy. Agents are all identical, that is, there is one representative agent. The representative agent is alive at time t and $t + 1$, and has preferences:*

$$\ln x_t + \beta \ln x_{t+1}, \quad \beta < 1.$$

This agent is endowed with w_t units of the consumption good at time t and w_{t+1} at time $t + 1$. Write down the consumer maximization problem of the representative agent, under the assumption that he/she can trade a bond (borrow or lend) at an equilibrium rate r . [It's a real interest rate - no inflation]. Write down the market clearing conditions (that is, feasibility conditions) for the whole economy. [Yes, there is only a representative agent and the portfolio positions of the bond must clear; it's a trivial economy]. Write down the Pareto planning problem for this economy. [Once again, there is only a representative agent; it's a trivial economy]. Solve for the Pareto optimal allocation. Solve for the equilibrium interest rate and for the representative agent equilibrium allocation. [Help yourself with the Pareto optimal allocations; the competitive equilibrium allocations are Pareto optimal in this economy, right?]. Suppose now that the bond market is closed (does not exist). But the agent can, at time t , put goods in a bucket which he/she find then at time $t + 1$. Any unit put in the bucket at time t depreciates to $1 - \delta$ at time $t + 1$. Assume that the agent has no endowment at time $t + 1$, $w_{t+1} = 0$, and solve for the Pareto optimal allocations [The solution will be in terms of w_t , β , and δ]. Suppose now that the Government taxes the goods in the bucket, at time $t + 1$, at the rate τ . Will agent now put more or less goods in the bucket at time t ? The utility of the Government is linear in tax revenues. Will it tax the representative agent maximally? That is, will the Government choose $\tau = 1$? Suppose now that the Government cannot look in the bucket. He still wants to tax the goods in the bucket, but just does not know how many there are. Write down the Pareto optimal allocation problem for this economy [choose the form in which the utility of the Government is maximized for a given level of utility U of the representative agent; be careful to write down the incentive compatibility constraint]. Do not try and solve the problem formally; but, given what you understood about this economy, can you guess what the Pareto optimal tax rate will be in this case?

Chapter 10

Fiscal policy

What are the effects of fiscal policies?

10.1 Ricardian Equivalence - a.k.a. Arithmetics



Robert Barro



Tom Sargent



Neil Wallace

Consider a two-period economy with a representative agent. A production (or a financial) technology allows the agent to save any non-negative amount S_t at t and receive $(1 + r)S_t$ at $t + 1$. The representative agent's budget constraints are:

$$\begin{aligned}c_t + S_t &= w_t \\c_{t+1} &= S_t(1+r) + w_{t+1}\end{aligned}$$

We have noted already that these budget constraints (solving and substituting for S_t) can be reduced to

$$c_t + \frac{1}{1+r}c_{t+1} = w_t + \frac{1}{1+r}w_{t+1}$$

We have also computed a solution for the consumption-saving problem of the representative agent with preferences $u(c_t) + \beta u(c_{t+1})$. If we assume that:

$$u(c_t) = \ln c_t, \quad u(c_{t+1}) = \ln c_{t+1}, \quad \beta = \frac{1}{1+r},$$

we have the following consumption allocations:

$$\begin{aligned}c_t &= \frac{1+r}{2+r} \left(w_t + \frac{1}{1+r}w_{t+1} \right) \\c_{t+1} &= \frac{1+r}{2+r} \left(w_t + \frac{1}{1+r}w_{t+1} \right)\end{aligned}$$

Consider now a government which can tax/subsidize agents at either moment of time. Let a subsidy to the representative agent at time t be g_t , and a tax be τ_t . Similarly, subsidies and taxes at time $t+1$ are, respectively, g_{t+1} and τ_{t+1} . Note that taxes are assumed to be *lump-sum*, that is, not indexed to income, savings, consumption, or any other of the agent's choices.¹ The government borrows at the interest rate r and balances budgets dynamically (not necessarily period by period). Therefore, the government dynamic budget constraint is:

$$g_t + \frac{1}{1+r}g_{t+1} = \tau_t + \frac{1}{1+r}\tau_{t+1}$$

¹In other words taxes are the same in absolute amount independently on the agent's choices. In an economy with heterogeneous agents, rather than a representative agent, lump-sum taxes are the same in units of consumption goods for each agent. See the next section for a relaxation of the *lump-sum* assumption.

10.1 RICARDIAN EQUIVALENCE - A.K.A. ARITHMETICS¹⁷⁷

With government intervention, the representative agent's budget constraints are:

$$\begin{aligned}c_t + S_t &= w_t + g_t - \tau_t \\c_{t+1} &= S_t(1 + r) + w_{t+1} + g_{t+1} - \tau_{t+1}\end{aligned}$$

which can be reduced to

$$c_t + \frac{1}{1+r}c_{t+1} = w_t + \frac{1}{1+r}w_{t+1} + \left(g_t + \frac{1}{1+r}g_{t+1} - \tau_t - \frac{1}{1+r}\tau_{t+1}\right)$$

The terms in parenthesis on the right-hand-side is $= 0$ because of the government budget constraint.

What does this mean? That in this economy,

government expenditures have no effect - because the representative agent rationally anticipates that every expenditure at t is financed through taxes at $t + 1$.

This proposition, forgotten by keynesian economists, has been returned to light by Robert Barro (who called it *Ricardian Equivalence* - from David Ricardo, who apparently first stated it informally) and by Tom Sargent and Neil Wallace (who properly called it *arithmetic*).

10.1.1 Borrowing-Constraints

The analysis in the previous section assumes that the representative agent does not face any borrowing constraints.² Suppose on the contrary that the agent can lend at rate r , but cannot borrow. That is,

$$S_t \geq 0,$$

or, equivalently:

$$c_t \leq w_t + g_t - \tau_t$$

In this case, obviously, positive government expenditures net of taxes at time t , $g_t - \tau_t > 0$, have an effect on consumption at time t , as long as the borrowing constraint is binding, that is, $c_t = w_t + g_t - \tau_t$.

²Similar (but less clean) results are obtained if we assume that the representative agent can borrow but at a rate $r' > r$.

We conclude that in this case government expenditure has efficiency gains, it allows for consumption smoothing which would be impossible without it. In other words (can you show this formally?), government expenditure can support the optimal consumption allocation

$$\begin{aligned} c_t &= \frac{1+r}{2+r} \left(w_t + \frac{1}{1+r} w_{t+1} \right) \\ c_{t+1} &= \frac{1+r}{2+r} \left(w_t + \frac{1}{1+r} w_{t+1} \right) \end{aligned}$$

even when such allocation would not satisfy the borrowing constraint with no government ($g_t = \tau_t = 0$), that is, when

$$w_t < \frac{1+r}{2+r} \left(w_t + \frac{1}{1+r} w_{t+1} \right)$$

or (better written), when:

$$w_t < w_{t+1}$$

Before concluding that, because agents are in fact borrowing constrained especially during recessions, government expenditure is a desirable policy, you should ask yourself why is it that the representative agent is constrained and the government is not. Why don't private banks lend to the agent? Well, the government has an advantage with respect to banks, in principle, in making sure agents re-pay debts (the government can tax and imprison those who do not pay taxes; the bank cannot).³ But at the same time the government has a disadvantage with respect to banks, normally, in screening and servicing debtors.

10.1.2 Distortionary Taxes

Suppose taxes are not lump-sum but *distortionary*. Taxes are distortionary when they are given as a function of an endogenous variable, a variable chosen by the agent in the economy.

³In fact, even this advantage is debatable. See the notes on Time Inconsistency of Government Policies.

10.1 RICARDIAN EQUIVALENCE - A.K.A. ARITHMETICS 179

Typically, income taxes are distortionary, because income depends on labor supply which is chosen by the agent. In our economy income is exogenous, it takes the form of endowments, and hence distortionary taxes are e.g., taxes on savings:

$$\tau_{t+1} = \tau(1+r)S_t$$

The government budget constraint is

$$\tau_t + \frac{1}{1+r}\tau_{t+1} = g_t + \frac{1}{1+r}g_{t+1}$$

Suppose also, without loss of generality, that $\tau_t = 0$, we have

$$\tau_{t+1} = (1+r)g_t + g_{t+1}.$$

The representative agent's budget constraints are

$$\begin{aligned} c_t + S_t &= w_t + g_t \\ c_{t+1} &= S_t(1+r)(1-\tau) + w_{t+1} + g_{t+1} \end{aligned}$$

Under the maintained assumptions that:

$$u(c_t) = \ln c_t, \quad u(c_{t+1}) = \ln c_{t+1}, \quad \beta = \frac{1}{1+r},$$

we have the following consumption allocations (*Make sure you derive these yourself!*):⁴

$$\begin{aligned} c_t &= \frac{1+r}{2+r-\tau} \left(w_t + \frac{1}{1+r}w_{t+1} \right) \\ c_{t+1} &= \frac{(1+r)(1-\tau)}{2+r-\tau} \left(w_t + \frac{1}{1+r}w_{t+1} \right) \end{aligned}$$

Taxes are derived from the intertemporal government budget constraint:

⁴In fact, be careful. If the agent is borrowing, $S_t < 0$, he receives a subsidy from the government, the way we have written the agent's maximization problem (the interest rate he faces is lower than $1+r$, it's $(1+r)(1-\tau)$ in fact. This is not typically the case in the fiscal code. The equations we obtain in the following only hold for $S \geq 0$. Suppose that the interest rate on borrowing is $1+r$ and derive the consumption equations in this case. How does the whole consumption function look like as a function of permanent income? Draw a figure.

$$\tau_{t+1} = \tau(1+r)S_t = (1+r)g_t + g_{t+1}$$

and hence

$$\begin{aligned}\tau &= \frac{g_t + \frac{1}{1+r}g_{t+1}}{S_t} \\ &= \frac{g_t + \frac{1}{1+r}g_{t+1}}{w_t + g_t - c_t} = \frac{g_t + \frac{1}{1+r}g_{t+1}}{w_t + g_t - \frac{1+r}{2+r-\tau} \left(w_t + \frac{1}{1+r}w_{t+1} \right)}\end{aligned}$$

We conclude that:

- $c_t > c_{t+1}$ (*Show this!*)
- τ increases with g_t and g_{t+1} (*Show this as well! Be careful, it requires some messy derivation and the use of the Implicit Function Theorem*)

Let the distortion due to taxes be measured by

$$\frac{c_t - c_{t+1}}{c_t}$$

Finally, and most importantly, we conclude from the analysis above that the distortion increases with g_t and g_{t+1} (*Try this! But don't be surprised if you can't do it*).

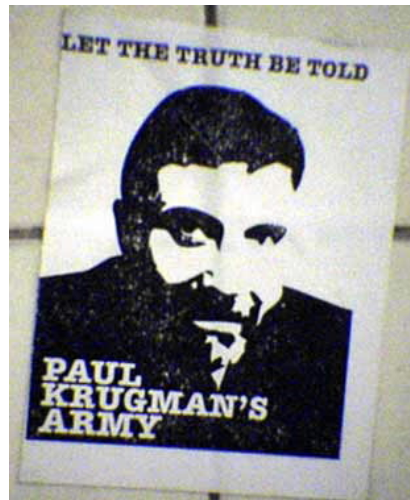
Finally, let's derive the Laffer curve, one of the main bullets of "small government" types:

An decrease in the tax rate τ could give rise to an increase in the government tax revenues, $\tau_{t+1} = \tau(1+r)S_t$.

How is this possible? Easy to prove: take $\tau = 1$ (all savings go to the government); show that in this case $S_t = 0$ and hence $\tau_{t+1} = 0$. Notice that the proof works at very high taxes; supply-side economics relies on an empirical question: at which tax rate τ^L does the Laffer curve effect kick in? Are tax rates now in the U.S. in the range of τ^L ? Answer: not even close.

10.1 RICARDIAN EQUIVALENCE - A.K.A. ARITHMETICS¹⁸¹

10.1.3 Krugman,, and the army of fiscal policy hawks



?? on the *New York Times* on November 11th, 2008, forgets all this; including the arithmetics:

The economic lesson is the importance of doing enough. F.D.R. thought he was being prudent by reining in his spending plans; in reality, he was taking big risks with the economy and with his legacy. My advice to the Obama people is to figure out how much help they think the economy needs, then add 50 percent. It's much better, in a depressed economy, to err on the side of too much stimulus than on the side of too little.

What's behind this statement? Any theory? The consumption multiplier! Let's see what it is.

The consumption multiplier

The consumption multiplier is an analytical concept which arises in John Hicks' representation of Keynesian economics, the famed IS-LM model. It works as follows:

Suppose agents' consumption is a constant fraction c of income:

$$\begin{aligned}c_t &= a + c(w_t + g_t) \\c_{t+1} &= a + c(w_{t+1} + g_{t+1})\end{aligned}$$

Also, income w_t at time t is endogenous, equal to consumption plus investment (the economy produces to satisfy its demand). Not having investment, we have $c_t = w_t$:

$$w_t = a + c(w_t + g_t) \implies w_t + g_t = \frac{a + g_t}{1 - c}$$

The consumption multiplier is

$$\frac{1}{1 - c} > 1:$$

Every extra unit of government spending g_t increases aggregate income $w_t + g_t$ of $\frac{1}{1-c} > 1$ units.

I think it's fair to say that the multiplier as a theoretical construct concept has been relegated to the disrespect of macroeconomists by Bob Lucas' "Critique." Here's why: if agents are rational and choose consumption to maximize present discounted expected utility, then the marginal consumption out of income, c , is not a constant - consumption depends on permanent income - and $\frac{dc_t}{dg_t}$ will depend on income today but also on income tomorrow. As a consequence, taxes tomorrow enter the determination of consumption today. In a world without frictions (e.g., when nobody is borrowing constrained) then government expenditure has no effect (Ricardian equivalence). But even with frictions, the multiplier depends on the effect of government expenditure on the permanent income.

The dynamic multiplier

But it takes much more than this to wear Paul Krugman down. He can take all sorts of ?? :

Even the claim that we'll have to pay for stimulus spending now with higher taxes later is mostly wrong. Spending more on recovery will lead to a stronger economy, both now and in the future — and a stronger economy means more government revenue.

10.1 RICARDIAN EQUIVALENCE - A.K.A. ARITHMETICS¹⁸³

Stimulus spending probably doesn't pay for itself, but its true cost, even in a narrow fiscal sense, is only a fraction of the headline number.

What is behind this outlandish statement that government expenditures now have a positive effect on GDP in the future: the dynamic multiplier of Blinder-Solow.⁵

Here's the story. Suppose agent's consumption depends on wealth, rather than on income:

$$c_t = cW_t.$$

If we define wealth as permanent income,

$$W_t = w_t + \frac{1}{1+r}w_{t+1}$$

we have the consumption equation we have derived from first principles (dynamic optimization). Adding a government sector, this notion of permanent income implies that wealth is as follows:

$$W_t^{RE} = w_t + \frac{1}{1+r}w_{t+1} + \left(g_t + \frac{1}{1+r}g_{t+1} - \tau_t + \frac{1}{1+r}\tau_{t+1} \right)$$

We call this notion of wealth W_t^{RE} as a reminder that it implies Ricardian equivalence. Recall in fact that the government budget balance

$$\begin{aligned} g_t - \tau_t &= B_t \\ g_{t+1} + (1+r)B_t &= \tau_{t+1} \end{aligned}$$

where B_t is government debt at t , can be written as:

$$g_t + \frac{1}{1+r}g_{t+1} - \tau_t + \frac{1}{1+r}\tau_{t+1} = 0$$

Recognizing that government balance implies that $(g_t + \frac{1}{1+r}g_{t+1} - \tau_t + \frac{1}{1+r}\tau_{t+1}) = 0$, the representative agent permanent income is not affected by government expenditures, which have then has no effects.

⁵Though the first formulation apparently is in Carl Christ (1968). Thanks to Alessandro Lizzeri for reminding me of this concept - which I had not heard of anymore since my undergraduate times of Dornbush-Fisher economics).

Naturally, the notion of wealth in Blinder-Solow's formulation of the dynamic multiplier, let us be denoted W_t^{BS} must be different.⁶ They assume the following:

$$W_t^{BS} = w_t + \frac{1}{1+r}w_{t+1} + B_t.$$

The rationale for W_t^{BS} is that agents in equilibrium must hold the government debt, B_t , in their portfolios, and hence consider it as part of their wealth. Once again, however, they do not anticipate the taxes at time $t+1$ necessary to balance the budget.

Suppose that $\tau_t = 0$ for simplicity. Then

$$c_t = c \left(w_t + \frac{1}{1+r}w_{t+1} + g_t \right)$$

and government expenditures have an effect at time t .

Kevin Murphy's static set-up



(Photo by Dan Dry)

Kevin Murphy

⁶Blinder-Solow's formulation also accounts for changes in the interest rate at equilibrium - which we disregard in the analysis here without losing the thrust of the argument.

10.1 RICARDIAN EQUIVALENCE - A.K.A. ARITHMETICS¹⁸⁵

Kevin Murphy has a very clean set-up to collect all the costs and benefits of government expenditures. We reproduce it here.

Let's set some notation:

g = increase in government spending with respect to the steady state spending (the level of spending which would have obtained were there not a recession); that is, size of the fiscal stimulus;

$1 - a$ = value of a dollar of government spending; that is, a measures the resource allocation inefficiency due to the fact that the government chooses what to spend on rather than households;

f = fraction of government spending obtained by borrowing constrained agents;

l = relative value of relaxing the borrowing constraint;

d = deadweight cost per dollar of the revenue from future taxation required to finance G .

Then, the benefits in value due to a fiscal stimulus of size g are:

$$(1 - a)g + lfg$$

The costs are:

$$(1 - f)g + dg$$

The net gains (benefits minus costs) of a fiscal stimulus g are then:

$$\text{Net Gain} = (1 - a)g + lfg - [(1 - f)g + dg] = [f(1 + l) - (a + d)]g$$

Paul Krugman, Brad De Long, and the economist who are more openly in favor of the stimulus tacitly assume:

$a = 0$, that is, the government does what's best for agents, in terms of their own (the agents') preferences;

l is high, that is, borrowing constrained agents have a very high marginal utility of consumption;

$d = 0$, that is, taxes are not distortionary

They also assume f is large – close to 1.

10.2 Optimal taxation



James Mirlees

We study optimal taxation in this Chapter. By "optimal" we mean that the government raising taxes does this in the interest of the agents of the economy and not in the personal interest of the bureaucracies that constitute the government itself. This is a strong assumption, but we will show that, even in this case, *we need to find ways to control the sovereign* (Ed Prescott, who has been awarded the Nobel Prize in 2005 for his work with F. Kydland on Time Inconsistency).

10.2.1 Distortionary Taxation

Suppose a government has to raise an amount g to provide the economy with a public good (e.g., military defence, police protection, communication infrastructures, education and health care,...). The public good consists of units of the single consumption good in the economy at time $t + 1$. The government raises taxes. We consider two different tax systems:

taxes on savings, and

lump-sum taxes.

We will show that lump-sum taxes are to be preferred in terms of Pareto efficiency. Taxes on savings are most commonly used though because they are easier to implement in real economies.

Consider a two-period economy with a representative agent. A production technology allows the agent to save any non-negative amount S_t at t and receive $(1+r)S_t$ at $t+1$. With lump-sum taxes τ the representative agent's budget constraints are:

$$c_t + S_t = w_t, \quad c_{t+1} = S_t(1+r) - \tau$$

$$\tau = g$$

With taxes on savings, instead:

$$c_t + S_t = w_t, \quad c_{t+1} = S_t(1+r)(1-\tau)$$

$$S_t(1+r)\tau = g$$

If the representative agent has period utility $u(c)$, and discounting β , the first order conditions of his/her maximization problem with taxes include:

$$\frac{u'(c_{t+1})}{u'(c_t)} = \beta(1+r)[\text{with lump-sum taxes}]$$

$$\frac{u'(c_{t+1})}{u'(c_t)} = \beta(1+r)(1-\tau)[\text{with taxes on savings}]$$

We say that the saving choice of the agent is distorted with taxes on savings but not with lump-sum taxes. What we mean is that the first order condition which determines savings at the margin is changed (with respect to the benchmark with no taxes) with taxes on savings but not with lump-sum taxes. This is the source of the inefficiency of taxes on savings: A planner with a constraint that g must be raised faces aggregate endowment y_t and $y_{t+1} - g$ at $t+1$ (and a saving technology with return $1+r$) and chooses the consumption allocation which results from lump-sum taxes. [Make sure you understood this point!]

10.2.2 Public Goods Provision

Consider the economy in the previous section. Assume that the government must now choose an amount g of public good provision. The public good has utility for the representative agent in the amount $v(g)$. The government

is aware that taxes on savings *distort the margin* (as we have shown in the previous section), but has no other financing mean available.

The government will have to choose public good provision g and tax rate τ to maximize the representative agent's utility while satisfying the constraint that the public good needs to be financed by taxes on capital: $g = (1 + r) \tau S_t$. The government, on the other hand, cannot prescribe a saving choice to the representative agent. (This is what distinguishes a government from a planner: The government has power of raising taxes, but agents make their own economic choices; the planner is instead a conceptual construct we use as a benchmark to define Pareto optimality, who chooses all of the agents' allocations, including savings). Write down the planner's problem and its first order conditions; recall that there is a representative agent and hence the objective function of the planner does not depend on parametric weights.

Recall that the public good is in units of the consumption good at time $t+1$. Consider the case in which at time t the government decides the amount of the public good to be provided and the tax rate to finance it. We solve this problem in two steps: first we solve for the amount of savings that the agent chooses given any g and hence any τ (given g , τ is determined by the financing constraint $g = (1 + r) \tau S_t$); then we solve for the optimal level of g chosen by the government.

The first step, that is, the choice problem of the representative agent, is the following:

$$\max_{c_t, c_{t+1}, S_t} u(c_t) + \beta (u(c_{t+1}) + v(g))$$

subject to:

$$c_t + S_t = w_t, \quad c_{t+1} = S_t (1 + r) (1 - \tau) \quad (10.1)$$

$$(1 + r) \tau S_t = g \quad (10.2)$$

Assume $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$, with $\gamma < 1$; and assume $w_{t+1} = 0$. Then the first order condition with respect to savings S_t (substitute (10.1), not (10.2), in the utility function of the agent) is (derive it; it takes a bit of algebraic work):

$$\frac{S_t}{w_t - S_t} = (\beta (1 + r) (1 - \tau))^{\frac{1-\gamma}{\gamma}} \quad (10.3)$$

For this equation it is relatively easy to see that $\frac{\partial S_t}{\partial \tau} < 0$. [To prove it either use the Implicit Function Theorem or reason as follows: the right-hand-side

is increasing with S_t ; the left-hand-side is decreasing in τ if $\gamma < 1$, as we assumed; therefore, if τ increases, the left-hand-side decreases and hence the right-hand-side must also decrease, which can only be if S_t decreases.] We can also similarly establish, by substituting (10.2) into (10.3) and re-doing the argument, that $\frac{\partial S_t}{\partial g} < 0$ (do this; S_t appears in (10.2) and hence the argument needs be modified a little bit).

We conclude therefore that, for given public good provision g financed through taxes on savings, the representative agent saves an amount $S(g)$, with $S'(g) = \frac{\partial S_t}{\partial g} < 0$.

We can now solve for the second step, the optimal level of public good provision g of the government. It is the solution of the following maximization problem (note that we have substituted the expression for consumption into the utility function, as well as the expression for savings, $S(g)$, and for taxes, $\frac{g}{(1+r)S(g)}$, from the financing constraint):

$$\max_g u(w_t - S(g)) + \beta u \left(S(g) (1+r) \left(1 - \frac{g}{(1+r)S(g)} \right) \right) + v(g)$$

The first order condition can be written as follows:

$$\beta v'(g) = \beta u'(c_{t+1}) - S'(g) (\beta u'(c_{t+1})(1+r) - u'(c_t)) \quad (10.4)$$

Let's analyze this condition carefully. Note the following:

the left-hand-side is the marginal benefit of g and the right-hand-side is its marginal cost;

the term $-S'(g) (\beta u'(c_{t+1})(1+r) - u'(c_t))$ is positive if $\tau > 0$:

$$S'(g) < 0;$$

$$\beta u'(c_{t+1})(1+r) - u'(c_t) > 0 \text{ if } \tau > 0, \text{ from the first order condition of the representative agent problem, which include } \frac{u'(c_t)}{u'(c_{t+1})} = \beta(1+r)(1-\tau);$$

it increases with the absolute value of $S'(g)$, and with $\beta u'(c_{t+1})(1+r) - u'(c_t)$.

We can now compare the solution of the government problem with the solution of the planner's problem (that I have asked you to write down at the

beginning of this section). The first order conditions of the planning problem (or of the government problem with lump-sum taxes) include the following:

$$\beta v'(g) = \beta u'(c_{t+1})$$

(check this! convince yourself, this is important). In this equation, also, the left-hand-side represents the marginal benefit of g and the right-hand-side its marginal cost. Therefore we have shown above that the marginal cost of g in the government problem is greater than the marginal cost associated to the planner's problem, $\beta u'(c_{t+1})$, if $S'(g) < 0$ and $\tau > 0$. We conclude that the provision of the public good in the government problem is less than the efficient amount (resulting from the planner's problem). [This is not straightforward, because the levels of consumption c_t and c_{t+1} in the two problems are also different; in other words, it is not enough to compare the first order conditions we did compare, but we need to look at the whole set of first order conditions. The result is true, however, because savings is smaller in government's problem, and hence c_{t+1} smaller. Try to prove this if you feel strong.] Equivalently, the choice of public provision g by the government with taxes on savings is smaller than the choice that would derive from a government problem with lump-sum taxes. [Perhaps it is worth convincing yourself of this; write down the problem with lump-sum taxes and solve it; show the solution coincides with the solution of the planning problem.]

Note that $\beta u'(c_{t+1})(1+r) - u'(c_t)$ is a measure of the distortion due to taxes on savings; that is, when taxes are lump-sum it is always true that $\beta u'(c_{t+1})(1+r) = u'(c_t)$, while with taxes on savings $\beta u'(c_{t+1})(1+r) = \frac{u'(c_t)}{1-\tau} > u'(c_t)$, and hence $\beta u'(c_{t+1})(1+r) - u'(c_t)$ increases with τ .

10.2.3 Time Inconsistency



Ed Prescott and Finn Kydland

Remember we considered *the case in which at time t the government decides the amount of the public good to be provided and the tax rate to finance it*. Suppose now we consider the case in which at time $t + 1$, after savings decisions have been made by the representative agent on the basis of the levels of g and τ which we derived in the previous section, the government can choose possibly a different level of public good provision G and associated tax T to satisfy the financing constraint. What is the government problem now?

The government at time $t + 1$ takes as given the amount of saving $S(g)$, and chooses G as the solution to the following problem:

$$\max_G \beta \left(u \left(S(g) (1 + r) \left(1 - \frac{G}{(1 + r) S(g)} \right) \right) + v(G) \right)$$

The first order condition of this problem is:

$$u'(c_{t+1}) = v'(g)$$

Convince yourself that the G which solves this problem is greater than G . (Now, this is easy, because $S(g)$ does not change, and hence c_{t+1} only changes in account of T .)

We conclude that, if the government would be allowed to change its decision at time $t+1$, it would. Even though the government is always maximizing

the utility of the representative agent. The intuition is that at time t the government limits the provision of the public good to limit the distortions that taxes impose on savings. At time $t + 1$ instead savings has already happened, and hence raising the public good provision (and taxes) does not distort savings.

Of course the agents at time t might not trust the government not to raise taxes at time $t + 1$, and hence might reduce savings nonetheless. What do you expect it will happen? You should be able to answer this informally. For a formal answer we need *Game Theory* - coming in the next classes.

10.3 Problems

Problem 10.3.1 *What do we (beginner economists) expect the effects of President-elect Barack Obama's proposed government expenditure plan to be? In particular, what will the effect on consumption be? IMPORTANT: Your answer should be no longer than 10 lines (and do not write tiny letters); Negative points are taken away for answers of the sorts I can hear on the street (-15 points is the most I will take away); I want a coherent argument - as an economist would do - there is no "correct answer", just good and bad arguments.*

Chapter 11

Empirical economics



Karl Pearson



Jacob Marschak



Tjalling Koopmans



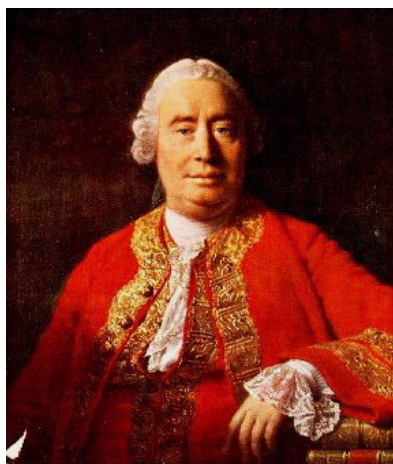
James Heckman

Understanding how economists test their models and how they measure fundamental economic quantities is a fundamental issue. As we noted in the Introduction, in fact, economics as a discipline is largely empirical, much more than it typically gets credit for.

The fundamental question of empirical analysis, in economics as in social sciences in general, consists in identifying *causal relationships*. It is a difficult task even to define what is a *causal relationships*, let alone identifying one in data.

The issue of causation is central to philosophical thought. While this is not the appropriate place for a discussion of the history of the philosophical thought on this issue, it is important at least to notice that the issue of

causation has been framed conceptually in a modern way by David Hume, *An Enquiry concerning Human Understanding and Other Writings*, written in 1748.¹ Cambridge: Cambridge UP, 2007. Print. (Section VII).



David Hume

Hume distinguishes "relations of ideas," which we would now call deductive arguments, from "matters of fact," inductive arguments. Hume also argues that "matters of fact" result from a mere habit or, in other words, form a subject's psychological mechanism, his idea of a causal relation derived from a habitual transition in the mind. In Hume's view a

matter of fact [is] an object, followed by another, and whose appearance always conveys [by habit] the thought to that other.

We facing a "matter of fact" then, according to Hume, we should not properly talk of causation, which is instead a property of "relation of ideas."

Economists refer instead to causality in the context of inductive reasoning. If Hume would argue that the relationship between the ingestion of food and the feeling of nourishment is not causal, economists would on the contrary accept this relationship as a causal one. Economists and statisticians make causal inferences from correlations in the data, while being aware of the limitations of inference methods.

¹See Hume, David, and Stephen Buckle, *An Enquiry concerning Human Understanding and Other Writings*, Cambridge: Cambridge University Press, 2007.

Economists define causation as the effects of changes in exogenous variables, parameters, and constraints on outcomes. Within a model, the effects on outcomes of variation in exogenous variables, parameters, and constraints is formalized by the Marshallian notion of *comparative statics*, i.e., the notion of a *ceteris paribus* change. Let e.g., a model represent the relationship between an outcome y and exogenous variables x_1 and x_2 as a functional relationship

$$y = f(x, z).$$

We say then that x (positively) causes y at (x, z) if $\frac{\partial f(x, z)}{\partial x} > 0$. Note that the partial derivative captures the notion of *coeteris paribus*: $\frac{\partial f(x, z)}{\partial x}$ encodes the effect of an (infinitesimal) change in x on y , *keeping the variable z constant*.

The requirement that x and z be exogenous variables is tantamount to a requirement that x and z can be controlled independently from y . Suppose instead that x and z were determined (in the model) jointly with y . They would then represent endogenous variables; the partial derivative $\frac{\partial f(x, z)}{\partial x}$ would then be well-defined mathematically, but would not represent a causal relationship.

11.1 Correlation and regression analysis

How can we then identify causal relationships in the data?² How can we observe independent changes of one variable x on an outcome y ?

Suppose we observe a series of observations of triples (x, z, y) e.g., over time or over space, or for different economic agents. Let observations of (x, z, y) be indexed by $i = 1, \dots, N$: the data are then a list $(x_i, z_i, y_i)_{i=1}^N$. We can then define the *correlation between x and y* in the data by means of the sample correlation coefficient,

$$\rho_{x,y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y},$$

where \bar{x} and \bar{y} are sample means, $\text{cov}(x, y)$ is the sample covariance, and σ_x

²This is the problem that the Cowles Commission for Research in Economics set out to explore in the late 30's in Chicago under the directorship of Jacob Marschak, first, and Tjalling C. Koopmans, later. The Commission's motto is illustrative of its purposes: *Science is Measurement*.

and σ_y are sample standard deviations:

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \\ \text{cov}(x, y) &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}), \\ \sigma_x &= \frac{1}{N} \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}, \quad \sigma_y = \frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}.\end{aligned}$$

A positive correlation between x and y , keeping z constant is an indication that on average, observations of x_i above its sample mean \bar{x} are on average associated to observations of y_i above its sample mean \bar{y} . Is this evidence of causation? Can we associate a positive correlation between x and y with a causal relationship between the two variables? The answer is negative, in general, because a positive correlation is silent about the exogeneity of either x or y , in exactly the same sense a positive partial derivative $\frac{\partial f(x,z)}{\partial x}$ is.

In the words of Karl Pearson, one of the fathers of mathematical statistics and in particular of the notion of correlation,

correlation does not imply causation.

Then again, in some specific empirical contexts, we can guarantee that some variables are exogenous. In this case, *regression analysis*, a generalization and an application of the notion of correlation is useful. Consider again data in the form of a list $(x_i, z_i, y_i)_{i=1}^N$ and an underlying model which implies that $y = f(x, z)$. Suppose we are willing to assume that the underlying model is (closely approximated to be) linear, modulo some non-systematic errors, e.g., measurement errors. That is,

$$y_i = \alpha + \beta_x x_i + \beta_z z_i + u_i,$$

where u_i is the residual error. The requirement that errors are non systematic implies that u_i assumed to be a random variable with zero mean. Define the best linear model as the one which minimizes the implied errors. More precisely, the implied errors are measured by the sum (across realizations) of squared errors, to have a measure which is independent of the sign of the errors themselves: $\sum_{i=1}^N (u_i)^2$.

In this context, regression analysis is the characterization of the best linear model for the given data, that is, the linear function which minimizes the the sum of squared errors in the data. It results as the solution of the following minimization problem:

$$\min_{\alpha, \beta_x, \beta_z} \sum_{i=1}^N [y_i - (\alpha + \beta_x x_i + \beta_z z_i)]^2.$$

Theorem 11.1.1 *Suppose we observe data $(x_i, z_i, y_i)_{i=1}^N$. The best linear model for the data is*

$$y_i = \alpha + \beta_x x_i + \beta_z z_i + u_i,$$

with

$$\begin{aligned} \alpha &= \bar{y} - \beta_x \bar{x} - \beta_z \bar{z} \\ \beta_x &= \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \\ \beta_z &= \frac{\text{cov}(z, y)}{\sigma_z^2} = \frac{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})(y_i - \bar{y})}{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2} \end{aligned}$$

Proof. *The f.o.c. of the minimization problem with respect to α, β_x, β_z are:*

$$\begin{aligned} 2 \sum_{i=1}^N (y_i - (\alpha + \beta_x x_i + \beta_z z_i)) &= 0 \\ -2 \sum_{i=1}^N (y_i - (\alpha + \beta_x x_i + \beta_z z_i)) x_i &= 0 \\ -2 \sum_{i=1}^N (y_i - (\alpha + \beta_x x_i + \beta_z z_i)) z_i &= 0 \end{aligned}$$

Some simple (and somewhat tedious) arithmetics allows now to obtain the expressions in the statement. Finally, second order conditions guarantee that the solution of this system represents a minimum of the problem (not a maximum), as the objective function $\sum_{i=1}^N [y_i - (\alpha + \beta_x x_i + \beta_z z_i)]^2$ is (easily shown to be) convex in $(\alpha, \beta_x, \beta_z)$. ■

11.2 Controlled experiments

A *controlled experiment* (or Randomized Controlled Trial - RCTs) compares the results obtained from an experimental group of subjects against a *control group*, which is identical to the experimental group except for the one aspect whose effect is being tested (the exogenous variable). A good example would be a drug trial. The group receiving the drug would be the experimental one; and the one receiving the placebo would be the control one.

What is essential for a controlled experiment is that

1. the sample of subjects be representative of the population under study, and
2. the distinction between the experimental and the control group be random, so that the two groups are statistically equivalent.

In general, a sample is not representative of the population if subjects are allowed to choose to participate in the experiment depending on some relevant characteristic of the experiment itself; e.g. in a drug test, if only subjects with some specific (intense) manifestations of a disease are choosing to participate.

The earliest controlled experiments were due to Charles Darwin who, with the help of Francis Galton, compared the growth of cross versus self-fertilized plants (Darwin, 1876).³ Controlled experiments are now common practice in medicine and other sciences: as of 2004, more than 150,000 RCTs were recorded in the Cochrane Library, an archive created to prepare, maintain, and disseminate systematic reviews of studies on the effects of health care programs.

³It is said that Darwin was influenced to some degree by path-breaking experiments in physical sciences due to Michael Faraday.

Controlled experiments are also typically used in what is called *experimental economics*, an important sub-discipline in the academic economic profession. More recently, controlled experiments have also been used in development economics, notably e.g., by the Mexican government in the PROGRESA (now called Oportunidades) project and at the Poverty Action Lab at MIT. We shall feel content with briefly introducing just a few examples illustrating the methodology.

Experimental economics. [...] Forsythe, Horowitz, Savin, and Sefton (1994). [...]

Development economics. The PROGRESA program in Mexico offers grants, distributed to women, conditional on children's school attendance and preventative health measures (nutrition supplementation, health care visits, and participation in health education programs). In 1998, when the program was launched, the Mexican government decided to start with a randomized pilot program in 506 of the 5,000 communities identified as potential beneficiaries of the program. Half of those were randomly selected to receive the program, and baseline and subsequent data were collected also in the remaining communities. Many researchers have studied the resulting data.

Duflo, Kremer, and Robinson (2006) evaluated a series of different interventions to understand the determinants of the adoption of fertilizer in Western Kenya. The design of the interventions made it possible to test some of the standard hypotheses of the hindrance to the adoption of new technology. Field demonstrations with treatment and control plots were conducted to evaluate the probatability of fertilizer in the local conditions. Because the farmers were randomly selected, those field trials also allowed them to study the impact of information provision and the channels of information transmission; other ways to provide information (starter kits, school based demonstrations) were also examined; finally, financing constraints and difficulties in saving were also explored, with interventions helping farmers to buy fertilizer.⁴

11.3 Natural experiments

Controlled experiments are a useful inference technique because, by controlling the sampling procedures, the experimenter can guarantee that the

⁴This paragraph is taken verbatim from Duflo, Glennerster and, Kremer (2006), Using Randomization in Development Economics Research: A Toolkit, mimeo MIT.

sample of subjects be representative of the population under study and that the distinction between the experimental and the control group be random. Even though an experiment is not controlled, however, it might nonetheless be that the sample satisfies these properties. In this case we say that we have discovered a *natural experiment*.

One of the earliest natural experiments was the 1854 Broad Street cholera outbreak. On August 1854, a major outbreak of cholera struck London. Deaths appeared concentrated near Broad Street. The physician John Snow identified the source of the outbreak as the polluted public water offered by pumps managed by one of the several companies active in London. Given the near random patchwork development of the water supply in mid-Nineteenth Century London, Snow considered the circumstance as a (natural) "experiment...on the grandest scale" on the public health effects of the exposure to the polluted water. The treatment group, in the experiment, is composed of the citizens living in a neighborhood whose water is supplied by the company which supplies Broad Street, while the control group is composed of all citizens whose water is supplied by a different company.

A more recent famous natural experiment regarding the effect of family size on the labor market outcomes of the mother has been studied by Angrist and Evans (1998). The correlations between family size and various labor market outcomes do not tell us how family size causally affects labor market outcomes because both labor market outcomes and family size may be affected by unobserved variables such as preferences and because labor market outcomes may itself affect family size. But two-children families with either two boys or two girls are substantially more likely to have a third child than two-children families with one boy and one girl. The gender of the first two children, then, forms a natural experiment: it is as if an experimenter has randomly assigned some families to have two children and others to have three or more. The authors are then able to estimate the causal effect of having a third child on labor market outcomes by comparing the labor market outcome of the mothers in the treatment group (three-children families with the first two of the same gender) with those of mothers in the control group (two-children-of-different-gender families).

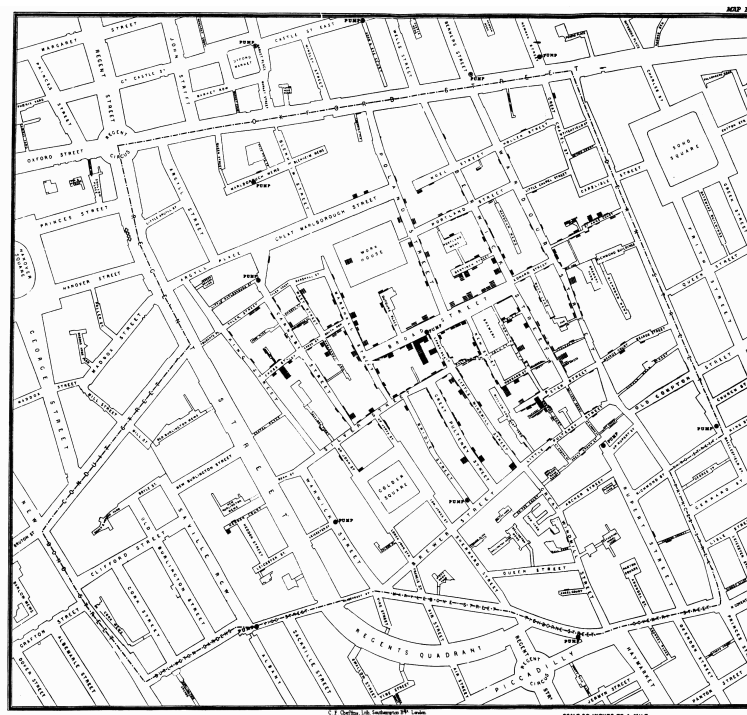


Figure 11.1: Original map by John Snow showing the clusters of cholera cases in the London epidemic of 1854

11.4 Structural econometrics

What then of an empirical context in which controlled experiments are not possible and the data has not been generated by what could be considered as a natural experiment? As we noted above, correlation and regression analysis are typically not sufficient to identify causal relationships in the data, as we have no guarantee that the necessary exogeneity assumptions apply. In particular, in this case several classical problems plague regression analysis. We first illustrate the most important of them. We shall then argue that the combination of models and statistical methods, which is called *structural econometrics*, can be used to produce a useful inference methodology.

11.4.1 Problems with regression analysis

What happens if we run a regression on data $(x_i, z_i, y_i)_{i=1}^N$ even though (x_i, z_i) are not exogenous? Here are some classic examples.

Identification

Suppose you run a regression

$$y_i = \alpha + \beta_x x_i + u_i$$

where y_i is the aggregate quantity demanded of a specific commodity - say, oil - at time i , and x_i is the market price of oil at time i . Suppose you find that the parameter β_x is positive. Do you conclude that the demand for oil is positively sloped, that consumers are irrational (or at least have strange preferences), that economics is useless, and you pass to sociology? No. The fact is that the parameter β_x is not identified, you might be picking up supply rather than demand. Here's the explanation.

The observed relationship between y_i and x_i in the data is an equilibrium relationship. It's the outcome of

$$\begin{aligned} y_i^d &= \alpha^d + \beta_x^d x_i + u_i^d \\ y_i^s &= \alpha^s + \beta_x^s x_i + u_i^s \\ y_i &= y_i^d = y_i^s \end{aligned}$$

where the apex d (resp. s) indicates demand (resp. supply). As a consequence, we do not know if variation in (x_i, y_i) over time i is due to variation

of u_i^d or u_i^s or both. To make the contrast as stark as possible, consider the case in which $u_i^s = 0$, for any $i = 1, \dots, N$. In this case the variation is due to demand and the regression we have run estimates the supply function, which is in fact naturally positively sloped. The figure should illustrate.

Formally, the solution of the demand and supply system is:

$$\begin{aligned} x_i &= \frac{\alpha^d - \alpha^s + u_i^d - u_i^s}{\beta_x^s - \beta_x^d} \\ y_i &= \alpha^s + \beta_x^s x_i + u_i^s. \end{aligned}$$

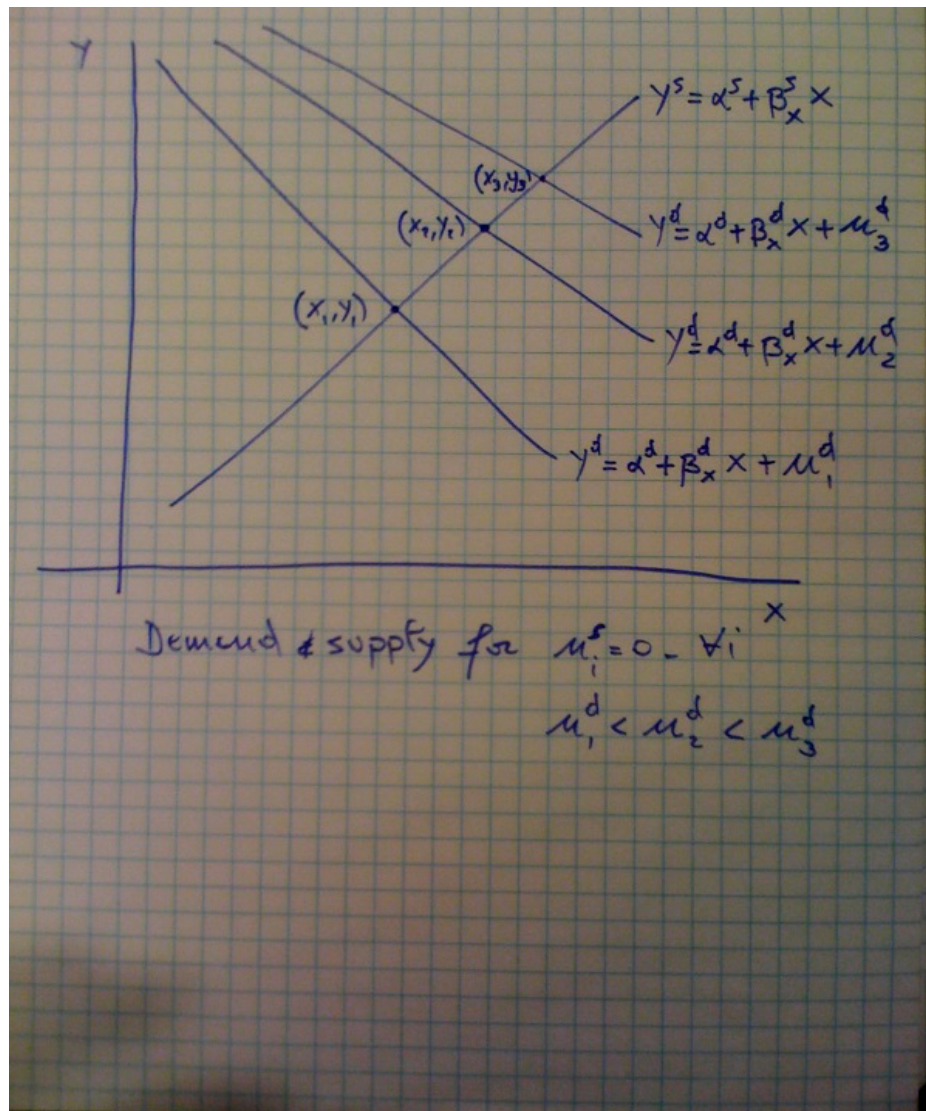
Suppose that $\beta_x^s > 0$ and $\beta_x^d < 0$, so that demand has a negative slope and supply a positive one, as to be expected. But suppose that $u_i^s = 0$, for any $i = 1, \dots, N$. Then $x_i = \frac{\alpha^d - \alpha^s + u_i^d}{\beta_x^s - \beta_x^d}$ and $y_i = \alpha^s + \beta_x^s x_i$. It follows that $\frac{\partial x_i}{\partial u_i^d} > 0$ and $\frac{\partial y_i}{\partial u_i^d} = \beta_x^s \frac{\partial x_i}{\partial u_i^d} > 0$. As a consequence positive demand shocks u_i^d are associated with higher prices x_i and higher quantities y_i , resulting in a positive slope in the regression, as in the figure. Convince yourself that, if $u_i^d = 0$, for any $i = 1, \dots, N$, we would find a negative slope in the regression. In fact, nothing can be said about the parameters of interest, the slope of the demand and the supply unless we observe (u_i^d, u_i^s) . We say that the slope of the demand and the supply are *not identified*.

Selection

Suppose you run a regression

$$y_i = \alpha + \beta_x x_i + u_i$$

where y_i is the monthly wage of individual i and x_i is a dummy variable equal to 1 if the individual has a college degree and 0 otherwise. Suppose you estimate $\beta_x = 2,347$. Can you conclude that the return of going to college is on average 2,347 dollars a month? The answer is once again negative. The reason is that the population of those who attend college is self-selected: they are the children of relatively rich parents who have connections and find them better jobs. (Ok, another possible selection argument - just in theory - is that those who go to college are smarter and this is why they get higher wages).



Reverse causality

Suppose you run a regression

$$y_i = \alpha + \beta_x x_i + u_i$$

where y_i is the growth rate of country i and x_i is the average income tax-rate in country i . Suppose you find a positive estimate for β_x . Can you conclude that lowering taxes induces higher growth rates? The answer is again negative. The reason is that causation could be reversed, it could well be that it is growth rates which causes taxes, e.g., because a state government whose economy grows fast might afford lower taxes. Remember: *regressions only document correlation, not causation!*

11.4.2 Back to models

Well, what do we do then if we do not have a natural (nor a controlled) experiment and if regressions are problematic? Well, we go back to models. Exogeneity is a property of variables *in the model*, never independently of the model. Empirical analysis is only as good as the model is: empirical results are conditional to the model. Statistical inference procedures (including regressions) assume validity (or lose it) in the context of a model. We cannot pronounce statements like,

x and z cause y and in our data $\frac{\partial y}{\partial x} = 1.3$, $\frac{\partial y}{\partial z} = 2.9$.

We can only say,

conditionally on the model $y_i = \alpha + \beta_x x_i + \beta_z z_i + u_i$, in our data $\frac{\partial y}{\partial x} = 1.3$, $\frac{\partial y}{\partial z} = 2.9$.

Nothing we can do. We ought to recognize our limitations (in statistical inference) and live with them.

11.5 Problems

Problem 11.5.1 *Acemoglu, Cantoni, Robinson (2009) ask if Napoleon can be considered a modernizing force in the territories he conquered. They use*

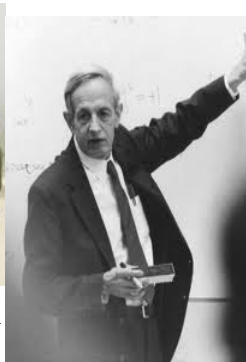
Germany as a test case, asking how regions controlled by Napoleon fared compared to those that were never under French rule. Occupied areas that kept the reforms Napoleon put in place (a written civil code; abolition of serfdom and, in the cities, of guilds that hampered economic freedom) subsequently experienced the most rapid economic development. Areas that the French never controlled developed less quickly (measured in terms of urbanization), and those regions that were invaded, but returned to their old ways after liberation, developed most slowly of all. Is this a natural experiment? Discuss.

Chapter 12

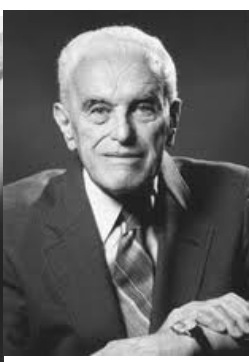
Game theory



Oskar Morgenstern
and John von
Neumann



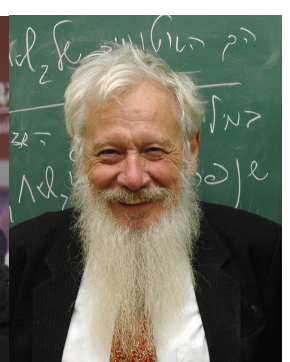
John Nash



John Harsanyi



Reinhard Selten



Robert Aumann

Game theory is the study of interacting decision-makers. In earlier sections we studied the theory of the consumer. That is, the optimal decision by a single decision-maker, who is called the consumer, in a simple environment in which she is a price taker. Thus, Game theory, is a natural generalization of the single decision-maker theory which deals with how a utility maximizer behave in a situation in which her payoff depends on the choices of another utility maximizer. Many fields, such as sociology, psychology and biology, study interacting decision-makers. Game theory focus on rational decision-making, which is the most appropriate model for a wide variety of economic contexts.

12.1 Strategic Games

A strategic game is a model of interactive decision-making in which each decision-maker chooses his plan of action, his strategy.

Any strategic game is defined by:

- (i) a (finite) set of players, $I = \{1, \dots, m\}$;
- (ii) a (finite) set of (pure) strategies which are the choices that each player can make: $s_i \in S^i$ is a pure strategy of player i ; $S = S^1 \times \dots \times S^m$ is the strategy space and $s \in S$ is a strategy profile.
- (iii) a set of payoffs that indicates the utility that each player has in any combination of strategies, $u^i : S \rightarrow R$.

Consider 2×2 games - 2 players and 2 pure strategies to each player. For such a game, the following bi-matrix representation is very convenient.

		Pl. 2	
		L	R
Pl. 1	T	A	C
	B	E	G
		B	D
		F	H

Representation of a 2×2 game

Player 1's actions are identified with the rows and the other player by the columns. The two numbers in a box formed by a specific row and column are the players' payoffs given that these actions were chosen. For example, in the game above A and B are the payoffs of player 1 and player 2 respectively when player 1 is choosing strategy T and player 2 strategy L .

Applying the definition of a strategic game to the game above yields:
 $I = 1, 2$ - the set of players
 $S^1 = \{T, B\}$, $S^2 = \{L, R\}$ and $S = S^1 \times S^2 = \{(T, L), (T, R), (B, L), (B, R)\}$
 Payoffs are given by the bi-matrix.

12.1.1 Nash equilibrium

A strategy profile s^* is Nash equilibrium if, for any player $i \in I$,

$$u^i(s_i^*, s_{-i}^*) \geq u^i(s_i, s_{-i}^*)$$

for all $s_i \in S^i$.

In other words, at a Nash equilibrium no player can profitably deviate from his strategy, given all other players strategies. The notion of a *best response* is most useful in finding the set of Nash equilibria.

Player i 's best response to a strategy s_j of player j is defined as

$$BR^i(s_j \in S^j) = \{(s_i \in S^i) : u^i(s_i, s_j) \geq u^i(s_{i'}, s_j)\}$$

A Nash equilibrium has the property that

$$s_j = BR^j(BR^i(s_j \in S^j))$$

12.1.2 Examples

The following are classical games representing a variety of strategic situations.

Two people wish to go out together to a ball game or to a concert. Their main concern is to be together, but one of them prefers the game and the other the concert.

Battle of the sexes

	G	C
G	2	0
	1	0
C	0	1
	0	2

Battle of the sexes

The game has two Nash equilibriums in pure strategies: (G, G) and (C, C) which is written as

$$N(G) = \{(G, G), (C, C)\}$$

Coordination game

As in the *BoS*, the two players wish to go together but this time they agree that to see a ball game is much better.

Coordination

		G	C
G	2	0	
	2	0	
C	0	1	
	0	1	

Coordination

Find the set of Nash equilibrium $N(G)$. In addition, discuss its reasonableness.

The prisoner dilemma

Two suspects in a crime are under a police investigation (in different cells). If they both confess, each will be sentenced for three years. If one of them doesn't confess and the other does, the first one will be used as a witness against the other, who will receive a sentence of four years. If both don't confess, they will be convicted in minor offense and spend only one year in prison. As it is standard, we shall encode actions in a confusing way. Action C is not "confess" but rather "cooperate" with the other prisoner, that is, "don't confess". It gets worse: action D is "defect", that is "confess".

Prisoner Dilemma

	C	D
C	0, 0	-4, 1
D	1, -4	-3, -3

Prisoner Dilemma

Find the set of Nash equilibrium $N(G)$.

Cournot Oligopoly

Two firms compete on the quantity they produce of good a single commodity. They face a demand function

$$p = f(x)$$

where p is the price at which they will sell the good, which depends on the total quantity produced, $x = x_1 + x_2$ (x_i is the quantity produced by firm $i = 1, 2$). Let the demand be linear:

$$p = \beta - \gamma(x_1 + x_2)$$

The production cost for firm i is

$$c(x_i) = -\alpha(x_i)^2$$

and its revenues are

$$px_i.$$

12.1.3 Mixed Strategies

Some games do not have a Nash equilibrium *in pure strategies*, as we defined it up to now. In those games we study *mixed strategies*, constructed as follows. Let $S^i = \{s_1^i, \dots, s_N^i\}$ denote the pure strategies of agent i . The corresponding set of mixed strategies is the set of all probability vectors

$$\pi^i = \{\pi_1^i, \dots, \pi_N^i\}, \quad \sum_{n=1}^N \pi_n^i = 1$$

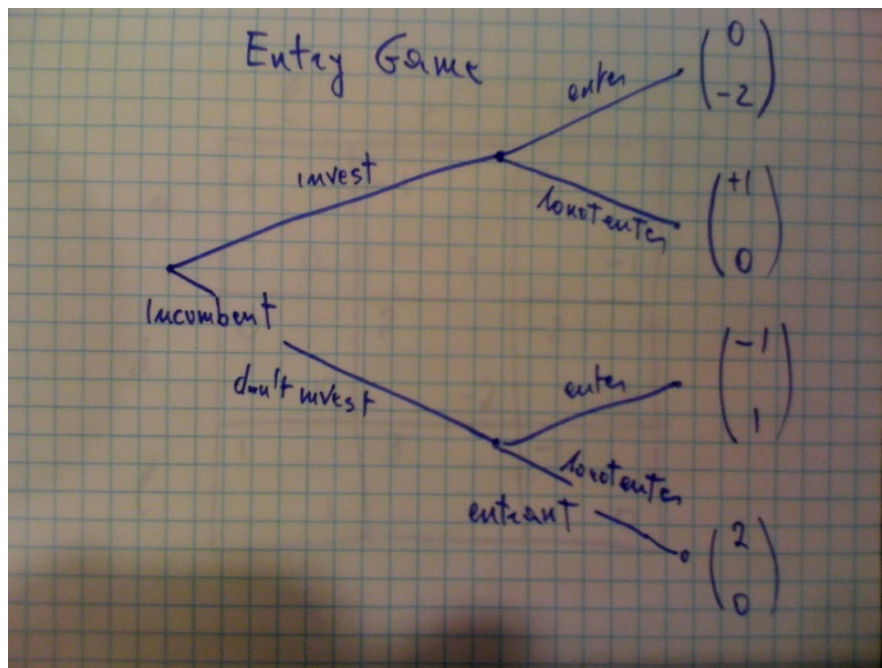
where π_n^i takes the interpretation of the probability that the pure strategy s_n^i is played by player i .

Matching pennies

Each of the two players chose either head or tail. If they choose differently player 1 pays \$1 to player 2 and if they choose the same player 2 pays \$1 to player 1. This kind of a game is also called strictly competitive. Moreover, this game is also a zero sum game.

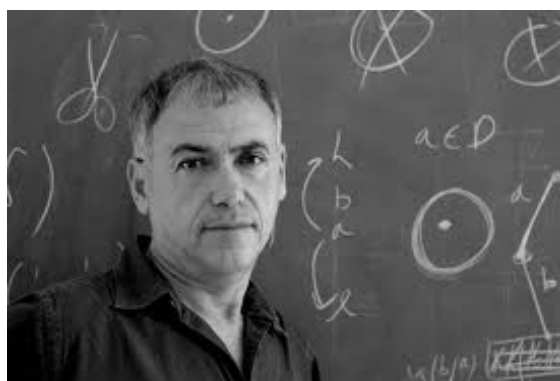
12.2 Dynamic Games

Consider a game in which decisions occur over time. A typical example is the entry game, in the figure.



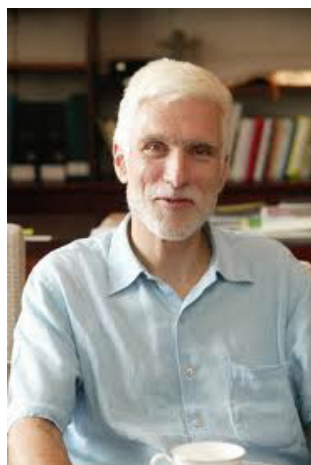
Entry

12.3 Bargaining



Ariel Rubinstein

12.4 Reputation and repeated games



Drew Fudenberg



David Levine

12.5 Problems

Problem 12.5.1 Consider the following 3×2 game. Find first all (if any) pure strategy Nash equilibria of this game. Then look for all mixed strategy equilibria.

	l	c	r
u	1	-1	2
d	0	2	-2
l	1	3	-1

Problem 12.5.2 Consider a duopoly (an oligopoly with two firms), $i = 1, 2$. They are symmetric and produce the same good. Each firm i chooses the dollar price she charges, p_i , and then produces to satisfy the demand at this price. Production of x_i units involve dollar costs equal to $c_i = \alpha x_i$. Demand d comes from a representative agent and is a declining function of the price p the agent is charged:

$$d = f(p).$$

The agent chooses the firm he patronizes. Write down the profit function of firm 1 as a function of her own and firm 2 price. Write down the profit maximizing price p_1 as a function of p_2 (use also your economic understanding of the problem; the mathematical maximization problem is not standard!) Exploiting symmetry, write down the profit maximizing price p_2 as a function of p_1 . Compute the unique Nash equilibrium of the game.

Chapter 13

Mathematical Appendix

We collect here some useful mathematical preliminaries.

13.1 Sets

A *set* is any well-specified collection of *elements*. A set may contain finitely many or infinitely many elements. For example

$$A = \{1, 2, 3\}$$

$$B = \{1, 2, 3, \dots\}$$

The term element is left as an abstract concept. For our purposes an element will usually be a number, but there is no need to restrict it to be so in its definition. The notation $a \in A$ means a is a member of the set A . Conversely, $a \notin A$ means a is not a member of the set A .

The following is an example of standard notation used to define sets: $C := \{x \in A : x > 1\}$. This means the set C is defined as the elements in A that are larger than 1 (i.e. $C = \{2, 3\}$). A set that contains no elements is called the *empty set* or *null set* and is denoted by \emptyset . Let A, B , be the sets defined above. The set A is contained in B (i.e. all the elements of the set A are also in the set B) and is called a *subset* of B . This is written $A \subseteq B$. This can also be written the other way round, $B \supseteq A$ and reads " B is a *superset* of A ".

Using this definition two set, say A and A' , are *equal* if $A \subseteq A'$ and $A' \subseteq A$. A set A is a *proper subset* of a set A' if $A \subseteq A'$ and $A' \not\subseteq A$. This is

denoted $A \subset A'$. Given two sets, say A and B , new sets can be formed using the following operations

Union: $A \cup B := \{x : x \in A \text{ or } x \in B\}$

Intersect: $A \cap B := \{x : x \in A \text{ and } x \in B\}$

Set subtraction: $A \setminus B := \{x : x \in A \text{ and } x \notin B\}$

If it is clear that the sets under discussion are subsets of some ("universal") set, say \mathbb{U} , then it is possible to define the operation *complement*:

$A^c := \{x \in \mathbb{U} : x \notin A\}$

13.2 Numbers

The most basic numbers are the counting numbers $\{1, 2, 3, \dots\}$, usually called the *natural numbers* and denoted by the script letter \mathbb{N} . The natural numbers augmented with the number zero and all the negatives of the natural numbers is called the set of *integers*, and is denoted $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$. Since the sum, difference and product of two integers is another integer but the quotient of two integers is not in general, the natural extension is to the set that includes such values. This leads to the definition of the *rational numbers*. This is the set of all numbers that can be formed by the quotient of an element of the integers with an element of the natural numbers, i.e.

$$\mathbb{Q} = \left\{ \frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N} \right\}$$

Things do not end at the rationals. There is a sense in which the set of rationals has a great deal of "holes", numbers that cannot be written as a quotient of an integer and a natural number (e.g. $\sqrt{2}, \pi \notin \mathbb{Q}$). This leads to the set of *real numbers*, denoted \mathbb{R} . The construction of the real numbers is not for this class and as such there will be no formal definition of the reals.

Some commonly used sets of numbers:

$\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$, called the *non-negative* real number;

$\mathbb{R}_{++} = \{x \in \mathbb{R} : x > 0\}$, called the *positive* real numbers;

$[0, 1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$, called the *closed interval* between zero and one;

$(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$, called the *open interval* between zero and one;

The sets $\mathbb{R}_-, \mathbb{R}_{--}, (a, b), [a, b], [a, b), (a, b]$ are similarly defined, where $a, b \in \mathbb{R}$ such that $a > b$.

13.3 Vectors

A *vector of size n* is an n -tuple which in general is written as $\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$.

For example, if $n = 2$ the general form of the vector is $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. If $x_1 \in \mathbb{R}$

and $x_2 \in \mathbb{R}$, then $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ where $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. The " \times " operator is the *Cartesian Product*.

Set notation, vectors and graphs: Here are some examples of sets of vectors (of size 2) that can be represented on a 2-D graph

$$\begin{aligned} &\mathbb{R}_+ \\ &\mathbb{R}_{++}^2 \\ &\{(x_1, x_2) \in \mathbb{R}_+^2 : x_2 \leq m + bx_1\} \end{aligned}$$

13.4 Functions

A *function* is a rule that describes a relationship between numbers. For each number x a function assigns a *unique* number y according to some rule. The general notation, when the specific algebraic rule is not known, is $y = f(x)$ where the number x is called the *independent variable* and the number y is called the *dependent variable*.

Often some variable y depends on several other variables, say x_1, x_2 . In this case, we write $y = f(x_1, x_2)$ to indicate that both variables together determine the value of y . To formalize matters, a function maps *each* element of a set, called the *domain* of the function, into a *unique* element of another set, called the *co-domain*. This is written $f : D \rightarrow C$, where D is the domain and C is the co-domain. The function is only defined for the elements of the domain. Related to the co-domain is the *range*. This is the set of all values that the function takes, $\{y \in C : \exists x \in D \text{ such that } f(x) = y\}$ (NB: the symbol " \exists " is common short hand for "there exists"). The range is always a

subset of the co-domain.

The functions that are used in this book are *real valued functions*. These are functions whose co-domain is a subset of the real numbers, i.e. $C \subset \mathbb{R}$. This hand-out will focus at first on *real valued functions on the real line*, referred to as *univariate functions*. Here the domain is a subset of the real numbers, i.e. $D \subset \mathbb{R}$. Later we will deal with multivariate functions, where $D \subset \mathbb{R}^n$ for $n \in \mathbb{N}$.

Some examples of real valued functions on the real line (for each of the functions below you should think of suitable domains, co-domains and the corresponding ranges):

- The simplest function is called a *monomial* and is written $y = \alpha x^k$ where $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}$ is called the *degree* of the monomial. A *polynomial* is a function which is formed by adding monomials. For example a polynomial of degree $k = 2$ is written $y = a_2x^2 + a_1x + a_0$. More generally a polynomial of degree k is $y = \sum_{i=0}^k a_i x^i$.
- A *rational function* is a ratio of polynomials. For example, $y = \frac{x^5+1}{x-5}$.
- A polynomial of degree $k = 1$ is a *linear function*. It has the form

$$f(x) = mx + b$$

where $m, b \in \mathbb{R}$. Note that a polynomial of degree $k = 0$ is also a linear function which assigns the same number for any choice of x . It is called a *constant function*.

13.4.1 Some Properties of Functions

Increasing and decreasing functions:

A function $f(x)$ is an *increasing function* if

$$x_1 > x_2 \implies f(x_1) \geq f(x_2)$$

and a *strictly increasing function* if

$$x_1 > x_2 \implies f(x_1) > f(x_2)$$

A function $f(x)$ is a *decreasing function* if

$$x_1 > x_2 \implies f(x_1) \leq f(x_2)$$

and a *strictly decreasing function* if

$$x_1 > x_2 \implies f(x_1) < f(x_2)$$

Concavity and convexity:

A function of one variable is *concave* if

$$f(tx + (1-t)x') \geq tf(x) + (1-t)f(x')$$

for all x and x' and all $0 \leq t \leq 1$.

A function of one variable is *strictly concave* if

$$f(tx + (1-t)x') > tf(x) + (1-t)f(x')$$

for all x and x' such that $x \neq x'$ and all $0 < t < 1$.

A function of one variable is *convex* if

$$f(tx + (1-t)x') \leq tf(x) + (1-t)f(x')$$

for all x and x' and all $0 \leq t \leq 1$.

A function of one variable is *strictly convex* if

$$f(tx + (1-t)x') < tf(x) + (1-t)f(x')$$

for all x and x' such that $x \neq x'$ and all $0 < t < 1$.

A function $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ is *homogeneous* of degree k if

$$f(tx) = t^k f(x) \text{ for all } t > 0.$$

Note that if we double all the arguments of a function that is homogeneous of degree zero ($k = 0$), the value of the function doesn't change.

A function $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ is *continuous* when "the graph can be drawn without lifting a pencil from the paper".

The slope

The slope of a linear function $f(x) = mx + b$ is the ratio $\frac{y_1 - y_0}{x_1 - x_0}$ where the points (x_0, y_0) and (x_1, y_1) , are any arbitrary points. Note that a linear function has a constant slope at all points along its graph. Moreover, $f(x) = mx + b$ the line has the slope m and its y -intersect is at $(0, b)$.

The slope of any function, linear or non-linear, is its derivative, written $f'(x_0)$ or $\frac{df(x)}{dx}$. Note that only the derivative of a linear function is the same for any x . On the other hand, a nonlinear function has the property that its slope changes as x changes. A *tangent* to a function at some point x_0 is a linear function that has the same slope.

Derivatives

If a function is *differentiable* at a point x_0 , the derivative of a function $y = f(x)$ at the point x_0 is defined to be

$$\frac{df(x_0)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

In words, the derivative is the rate of change of y with respect to x as the change in x goes to zero. The derivative gives the precise meaning to the phrase "the rate of change of y with respect to x for small changes in x ".

For any integer k the derivative of $f(x) = x^k$ at a point $x = x_0$ is $f'(x_0) = k(x_0)^{k-1}$. Now, assume that k is an arbitrary constant and f and g are continuous differentiable functions at $x = x_0$. Then,

$$\begin{aligned} (f \pm g)'(x_0) &= f'(x_0) \pm g'(x_0) \\ (kf)'(x_0) &= kf'(x_0) \\ (f \times g)'(x_0) &= f'(x_0)g(x_0) + g'(x_0)f(x_0) \\ (f/g)'(x_0) &= \frac{f'(x_0)g(x_0) - g'(x_0)f(x_0)}{g(x_0)^2} \\ ((f(x_0))^n)' &= n(f(x_0))^{n-1}f'(x) \\ (f \circ g)(x_0) &= f'(g(x_0))g'(x_0) \end{aligned}$$

Examples

(i) $f(x) = 5x^3$

(ii) $f(x) = 3x^{-3/2}$

$$(iii) f(x) = 3x^2 - 9x + 7x^{2/5}$$

$$(iv) f(x) = \frac{x+1}{x-1}$$

$$(v) f(x) = \frac{1}{2}\sqrt{x}$$

$$(vi) f(x) = (x^2 + 3x + 1)(x^3 - 2x^2 + 5)$$

A second derivative of a function of one variable is just the derivative of the derivative. It is denoted as follows

$$\frac{d^2 f(x)}{dx^2} = \frac{d \frac{df(x)}{dx}}{dx}$$

Examples

$$(i) f(x) = 5x^3$$

$$(ii) f(x) = 3x^{-3/2}$$

If a function has more than one argument, then we can take the *partial derivative*. This measures how the value of a function changes as one of its arguments changes, keeping all other arguments constant. Formally the partial derivative of the function $y = f(x_1, \dots, x_n)$ with respect to x_i at point $y = f(\bar{x}_1, \dots, \bar{x}_n)$ is defined as

$$\frac{\partial f(\bar{x}_1, \dots, \bar{x}_n)}{\partial x_i} = \lim_{\Delta x_i \rightarrow 0} \frac{f(\bar{x}_1, \dots, \bar{x}_i + \Delta x_i, \dots, \bar{x}_n) - f(\bar{x}_1, \dots, \bar{x}_n)}{\Delta x_i}$$

In calculating partial derivatives, the key is to remember to treat the other arguments as constants, then use the rules above

Examples

$$(i) f(x, w) = 4x^2w - 3xw^3 + 6x$$

$$(ii) f(x, w, z) = 3x^{-3/2}z^4 + w^2z^2$$

One important concept that uses the notion of partial derivatives is that of the total derivative, which relates changes in the arguments of a function to changes in the value of the function. For the differentiable function $y = f(x, z)$

$$dy = \frac{\partial f(x, z)}{\partial x} dx + \frac{\partial f(x, z)}{\partial z} dz$$

The Isoquant

An isoquant is like a contour on a map. It is the set of points in the domain of a function for which the function is equal to some constant value. There will be an isoquant associated with each value in the range of the function.

Formally, the isoquant associated with the element y in the range of the function $f(x_1, \dots, x_n)$ we will denote as $I(y)$

$$I(y) = \{x_1 \dots x_n : f(x_1, \dots, x_n) = y\}$$

We can use the total derivative to calculate the slope of the isoquant. We can do this because we know that, on the isoquant, the value of the function does not change. Consider a function with two arguments $f(x_1, x_2)$, and the isoquant $f(x_1, x_2) = \bar{y}$. We can write

$$d\bar{y} = \frac{\partial f(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial f(x_1, x_2)}{\partial x_2} dx_2 = 0$$

which we can rearrange to give

$$-\frac{\frac{\partial f(x_1, x_2)}{\partial x_1}}{\frac{\partial f(x_1, x_2)}{\partial x_2}} = \frac{dx_2}{dx_1}$$

13.4.2 Optimization

Say we have a function $f : \mathbb{R} \rightarrow \mathbb{R}$, such that $y = f(x)$. How can we find out which value of x will maximize y ? First we ask the question: what properties will such a maximum have? First some notation. We describe the above problem as *maximising y with respect to x* . If y^* is the maximum value of $f(x)$ then we write $y^* = \max_x f(x)$. If x^* is the value of x that maximises $f(x)$ we write $x^* = \arg \max_x f(x)$

Not all functions will have a maximum. For example, $f(x) = -x^2$ does have a maximum, while $f(x) = x$ does not (see figure 1). However, if the function has an interior maximum, then the slope of the function at that point will be zero (see figure 2). This is called the *first order condition* of the maximization problem

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function. If \bar{x} is the argmax of f , then

$$\frac{\delta f(\bar{x})}{\delta x} = 0$$

We have stated that if \bar{x} is the argmax of f the slope of f at \bar{x} must be 0. In other words, it is a *necessary condition* of \bar{x} being the argmax that the slope is zero at that point. Is it also true that, if the slope of f at some point

x^* is zero, then x^* is the argmax of f ? In other words, is this a *sufficient condition*? The answer is no. If the slope is zero, then the point we are looking at could be a maximum, minimum or neither - see figure 3. We can differentiate between these cases by looking at the *second derivative* of the function. This measures the 'slope of the slope' of the function.

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a twice-differentiable function. Let \bar{x} be a critical point of f , (i.e. $\frac{\delta f(\bar{x})}{\delta x} = 0$) then if:

$\frac{d^2 f(x)}{dx^2} > 0$ then \bar{x} is a minimum point

$\frac{d^2 f(x)}{dx^2} < 0$ then \bar{x} is a maximum point

$\frac{d^2 f(x)}{dx^2} = 0$ then \bar{x} can be a maximum point, minimum point or a point of inflection

Note that, for twice differentiable functions, $\frac{d^2 f(x)}{dx^2} > 0$ implies that the function is convex, and $\frac{d^2 f(x)}{dx^2} < 0$ implies that the function is concave (at that point). This leads to a refinement of theorem 1.

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a twice differentiable function if \bar{x} is the argmax of f , then

$$a: \quad \frac{\delta f(\bar{x})}{\delta x} = 0$$

$$b: \quad \frac{d^2 f(x)}{dx^2} \leq 0$$

Condition b is known as the *second order condition* of the maximization problem. This is a stronger set of necessary conditions, but are they now sufficient conditions? Again the answer is no. The reason is that the conditions are enough to find a maximum of the function, but not necessarily *the* maximum of the function. Figure 4 highlights the difference between a *local* and a *global* maximum.

To get sufficient conditions, we need something stronger. For example, we could demand that our candidate point is the only critical point.

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a twice differentiable function if there exists a point \bar{x} such that

$$a: \quad \frac{\delta f(\bar{x})}{\delta x} = 0$$

$$b: \quad \frac{d^2 f(x)}{dx^2} \leq 0$$

$$c: \quad \frac{\delta f(x)}{\delta x} \neq 0 \quad \forall \quad x \neq \bar{x}$$

Then \bar{x} is the global argmax of f

Unconstrained optimization - many variables

We can find analogous results to those above when maximizing a function of many variables. Just as in the single variable case, the slope of the function must be flat at an interior maximum. However, now it has to be flat with respect to all the arguments of the function. In other words, all the partial derivatives of the function will be zero at the interior maximum.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function. Let x_1, \dots, x_n be the argmax of f then

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_i}$$

Often in economics we will be interested in optimizing a function subject to some constraints. For example, we might want to know how a consumer can optimize utility within a given budget set, or how a firm can maximize output with given labour and capital inputs. While such problems are many and varied, they all have the same setup.

A constrained optimization problem is a problem of the following form:

Choose x_1, \dots, x_n *(arguments)*
 to maximize $f(x_1, \dots, x_n)$ *(objective function)*
 subject to $h_1(x_1, \dots, x_n) = 0$,
 \vdots
 $h_I(x_1, \dots, x_n) = 0$ *(equality constraints)*
 $g_1(x_1, \dots, x_n) \leq 0$,
 \vdots
 $g_J(x_1, \dots, x_n) \leq 0$ *(inequality constraints)*

How do we solve such problems? We can gain some intuition from looking at the graph of a simple problem. Say we have only two variables and one equality constraint. In other words, we want to solve the problem:

choose x_1, x_2 *(arguments)*
 to maximize $x_1 x_2$ *(objective function)*
 subject to $x_1 + x_2 - 1 = 0$.

Because this problem is in two dimensions, we can graph it, which we do in figure 5. The solid line shows the constraint, while the dotted line shows the isoquants of the objective function

It is clear that the maximum occurs at the point where the slope of the isoquant of the objective function is equal to the slope of the constraint function. As we know from class 1, we can write this in terms of partial

derivatives.

$$\frac{\frac{\partial f(x_1, x_2)}{\partial x_2}}{\frac{\partial f(x_1, x_2)}{\partial x_1}} = \frac{\frac{\partial h(x_1, x_2)}{\partial x_2}}{\frac{\partial h(x_1, x_2)}{\partial x_1}}$$

The trick we use to find these points is to rewrite the problem as a *Lagrangian Function*.

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu h(x_1, x_2)$$

Under the right conditions, finding the unconstrained optimum of this function is the same as solving the constrained optimum.

To solve the unconstrained optimization problem, we take first derivatives with respect to the three arguments and set them equal to zero:

$$\begin{aligned}\frac{\partial L(x_1, x_2, \mu)}{\partial x_1} &= \frac{\partial f(x_1, x_2)}{\partial x_1} - \mu \frac{\partial h(x_1, x_2)}{\partial x_1} = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial x_2} &= \frac{\partial f(x_1, x_2)}{\partial x_2} - \mu \frac{\partial h(x_1, x_2)}{\partial x_2} = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial \mu} &= h(x_1, x_2) = 0\end{aligned}$$

Where the third equation simply repeats our constraint. These are the *first order conditions* of the constrained optimization problem. We can see that this uses the finding we made above about the ratio of the partial derivatives. Taking the ratio of the first two first order conditions gives us the condition that we stated above.

In the case of our example, our first order conditions become

$$\begin{aligned}\frac{\partial L(x_1, x_2, \mu)}{\partial x_1} &= x_2 - \mu = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial x_2} &= x_1 - \mu = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial \mu} &= x_1 + x_2 - 1 = 0\end{aligned}$$

Using the first two conditions tells us that $x_2 = x_1$, and substituting this into the third condition gives $x_2 = x_1 = \frac{1}{2}$. This is the solution to the constrained optimization problem. If we have more constraints, we can add more Lagrange multipliers. So consider the problem:

Choose x_1, \dots, x_n (arguments)
to maximize $f(x_1, \dots, x_n)$ (objective function)
subject to $h_1(x_1, \dots, x_n) = 0$,
 $h_2(x_1, \dots, x_n) = 0$
 $h_3(x_1, \dots, x_n) = 0$ (equality constraints)

We would form the Lagrangian

$$L(x_1 \dots x_n, \mu_1, \mu_2, \mu_3) = f(x_1 \dots x_n) - \mu_1 h_1(x_1, \dots, x_n) - \mu_2 h_2(x_1, \dots, x_n) - \mu_3 h_3(x_1, \dots, x_n)$$

and have the first order conditions

$$\begin{aligned} \frac{\partial L(x_1 \dots x_n, \mu_1, \mu_2, \mu_3)}{\partial x_i} &= \frac{\partial f(x_1 \dots x_n)}{\partial x_i} - \mu_1 \frac{\partial h_1(x_1 \dots x_n)}{\partial x_i} - \mu_2 \frac{\partial h_2(x_1 \dots x_n)}{\partial x_i} + \\ &\quad - \mu_3 \frac{\partial h_3(x_1 \dots x_n)}{\partial x_i} = 0 \\ &\quad \forall i \in 1 \dots n \end{aligned}$$

$$\frac{\partial L(x_1 \dots x_n, \mu_1, \mu_2, \mu_3)}{\partial \mu_1} = h_1(x_1, x_2) = 0$$

$$\frac{\partial L(x_1 \dots x_n, \mu_1, \mu_2, \mu_3)}{\partial \mu_2} = h_2(x_1, x_2) = 0$$

$$\frac{\partial L(x_1 \dots x_n, \mu_1, \mu_2, \mu_3)}{\partial \mu_3} = h_3(x_1, x_2) = 0$$

This leads us to the following theorem.

Let $\bar{x}_1, \dots, \bar{x}_n$ maximize $f(x_1, \dots, x_n)$ subject to $h_1(x_1, \dots, x_n) = 0, \dots, h_j(x_1, \dots, x_n) = 0^1$. Then there exists a set of j numbers μ_1, \dots, μ_j such that $\{\bar{x}_1, \dots, \bar{x}_n, \mu_1, \dots, \mu_j\}$ is a critical point of the Lagrangian

$$L(x_1 \dots x_n, \mu_1, \dots, \mu_j) = f(x_1 \dots x_n) - \mu_1 h_1(x_1, \dots, x_n) - \dots - \mu_j h_j(x_1, \dots, x_n)$$

As in the case of unconstrained optimization, this is a necessary but not sufficient condition for finding the constrained optimum. We have similar worries about finding minima instead of maxima, and local rather than global extreme points. There are second order conditions for constrained optimization problems. However, we will not cover them here.

What about inequality constraints? This makes the problem more difficult, as we do not know in advance whether the constraint will bind or

¹This theorem should really have another condition, that $\bar{x}_1, \dots, \bar{x}_n$ satisfy the non-degenerate constraint qualification. However, we will skip over this for now.

not at the maximum. If the constraint does not bind, then it is not really a constraint at all. If it does bind, then we can treat it as we would an equality constraint. While there are techniques for dealing with inequality constraints, in some cases it may be obvious whether or not a constraint will bind. For the purposes of this course, we will be able to use these 'common sense' methods, rather than grinding through the tedious maths.

13.4.3 Solving the Cobb-Douglas Consumer's Problem

The second problem of section 1.4 of the second maths refresher homework is a classic example of a consumer maximization problem. You will see this problem, or a version of it, time and time again. It is therefore worth getting familiar with it now, as it will save you lots of time later on. We will also use it here to demonstrate the 5 steps that, quite often, will lead you to the solution of a constrained optimization problem. While its not advisable to think of maths in a cookbook fashion, you can treat this as such as a last resort. First, to restate the problem:

Maximize $x^\alpha y^{(1-\alpha)}$ subject to $p_x x + p_y y = I$. (Note that α, p_x, p_y, I are parameters for this problem your answer should be a function of x and y in terms of these values)

The interpretation of this problem is as follows: A consumer is has to choose between bundles of good x and good y in order to maximize her utility function $x^\alpha y^{(1-\alpha)}$. However, she only has a certain amount of income to spend, I . The price of the two goods is denoted by p_x and p_y . So the consumer is trying to choose the bundles to maximize her utility function subject to the bundles being those that she can afford. As stated above, α, p_x, p_y, I are what we call *parameters* of the problem. In other words, they are the 'rules of the game', or things that the consumer cannot alter. Your solution to the problem will be a function of these parameters.

The cookbook method of solving these problems has 5 steps, described below:

Form the Lagrangian function

This always takes the form

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu h(x_1, x_2)$$

Where μ is the lagrangian multiplier, $f(x_1, x_2)$ is the objective function and $h(x_1, x_2)$ is the constraint. In this case, we get:

$$L(x, y, \mu) = x^\alpha y^{(1-\alpha)} - \mu(p_x x + p_y y - I)$$

Take partial derivatives and set them equal to zero

I've tried to convince you that solving the constrained optimization problem is the same as solving the unconstrained problem for the lagrangian. If you are solving an unconstrained optimization problem, then the first thing we do is take the partial derivatives and set them equal to zero. In general:

$$\begin{aligned}\frac{\partial L(x_1, x_2, \mu)}{\partial x_1} &= \frac{\partial f(x_1, x_2)}{\partial x_1} - \mu \frac{\partial h(x_1, x_2)}{\partial x_1} = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial x_2} &= \frac{\partial f(x_1, x_2)}{\partial x_2} - \mu \frac{\partial h(x_1, x_2)}{\partial x_2} = 0 \\ \frac{\partial L(x_1, x_2, \mu)}{\partial \mu} &= h(x_1, x_2) = 0\end{aligned}$$

And in this specific case:

$$\begin{aligned}\frac{\partial L(x, y, \mu)}{\partial x} &= \alpha x^{\alpha-1} y^{1-\alpha} - \mu p_x = 0 \\ \frac{\partial L(x, y, \mu)}{\partial y} &= (1 - \alpha) x^\alpha y^{-\alpha} - \mu p_y = 0 \\ \frac{\partial L(x, y, \mu)}{\partial \mu} &= p_x x + p_y y - I = 0\end{aligned}$$

Divide the FOC from the choice variables by each other to get rid of μ

We can always rewrite our partial derivatives from the choice variables by taking one term over to the other side of the equality. In general:

$$\begin{aligned}\frac{\partial f(x_1, x_2)}{\partial x_1} - \mu \frac{\partial h(x_1, x_2)}{\partial x_1} = 0 &\Rightarrow \frac{\partial f(x_1, x_2)}{\partial x_1} = \mu \frac{\partial h(x_1, x_2)}{\partial x_1} \\ \frac{\partial f(x_1, x_2)}{\partial x_2} - \mu \frac{\partial h(x_1, x_2)}{\partial x_2} = 0 &\Rightarrow \frac{\partial f(x_1, x_2)}{\partial x_2} = \mu \frac{\partial h(x_1, x_2)}{\partial x_2}\end{aligned}$$

We can then divide these equations by each other, giving

$$\frac{\frac{\partial f(x_1, x_2)}{\partial x_1}}{\frac{\partial f(x_1, x_2)}{\partial x_2}} = \frac{\mu \frac{\partial h(x_1, x_2)}{\partial x_1}}{\mu \frac{\partial h(x_1, x_2)}{\partial x_2}} \Rightarrow \frac{\frac{\partial f(x_1, x_2)}{\partial x_1}}{\frac{\partial f(x_1, x_2)}{\partial x_2}} = \frac{\frac{\partial h(x_1, x_2)}{\partial x_1}}{\frac{\partial h(x_1, x_2)}{\partial x_2}}$$

This gets rid of the μ from the top and bottom of our equation.

In this specific case, we get

$$\begin{aligned} \alpha x^{\alpha-1} y^{1-\alpha} - \mu p_x &= 0 \Rightarrow \alpha x^{\alpha-1} y^{1-\alpha} = \mu p_x \\ (1-\alpha) x^\alpha y^{-\alpha} - \mu p_y &= 0 \Rightarrow (1-\alpha) x^\alpha y^{-\alpha} = \mu p_y \end{aligned}$$

And so

$$\frac{\alpha x^{\alpha-1} y^{1-\alpha}}{(1-\alpha) x^\alpha y^{-\alpha}} = \frac{\mu p_x}{\mu p_y} \Rightarrow \frac{\alpha x^{\alpha-1} y^{1-\alpha}}{(1-\alpha) x^\alpha y^{-\alpha}} = \frac{p_x}{p_y}$$

Solve for x in terms of y (or viceversa)

This equation should now be in terms only of x, y and the parameters of the model, as we have got rid of μ . In the case of Cobb-Douglas Utilities, something magical happens. Note that, on the left hand side of the equation, the powers of α on x and y cancel, leaving a linear model!

$$\begin{aligned} \frac{\alpha x^{\alpha-1} y^{1-\alpha}}{(1-\alpha) x^\alpha y^{-\alpha}} &= \frac{\alpha y}{(1-\alpha) x} \\ &\Rightarrow \frac{\alpha y}{(1-\alpha) x} = \frac{p_x}{p_y} \\ &\Rightarrow y = \frac{(1-\alpha) x p_x}{\alpha p_y} \end{aligned}$$

Substitute into the constraint and solve

At present we have two unknowns. That means that we have to find another equation so we can solve the system. Luckily, we have such an equation - the constraint. We can now substitute the above into this equation

$$\begin{aligned} p_x x + p_y y &= I \\ p_x x + p_y \left(\frac{(1-\alpha) x p_x}{\alpha p_y} \right) &= I \end{aligned}$$

but here we can cancel the p_y terms, to give

$$p_x x + \frac{(1 - \alpha)}{\alpha} p_x x = I$$

We can neaten this up a bit by multiplying the left most term top and bottom by α

$$\begin{aligned} \frac{\alpha}{\alpha} p_x x + \frac{(1 - \alpha)}{\alpha} p_x x &= I \\ \frac{\alpha p_x x + (1 - \alpha) p_x x}{\alpha} &= I \\ \frac{p_x x}{\alpha} &= I \\ x &= \frac{I \alpha}{p_x} \end{aligned}$$

Similar work will tell you that

$$y = \frac{I(1 - \alpha)}{p_y}$$

We now have an expression for x and y in terms only of the parameters, so we have solved the model. Victory!. Note that this can be rewritten as $p_x x = \alpha I$ where the left hand side of the equation is just the amount spent on good x . This tells us that, with these preferences, the amount spent on good x will be a constant fraction of income, whatever the price!

13.5 Linear algebra

Linear algebra is about the solution of

$$Ax = b$$

where $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ is an $(m \times n)$ matrix, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is an $(n \times 1)$ vector, and $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ is an $(m \times 1)$ vector.

The equation

$$Ax = b$$

can also be represented by the system of equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

and by the equation

$$a_1x_1 + \cdots + a_nx_n = b$$

where a_1, \dots, a_n are the columns of the matrix A .

13.5.1 Reduction to echelon

The augmented matrix $[A \mid b]$ can be reduced to echelon and reduced echelon form.

Every row $[0 \cdots 0 \mid 0]$ in the echelon form corresponds to a free variable.

A row $[0 \cdots 0 \mid b_m > 0]$ in the echelon form implies inconsistency.

13.5.2 Independent columns

The study of

$$Ax = b, \text{ for any } b$$

depends crucially on the number of linear independent column of the matrix A . Let $\text{span}(A)$ denote the set of possible vectors b such that $Ax = b$ for some vector x . Then $\text{span}(A)$ is a subspace of dimension equal to the number of independent columns of A . The number of independent columns of A cannot be larger than $\min\{m, n\}$.

i) If $\text{span}(A) = \mathbb{R}^m$, then

$$Ax = b$$

has a solution for any b . In this case,

- if $n = m$, the solution has no free variables and the linear transformation $x \rightarrow T(x)$ is *one-to-one* and *onto*; equivalently, $Ax = 0$ has only the trivial solution $x = 0$;
- if $n > m$, the solution has $n - m$ free variables and the linear transformation $x \rightarrow T(x)$ is *onto* but *NOT one-to-one*; equivalently, $Ax = 0$ has an infinity of solutions.

ii) If $\text{span}(A) = \mathbb{R}^k$, $k < m$, then

$$Ax = b$$

does not have a solution for some b - not enough independent variables to satisfy the m equations. In this case, the linear transformation $x \rightarrow T(x)$ is *NOT onto*, but

- if $k = n$ it is *one-to-one* (the solution has NO free variables); $Ax = 0$ has only the trivial solution $x = 0$;
- if $k < n$, it is *NOT one-to-one* (the solution has free variables); $Ax = 0$ has an infinity of solutions.

13.5.3 Inverse matrix

A square $(n \times n)$ matrix A has an inverse A^{-1} such that

$$A \cdot A^{-1} = I_n$$

if and only if $\text{span}(A) = \mathbb{R}^n$, that is, the matrix A has n independent columns.

If A has an inverse then, $Ax = b$ has a unique solution for any b , only the the trivial solution for $b = 0$; equivalently, the linear transformation $x \rightarrow T(x)$ is *one-to-one* and *onto*.

The general solution of

$$Ax = b$$

in this case is

$$x = A^{-1}b.$$

13.5.4 Determinant

A square $(n \times n)$ matrix A has an inverse A^{-1} if and only if $\det(A) \neq 0$.

13.6 Problems**13.6.1 Sets and numbers**

Let $A = \{1, 2, 3, \dots, 10\}$, $B = \{0, 0.5, \sqrt{2}, 2, 10\}$, \mathbb{R} be the real numbers, \mathbb{N} be the natural numbers, \mathbb{Z} be the integers and \mathbb{Q} be the rationals. In each case below, describe the content of the set C

1. $C = \{x \in A : x > 6\}$

This is all the numbers in set A that are strictly greater than 6. In other words $C = \{7, 8, 9, 10\}$

2. $C = A \cup B$

All the elements that are either in set A or set B (or both). $C = \{0, 0.5, 1, \sqrt{2}, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

3. $C = A \cap B$

All the elements that are in both set A and set B . $C = \{2, 10\}$

4. $C = A \setminus B$

All the elements that are in set A but not in set B . $C = \{1, 3, 4, 5, 6, 7, 8, 9\}$

5. $C = B \cap \mathbb{R}$

All the elements of set B that are also real numbers. As all the numbers that we will come across in this course are real numbers, everything in set B is also in \mathbb{R} . So $C = \{0, 0.5, \sqrt{2}, 2, 10\} = B$

6. $C = B \cap \mathbb{N}$

All the elements of set B that are also natural numbers. Remember that the natural numbers are $\mathbb{N} = \{1, 2, 3, \dots\}$ so $C = \{2, 10\}$

7. $C = B \cap \mathbb{Z}$

All the elements of set B that are also integers. Remember that the natural numbers are $\mathbb{Z} = \{\dots -3, -2, -1, 0, 1, 2, 3, \dots\}$ so $C = \{0, 2, 10\}$

8. $C = B \cap \mathbb{Q}$

All the elements of set B that are also rational numbers. Remember that the rational numbers are all the numbers that can be expressed as a ratio of an integer and a natural number, and that $\sqrt{2}$ is not a rational number but 0.5 is, as it is equal to $\frac{1}{2}$. So $C = \{0, 0.5, 2, 10\}$

9. $C = \mathbb{Z} \setminus \mathbb{N}$

All the numbers which are integers but not natural numbers. As the integers are defined as $\mathbb{Z} = \{\dots -3, -2, -1, 0, 1, 2, 3, \dots\}$ and the natural numbers as $\mathbb{N} = \{1, 2, 3, \dots\}$, $C = \{\dots -3, -2, -1, 0\}$, or all the integers which are not strictly positive.

10. $C = \mathbb{N} \setminus \mathbb{Z}$

All the natural numbers that are not integers. We can see from above that any natural must also be an integer. The set of natural numbers which are not integers is therefore empty. We write this as $C = \emptyset$

13.6.2 Derivatives

Take the first derivatives of the following functions

1. $f(x) = (x^2 + x)(x^3 - 3x^2)$

Product rule. Let $g(x) = (x^2 + x)$ and $h(x) = (x^3 - 3x^2)$. Then $g'(x) = (2x + 1)$ and $h'(x) = (3x^2 - 6x)$. The product rule tells us that if $f(x) = g(x)h(x)$ then $f'(x) = g'(x)h(x) + h'(x)g(x)$. In this case,

$$f'(x) = (2x + 1)(x^3 - 3x^2) + (3x^2 - 6x)(x^2 + x)$$

2. $f(x) = \frac{(x^2+x)}{(x^3-3x^2)}$

Quotient rule. Let $g(x) = (x^2 + x)$ and $h(x) = (x^3 - 3x^2)$. Then $g'(x) = (2x + 1)$ and $h'(x) = (3x^2 - 6x)$. The quotient rule tells us that if $f(x) = \frac{g(x)}{h(x)}$ then $f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{(h(x))^2}$. In this case

$$f'(x) = \frac{(2x+1)(x^3-3x^2) + (x^2+x)(3x^2-6x)}{(x^3-3x^2)^2}$$

3. $f(x) = (2x^2 + 3x)^{3/2}$

Chain rule. let $g(x) = (2x^2 + 3x)$ and $h(y) = y^{3/2}$ then $f(x) = h(g(x))$. Then $g'(x) = (4x + 3)$ and $h'(y) = \frac{3}{2}y^{1/2}$. By chain rule, $f'(x) = h'(g(x))g'(x)$ in this case

$$f'(x) = \frac{3}{2}(2x^2 + 3x)^{1/2}(4x + 3)$$

4. $f(x) = \log(x^2)$

Two different methods. First, by chain rule, let $g(x) = x^2$ and $h(y) = \log y$ so that $f(x) = h(g(x))$. Then $g'(x) = 2x$ and $h'(y) = \frac{1}{y}$. By chain rule, $f'(x) = h'(g(x))g'(x)$ in this case

$$f'(x) = \frac{1}{x^2}2x = \frac{2}{x}$$

Alternatively, we can always rewrite $\log x^2$ as $2 \log x$, the differential of which is $\frac{2}{x}$

5. $f(x) = e^{x^2}$

Again, chain rule. let $g(x) = x^2$ and $h(y) = e^y$ then $f(x) = h(g(x))$. Then $g'(x) = 2x$ and $h'(y) = e^y$. By chain rule, $f'(x) = h'(g(x))g'(x)$ in this case

$$f'(x) = e^{x^2}2x$$

Take the first and second derivatives of these functions

1. $f(x) = 6x^3 + 4x^2 + x$

First derivative $f'(x) = 18x^2 + 8x + 1$

Second derivative $f''(x) = 36x + 8$

2. $f(x) = \log(x)$

First derivative $f'(x) = \frac{1}{x}$

Second derivative $f''(x) = -\frac{1}{x^2}$

3. $f(x) = e^x x^2$

First derivative $f'(x) = 2xe^x + x^2e^x$ (product rule)

Second derivative $f''(x) = 2xe^x + 2e^x + 2xe^x + x^2e^x = (2 + 4x + x^2)e^x$

Take all the partial derivatives of the following functions

Two things to remember:

1: There are the same number of partial derivatives as there are arguments in a function. If we have $f(x, y, z)$, then we will have three partial derivatives

2: When we are taking the partial derivatives, we take all the other arguments as constant. So in the case of $f(x, y, z) = xyz + x^2y + z^3$, if we are taking the derivative with respect to x , we can write

$f(x, y, z) = ax + bx^2 + c$, where $a = yz$, $b = y$ and $c = z^3$. The derivative of this function is $\frac{\partial f}{\partial x} = a + 2bx = yz + 2yx$

1. $f(x, y, z) = xyz + x^2y + z^3$

$$\frac{\partial f}{\partial x} = yz + 2xy$$

$$\frac{\partial f}{\partial y} = xz + x^2$$

$$\frac{\partial f}{\partial z} = xy + 3z^2$$

2. $f(x, y) = xy(y^2 + x^2)$

$$= xy^3 + yx^3$$

$$\frac{\partial f}{\partial x} = y^3 + 3yx^2$$

$$\frac{\partial f}{\partial y} = 3xy^2 + x^3$$

3. $f(x, y) = x \log y$

$$\frac{\partial f}{\partial x} = \log y$$

$$\frac{\partial f}{\partial y} = \frac{x}{y}$$

For the following two functions, sketch a graph of the isoquants for $f(x_1, x_2)$ equals 1, 2 and 5 in x_1, x_2 space. Write down an equation for the slope of the isoquants

1. $f(x_1, x_2) = x_1 + \frac{x_2}{2}$

For the case of $f(x_1, x_2) = 1$, we are looking for the set of x_1, x_2 such that $x_1 + \frac{x_2}{2} = 1$. Rearranging this gives $x_2 = 2 - 2x_1$. This is clearly a linear function with a slope of -2 . We can also calculate the slope of the isoquant using the result that

$$-\frac{\frac{\partial f(x_1, x_2)}{\partial x_1}}{\frac{\partial f(x_1, x_2)}{\partial x_2}} = \frac{dx_2}{dx_1}$$

as $\frac{\partial f(x_1, x_2)}{\partial x_1} = 1$ and $\frac{\partial f(x_1, x_2)}{\partial x_2} = \frac{1}{2}$, this confirms our result. Note that the slope is constant along the graph, and doesn't vary between the isoquants 1, 2 and 5.

2. $f(x_1, x_2) = x_1 x_2$

For the case of $f(x_1, x_2) = 1$, we are looking for the set of points such that $x_1 x_2 = 1$. Rearranging, this gives $x_2 = \frac{1}{x_1}$. This shape is a hyperbola. While you may not know this, you should be able to work out that, as x_1 gets very large, x_2 gets very small (but not negative) and as x_1 gets close to zero, x_2 gets very large.

We can work out the slope of the isoquant either directly, or using the ratio of the partial derivatives. As $\frac{\partial f(x_1, x_2)}{\partial x_1} = x_2$ and $\frac{\partial f(x_1, x_2)}{\partial x_2} = x_1$, the above formula tells us that $\frac{dx_2}{dx_1} = -\frac{x_2}{x_1}$. But we can do more. Note that $x_2 = \frac{1}{x_1}$ in this case, so we can substitute back in to give $\frac{dx_2}{dx_1} = -\frac{1}{x_1^2}$. Note that the slope of the isoquant changes both along the curve and between isoquants.

Partial Derivatives - one more try

Take all the partial derivatives of the following functions

1. $f(x, y, z) = x^3 y^2 z + z^2 y + y^2 x$

$$f_x(x, y, z) = 3x^2 y^2 z + y^2$$

$$f_y(x, y, z) = 2x^3yz + z^2 + 2yx$$

$$f_z(x, y, z) = x^2y^2 + 2zy$$

$$2. f(x, y, z) = (xy + z)(x + y)^3$$

$$f_x(x, y, z) = y(x + y)^3 + 3(x + y)^2(xy + z)$$

$$f_y(x, y, z) = x(x + y)^3 + 3(x + y)^2(xy + z)$$

$$f_z(x, y, z) = (x + y)^3$$

$$3. f(x, y) = \log xy$$

There are two ways to solve this. One way is to note that

$$f(x, y) = \log xy = \log x + \log y$$

$$\implies f_x(x, y) = \frac{1}{x}$$

$$f_y(x, y) = \frac{1}{y}$$

Or by chain rule

$$f_x(x, y) = y \cdot \frac{1}{yx} = \frac{1}{x}$$

$$f_y(x, y) = x \cdot \frac{1}{yx} = \frac{1}{y}$$

$$4. f(x, y) = \log \frac{x}{y}$$

As above

$$f_x(x, y) = \frac{1}{x}$$

$$f_y(x, y) = -\frac{1}{y}$$

$$5. f(x, y) = (x^2 + (xy + y^2))^{\frac{3}{2}}$$

$$f_x(x, y) = \frac{3}{2}(x^2 + (xy + y^2))^{\frac{1}{2}}(2x + y)$$

$$f_y(x, y) = \frac{3}{2}(x^2 + (xy + y^2))^{\frac{1}{2}}(2y + x)$$

13.6.3 Maximization

We start with functions of one variable and we move to functions of two variables.

Maximizing functions of one variable

For each of the following functions:

Find the critical points (i.e. the points where the function is flat);

Categorize these points as local maxima or minima (if possible);

Explain whether any of these points are global maxima or minima.

Take the first derivatives of the following functions

1. $f(x) = -3x^2 + 9x + 8$

FOC:

$$f'(x) = -6x + 9 = 0$$

$$x = \frac{3}{2}$$

This is the only critical point

SOC:

$$f''(x) = -6 < 0$$

This is therefore a local maximum

As we only have one critical point, and it is a local maximum, then we know from theorem 4 from the Prelim handout 2 that we have a global maximum.

2. $f(x) = 7x^2 + 2x + 3$

FOC:

$$f'(x) = 14x + 2 = 0$$

$$x = -\frac{1}{7}$$

This is the only critical point

SOC:

$$f''(x) = 14 > 0$$

This is therefore a local minimum

As we only have one critical point, and it is a local minimum, then we know from theorem 4 from the Prelim handout 2 that we have a global minimum.

3. $f(x) = \frac{x^3}{3} + x^2 - 8x + 10$

FOC:

$$f'(x) = x^2 + 2x - 8 = (x + 4)(x - 2) = 0$$

$$x = -4, 2$$

There are two critical points

SOC:

$$f''(x) = 2x + 2$$

$$f''(-4) = -6 < 0$$

This is a local maximum

$$f''(2) = 6 > 0$$

This is therefore a local minimum

We can see that neither of these points are global extrema. The easiest way to see this is to note that, as x gets very large, $f(x)$ goes to $+\infty$ and as x gets very small, $f(x)$ goes to $-\infty$

Maximizing functions of two variables

For the following functions, find the critical points (i.e. the point where the function is flat with regard to all its arguments)

1. $f(x, y) = x^4 + x^2 - 6xy + 3y^2$

FOC

$$f_y(x, y) = -6x + 6y = 0$$

$$\Rightarrow x = y$$

$$f_x(x, y) = 4x^3 + 2x - 6y = 0$$

but as $x=y$

$$\Rightarrow 4x^3 + 2x - 6x = 0$$

$$\Rightarrow 4x(x^2 - 1) = 0$$

$$\Rightarrow x = 0, x = 1, x = -1$$

So critical points lie at

$$(0, 0), (1, 1), (-1, -1)$$

2. $f(x, y) = xy^2 + x^3y - xy$

FOC

$$f_x(x, y) = y^2 + 3x^2y - y = 0$$

$$\Rightarrow y(y + 3x^2 - 1) = 0$$

$$f_y(x, y) = 2xy + x^3 - x = 0$$

$$\Rightarrow x(2y + x^2 - 1) = 0$$

The FOC with respect to x tells us that *either* $y = 0$ *or* $y + 3x^2 - 1 = 0$. Similarly, the second equation tells us that either $x = 0$ or $2y + x^2 - 1 = 0$.

This gives us 4 different cases

(i) $x = 0, y = 0$

(ii) $y = 0, 2y + x^2 - 1 = 0$

$$\Rightarrow x^2 = 1$$

$$\Rightarrow x = 1, x = -1$$

(iii) $y + 3x^2 - 1 = 0, x = 0$

$$\Rightarrow y = 1$$

(iv) $y + 3x^2 - 1 = 0, 2y + x^2 - 1 = 0$

$$\Rightarrow x^2 = 1 - 2y$$

$$\Rightarrow y + 3(1 - 2y) - 1 = 0$$

$$\Rightarrow 5y = 2 \Rightarrow y = \frac{2}{5}$$

$$\Rightarrow x^2 = 1 - 2 \cdot \frac{2}{5} = \frac{1}{5}$$

$$\Rightarrow x = \frac{1}{\sqrt{5}}, x = -\frac{1}{\sqrt{5}}$$

So, collecting all the different cases, we get the following critical points

$$(0, 0), (-1, 1), (1, 1), (0, 1), \left(-\frac{1}{\sqrt{5}}, \frac{2}{5}\right), \left(\frac{1}{\sqrt{5}}, \frac{2}{5}\right)$$

Constrained optimization

In the following two problems there is only one critical point, which I shall tell you is the maximum. Find it.

1. Maximize $3 \log x + 2 \log y$ subject to $3x + 4y = 10$

(i) Form the Lagrangian

$$L(x, y, \mu) = 3 \log x + 2 \log y - \mu(3x + 4y - 10)$$

(ii) Take the first order conditions

$$L_x(x, y, \mu) = \frac{3}{x} - 3\mu = 0 \Rightarrow \frac{3}{x} = 3\mu$$

$$L_y(x, y, \mu) = \frac{2}{y} - 4\mu = 0 \Rightarrow \frac{2}{y} = 4\mu$$

$$L_\mu(x, y, \mu) = 3x + 4y - 10 = 0$$

(iii) Divide the x, y FOC by each other to get rid of μ

$$\frac{\frac{3}{x}}{\frac{2}{y}} = \frac{3\mu}{4\mu} \Rightarrow \frac{3y}{2x} = \frac{3}{4}$$

(iv) Get x in terms of y (or visa versa)

$$\frac{3y}{2x} = \frac{3}{4} \Rightarrow y = \frac{x}{2}$$

(v) Substitute into the constraint and solve

$$3x + 4\left(\frac{x}{2}\right) = 10 \Rightarrow 5x = 10 \Rightarrow x = 2 \Rightarrow y = 1$$

2. Maximize $x^\alpha y^{(1-\alpha)}$ subject to $p_x x + p_y y = I$. (Note that a, p_x, p_y, I are parameters for this problem your answer should be a function of x and y in terms of these values)

See separate sheet

If you use the Lagrangian method on the following problem, you will get two solutions. One will be a maximum and the other will be a minimum. Find them both and decide which is which.

First, a quick apology - there are actually 4 solutions to this problem, but as we shall see below, they come in two very similar pairs of two.

1. Maximize $x^2 + y^2$ subject to $x^2 + xy + y^2 = 3$

(i) Form the Lagrangian

$$L(x, y, \mu) = x^2 + y^2 - \mu(x^2 + xy + y^2 - 3)$$

(ii) Take the first order conditions

$$L_x(x, y, \mu) = 2x - \mu(2x + y) = 0 \Rightarrow 2x = \mu(2x + y)$$

$$L_y(x, y, \mu) = 2y - \mu(2y + x) = 0 \Rightarrow 2y = \mu(2y + x)$$

$$L_{\mu}(x, y, \mu) = x^2 + xy + y^2 - 3 = 0$$

(iii) Divide the x, y FOC by each other to get rid of μ

$$\frac{2x}{2y} = \frac{\mu(2x+y)}{\mu(2x+y)} \Rightarrow \frac{x}{y} = \frac{(2x+y)}{(2y+x)}$$

(iv) Get x in terms of y (or visa versa)

$$\frac{x}{y} = \frac{(2x+y)}{(2y+x)} \Rightarrow x(2y+x) = y(2x+y) \Rightarrow 2xy + x^2 = 2xy + y^2$$

$$\Rightarrow x^2 = y^2 \Rightarrow x = y \text{ or } x = -y$$

(v) Substitute into the constraint and solve

Case 1: $x = y$

$$x^2 + xy + y^2 = 3 \Rightarrow x^2 + x^2 + x^2 = 3 \Rightarrow 3x^2 = 3 \Rightarrow x^2 = 1$$

Which gives 2 solutions: $(1, 1)$ and $(-1, -1)$

Case 2: $x = -y$

$$x^2 + xy + y^2 = 3 \Rightarrow x^2 - x^2 + x^2 = 3 \Rightarrow x^2 = 3$$

Which gives 2 solutions: $(-\sqrt{3}, \sqrt{3})$ and $(\sqrt{3}, -\sqrt{3})$

To find out which of these are the max and which are the minima, substitute the solutions back into the objective function. This tells you that the two solutions from case 1 give you $f(x, y) = x^2 + y^2 = 1 + 1 = 2$, whereas the solutions from case 2 are $(x, y) = x^2 + y^2 = 3 + 3 = 6$. The two solutions from case 1 are therefore minima whereas the two solutions from case 2 are maxima.

13.6.4 Linear algebra

[... To be added...]