Tomasz Żylicz

Advanced Microeconomics for QF students

Tuesday, 16:45-18:20
(30-hour lecture in Winter semester 2020/2021)

My office hours: Monday and Thursday, 8:45-9:30

My web page: http://coin.wne.uw.edu.pl/tzylicz/

<u>Microeconomics</u> = economy seen from the point of view of individual decision-makers

<u>Macroeconomics</u> = economy seen from the point of view of relationships between such aggregates as: GDP, exchange rate, unemployment, growth rate etc.

Both microeconomics and macroeconomics study entire economies

The distinction between microeconomics and macroeconomics is in a perspective adopted – not in an area studied

Abstraction as a methodological approach to economics in general and microeconomics in particular

Homo oeconomicus – a convenient abstraction which does not preclude spontaneity, altruism, etc.

Optimization – a convenient behavioural assumption (not necessarily an empirical fact)

Economic agents maximize something (profit, utility, market share, satisfaction, 'power' etc.) or minimize something (costs, disutility, risk, etc.)

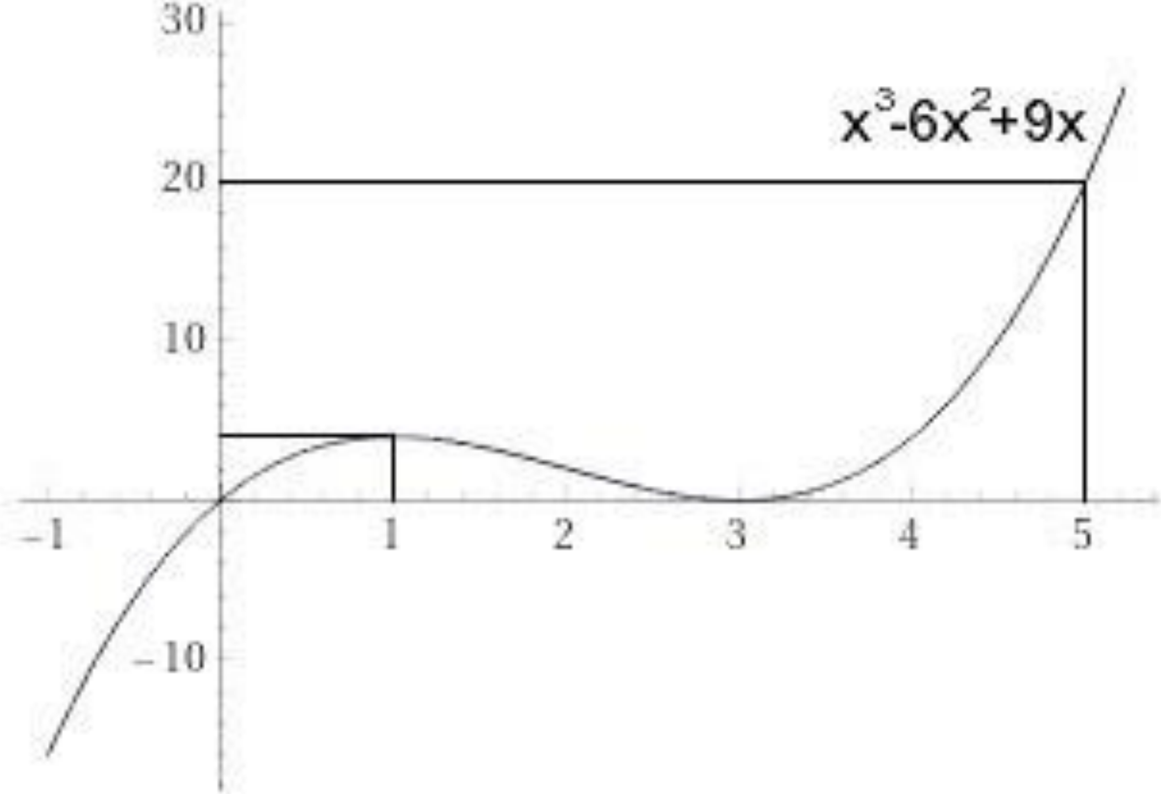
A typical problem indicates external constraints as well (production capacity, availability of credit, budget, etc.)

Homo oeconomicus

- 'Economic man'
- A convenient abstraction
- People make choices that let them achieve well-being
- Does not preclude spontaneity, altruism, etc.

Constrained maximisation

Maximize $\{x^3-6x^2+9x$: subject to $x-5\leq 0\}$



Mathematical representation of a microeconomic problem:

$$\text{Maximize } \{f(\mathbf{x}): \text{ subject to } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$$

Standard assumptions:

- $\mathbf{x} \in \Re^n$
- $f: \Re^n \rightarrow \Re$ – concave,
- $\mathbf{g}: \Re^n \rightarrow \Re^k$ – convex.

Then $\{\mathbf{x} \in \Re^n: \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$ is a convex set, and any local maximum of f in this set is a global one

Lagrange function ('Lagrangean') is defined as:

$$L(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x})-\boldsymbol{\lambda}\mathbf{g}(\mathbf{x}) = f(\mathbf{x})-(\lambda_1 g_1(\mathbf{x})+\ldots+\lambda_k g_k(\mathbf{x}))$$

Under the standard assumptions, economists calculate so-called 'First Order Conditions' (FOC), i.e. conditions for the derivatives of the Lagrange function to vanish:

$$\partial L/\partial \mathbf{x}=\mathbf{0},$$

i.e.

$$\partial f/\partial x_1-(\lambda_1\partial g_1(\mathbf{x})/\partial x_1+\ldots+\lambda_k\partial g_k(\mathbf{x})/\partial x_1)=0,$$

...

$$\partial f/\partial x_n-(\lambda_1\partial g_1(\mathbf{x})/\partial x_n+\ldots+\lambda_k\partial g_k(\mathbf{x})/\partial x_n)=0,$$

Partial derivatives of L with respect to **λ** return respective coordinates of the original 'constraint' function **g**

The result of optimization yields an equilibrium, i.e. an outcome that cannot be improved by the optimizing agent

Example:

Maximize $\{x^3-6x^2+9x: \text{ subject to } x-5\leq 0\}$

- $L(x,\lambda) = f(x)-\lambda g(x) = x^3-6x^2+9x-\lambda(x-5)$
- $x\in\Re$ and $\lambda\in\Re_+$ (they are numbers, not vectors)
- $\partial L/\partial x=3x^2-12x+9-\lambda$, and $\partial L/\partial\lambda=x-5$.
- $\partial L/\partial x=0 \Leftrightarrow 3x^2-12x+9-\lambda=0$

  - Case I: $\lambda=0 \Rightarrow 3x^2-12x+9=0$ – wrong.
  - Case II: $\lambda>0 \Rightarrow x-5=0$.

$\Rightarrow x^*=5$ and $\lambda^*=24$.

In typical economic problems, there is also a constraint:

$$\mathbf{x} \geq \mathbf{0}.$$

This – in the form of $-\mathbf{x} \leq \mathbf{0}$ – can be included in $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. Nevertheless economists prefer to deal with this additional constraint by applying the so-called Kuhn-Tucker theorem (see QF-15).

Equilibrium is merely a reference point. Microeconomists do not claim that economies are in equilibrium; they may be in dis-equilibrium. Nevertheless microeconomists identify forces which motivate economic agents to undertake certain actions

Questions:

Q-1 Unlike macroeconomics, microeconomics
[a] does not analyse the entire economy
[b] studies decisions of individual economic agents
[c] applies mathematics to derive equilibrium conditions
[d] studies decisions of small firms only
[e] none of the above

Exercises:

E-1 Prove that a strictly concave function cannot have two different local maxima over a convex set.

Consumer's choice

Consumption bundle – a collection of goods $\mathbf{x}=(x_1,\ldots,x_n)$ taken out of a consumption set X: $\mathbf{x}\in X\subset\Re^n$. Normally it is also assumed that $\mathbf{x}\geq\mathbf{0}$.

Preference relation (for a given consumer): a relation ≥ defined on the consumption set X

Definition: Relation ≥ is rational if

1. $\forall\ \mathbf{x},\mathbf{y}\in X\ [\mathbf{x}\geq\mathbf{y} \vee \mathbf{y}\geq\mathbf{x}]$ (completeness); and
2. $\forall\ \mathbf{x},\mathbf{y},\mathbf{z}\in X\ [(\mathbf{x}\geq\mathbf{y}\ \&\ \mathbf{y}\geq\mathbf{z}) \Rightarrow \mathbf{x}\geq\mathbf{z}]$ (transitivity).

Note: Preference (and strict preference) relations use the same symbols as arithmetic relations (greater than, and greater than or equal to). This, however, shall not lead to any doubts, since it will always be clear from the context whether the formula "$x \geq y$" means "x is preferred over y" or "x is greater than or equal to y". If x and y are consumption alternatives (bundles), i.e. $x, y \in X$ then the formula reads "x is preferred over y", but if x and y are numbers then the formula reads "x is greater than or equal to y".

Strict preference, >:

$$\mathbf{x} > \mathbf{y} \Leftrightarrow (\mathbf{x} \geq \mathbf{y} \;\&\; \neg \mathbf{y} \geq \mathbf{x})$$

Indifference, ~:

$$\mathbf{x} \sim \mathbf{y} \Leftrightarrow (\mathbf{x} \geq \mathbf{y} \;\&\; \mathbf{y} \geq \mathbf{x})$$

Theorem: If ≥ is rational, then:

1. $\forall$ **x**$\in$X [**x**$\geq$**x**] (reflexivity of $\geq$)
2. $\forall$ **x**$\in$X [$\neg$**x>x**] (counter-reflexivity of >)
3. $\forall$ **x**,**y**,**z**$\in$X [(**x**>**y** & **y**>**z**) $\Rightarrow$ **x**>**z**] (transitivity of >)
4. $\forall$ **x**$\in$X [**x**~**x**] (reflexivity of ~)
5. $\forall$ **x**,**y**,**z**$\in$X [(**x**~**y** & **y**~**z**) $\Rightarrow$ **x**~**z**] (transitivity of ~)
6. $\forall$ **x**,**y**$\in$X [**x**~**y** $\Rightarrow$ **y**~**x**] (symmetry of ~)
7. $\forall$ **x**,**y**,**z**$\in$X [(**x**>**y** & **y**$\geq$**z**) $\Rightarrow$ **x**>**z**]

Note:

4–6 (and completeness) imply that ~ is an equivalence on X

A <u>utility function</u> $u:X \rightarrow \Re$ representing $\geq$ is any function such that $\forall$ $\mathbf{x},\mathbf{y} \in X$ $[\mathbf{x} \geq \mathbf{y} \Leftrightarrow u(\mathbf{x}) \geq u(\mathbf{y})]$

<u>Theorem</u>: If there is a utility function representing $\geq$, then $\geq$ is rational

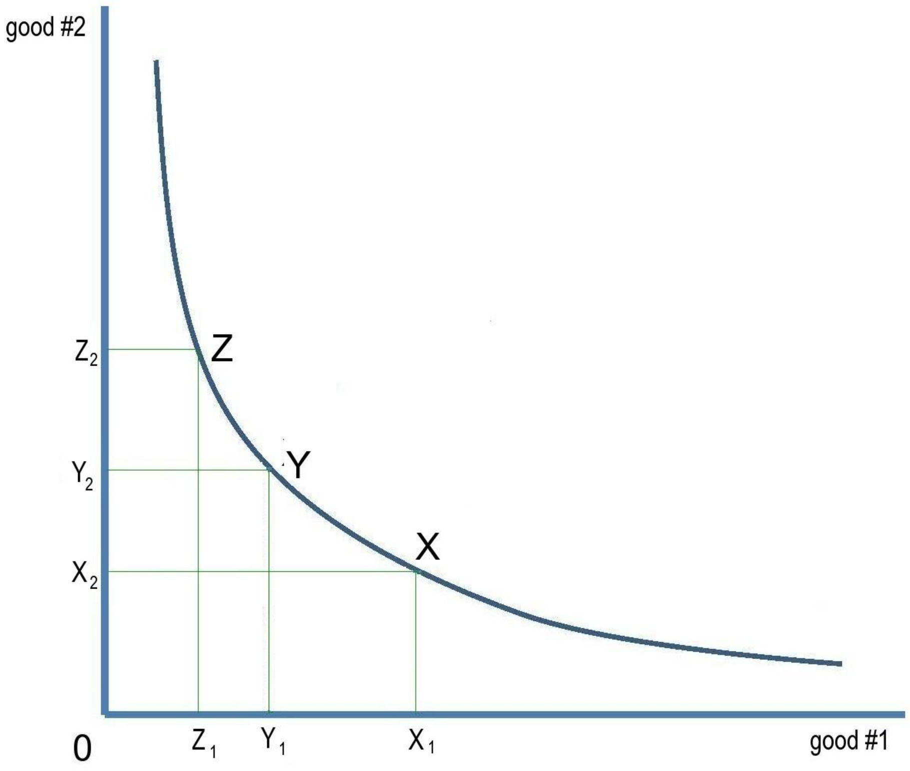
<u>Theorem</u>: If u is a utility function representing ≥, and f is a strictly increasing function, then f(u) is also a utility function representing ≥

Indifference curve $I(\mathbf{x})=\{\mathbf{y}\in X: \mathbf{y}\sim\mathbf{x}\}$

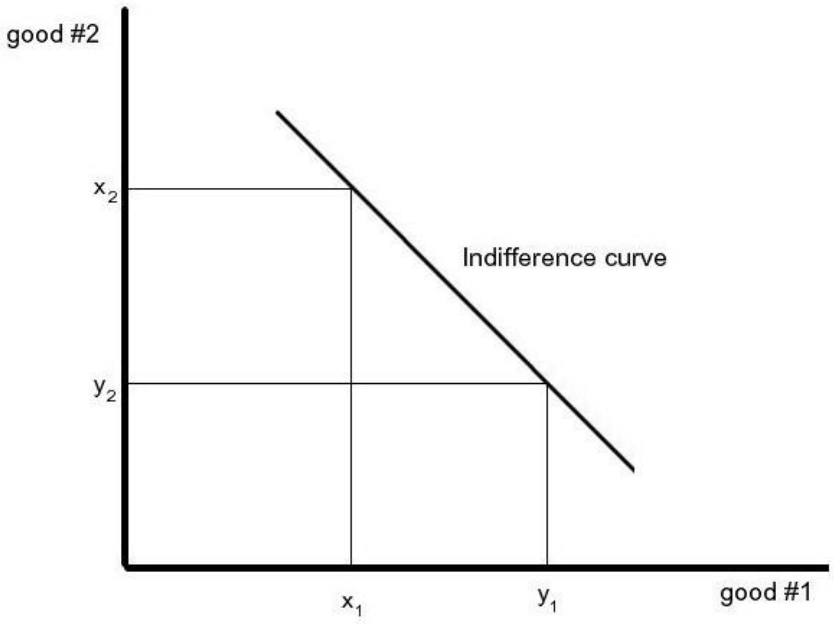
Theorem: If ≥ is rational, then every two indifference curves are either identical or disjoint

Definition. Let n=2. Perfect substitutes: each indifference curve is a segment of a straight line. Perfect complements: each indifference curve is L-shaped.

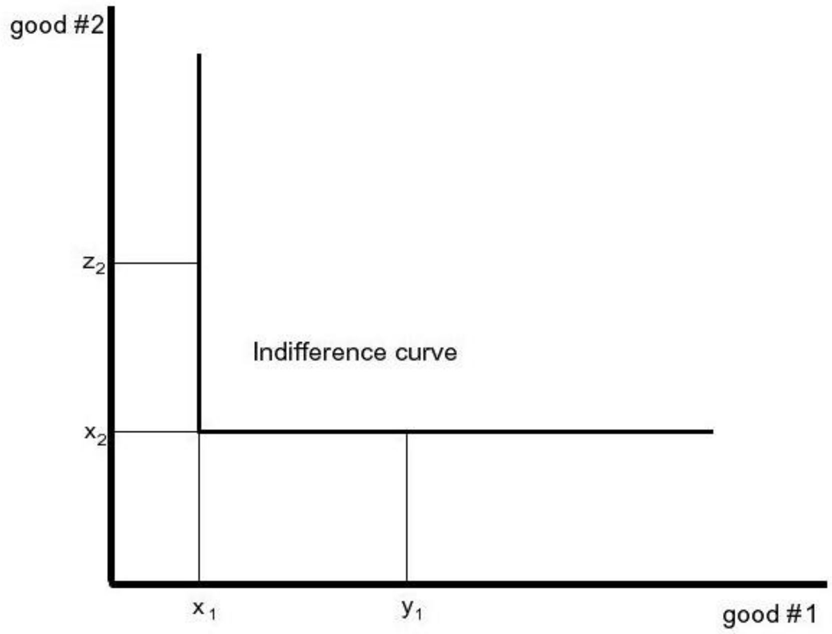
Indifference curves (1)



Indifference curves (2)



Perfect substitutes



Perfect complements

Preference axioms

- Monotonicity: If $\mathbf{x}=(x_1,\ldots,x_n)$, $\mathbf{y}=(y_1,x_2,\ldots,x_n)$, and $x_1 \geq y_1$, then $\mathbf{x} \geq \mathbf{y}$; likewise for goods 2, 3, ..., n
- Convexity: If $I(\mathbf{x})=I(\mathbf{z})$, and $\lambda \in [0,1]$, then $\lambda\mathbf{x}+(1-\lambda)\mathbf{z} \geq \mathbf{x}$ (indifference curves for bundles consisting of two goods are graphs of convex functions)
- Continuity: $\forall k=1,2,\ldots \forall \mathbf{x}_k,\mathbf{y}_k \in \Re^n$ $[(\mathbf{x}_k \geq \mathbf{y}_k \;\&\; \mathbf{x}=\lim_{n\to\infty}\mathbf{x}_n \;\&\; \mathbf{y}=\lim_{n\to\infty}\mathbf{y}_n) \Rightarrow \mathbf{x} \geq \mathbf{y}]$

Theorem. Let u be a utility function representing ≥. If **x**∈I(**y**) then u(**x**)=u(**y**)

Marginal Rate of Substitution (MRS): Coefficient of the tangent line to the indifference curve

Theorem. Let u be a differentiable utility function representing ≥. Then

$$MRS=-\partial u(x_1,x_2)/\partial x_1:\partial u(x_1,x_2)/\partial x_2.$$

An easy proof makes use of symbols "dy" and "dx".

Note. Economists (and engineers) apply the symbol "dy" in order to denote so-called differential of y. This notation was widely used by 17th and 18th century mathematicians who laid foundations for differential calculus. Later on it was used only in symbols such as "dy/dx", and examples were provided to prove that de-coupling "dy" from "dx" can lead to a nonsense. Indeed it can. Nevertheless, when applied with caution, it can simplify formulae. For practical purposes, in order to calculate df(x), one needs to calculate df(x)/dx=g(x) and "multiply" both sides of the equation by dx.

Proof:

$$du=(\partial u(x_1,x_2)/\partial x_1)dx_1+(\partial u(x_1,x_2)/\partial x_2)dx_2=0$$

$$(\partial u(x_1,x_2)/\partial x_1)dx_1=-(\partial u(x_1,x_2)/\partial x_2)dx_2$$

$$(\partial u(x_1,x_2)/\partial x_1)=-(\partial u(x_1,x_2)/\partial x_2)dx_2/dx_1$$

"Divide" this equation into $-(\partial u(x_1,x_2)/\partial x_2)$

Questions:

Q-2 Preferences are not rational when

[a] a consumer cannot determine whether two given bundles belong to the same indifference curve

[b] a bundle comprising more units of a good belongs to the same indifference curve that the original bundle

[c] a bundle comprising less units of a good belongs to the same indifference curve that the original bundle

[d] non-integer numbers of units are not considered "'legitimate" bundles

[e] none of the above

Exercises:

E-2 Can there be a preference relation such that any two bundles considered indifferent are identical? Please explain.

Choice under budgetary constraint

Let **p**=(p_1,...,p_n) be the vector of prices of goods from the consumer's bundle and let m be the consumer's income (money to spent on the bundle). Then the consumer's budgetary constraint is:

- $p_1x_1+\ldots+p_nx_n \leq m$, and
- $p_1x_1+\ldots+p_nx_n = m$

is called budget line. The set

- $\{\mathbf{x} \in X:\ p_1x_1+\ldots+p_nx_n \leq m\}$

is called budget set.

Consumer's choice under budgetary constraint is to:

- select a bundle in the budget set located at the highest indifference curve; or
- select a bundle in the budget set yielding the highest utility.

Both approaches are equivalent (if a utility function exists).

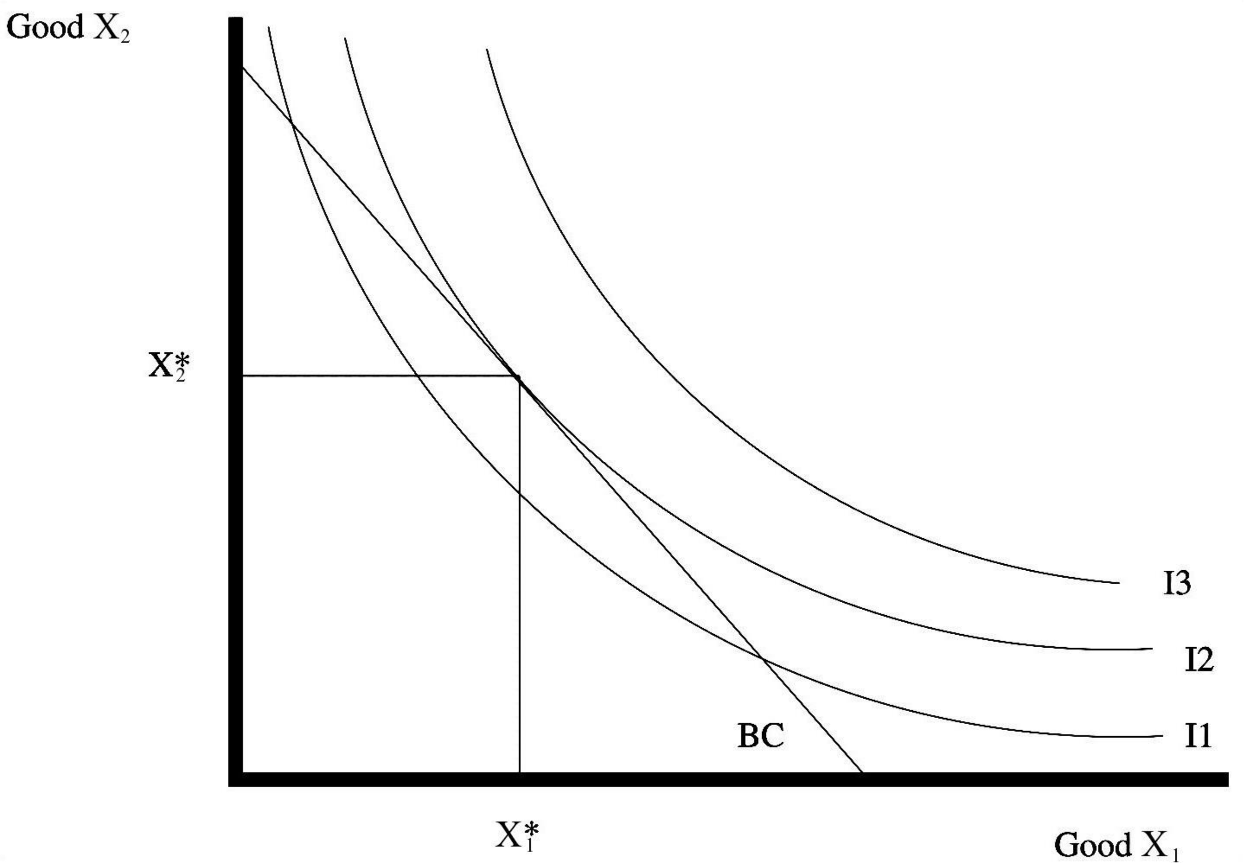
Utility Maximization Problem (UMP) is solving the problem of consumer's choice under budgetary constraint

Theorem: If u is a continuous function, and all prices are positive ($p_i>0$ for $i=1,\ldots,n$) then UMP has a solution $(x_1^*,\ldots,x_n^*)=\mathbf{x}^*(\mathbf{p},m)$

A solution $\mathbf{x}^*(\mathbf{p},m)$ to an UMP is called (primary) demand, also Marshallian or Walrasian demand.

Note: $\mathbf{x}^*(\mathbf{p},m)$ may not be unique

Consumer choice



Indirect utility is the utility of a bundle which solves an UMP. It is denoted by $v(\mathbf{p},m)$. By definition, $v(\mathbf{p},m)=u(\mathbf{x}^*(\mathbf{p},m))$; if $\mathbf{x}^*(\mathbf{p},m)$ is not unique then $u(\mathbf{x}^*(\mathbf{p},m))$ is understood as the utility of any bundle which solves the UMP.

Roy's identity:

$$x_i^*(\mathbf{p},m)=-(\partial v(\mathbf{p},m)/\partial p_i):(\partial v(\mathbf{p},m)/\partial m)$$

(can be proved for a continuous utility function representing locally non-satiated and strictly convex preferences)

Measures of welfare changes

Gross consumer surplus is the sum or integral of marginal benefits (measured by reservation prices) of the goods consumed. (Net) consumer surplus (CS) is the gross consumer surplus after subtracting the cost of purchasing the bundle.

Note: Calculating CS for welfare changes caused by price changes is difficult since – in general – changing prices imply changing consumer's relative wealth and thus his or her optimum choices

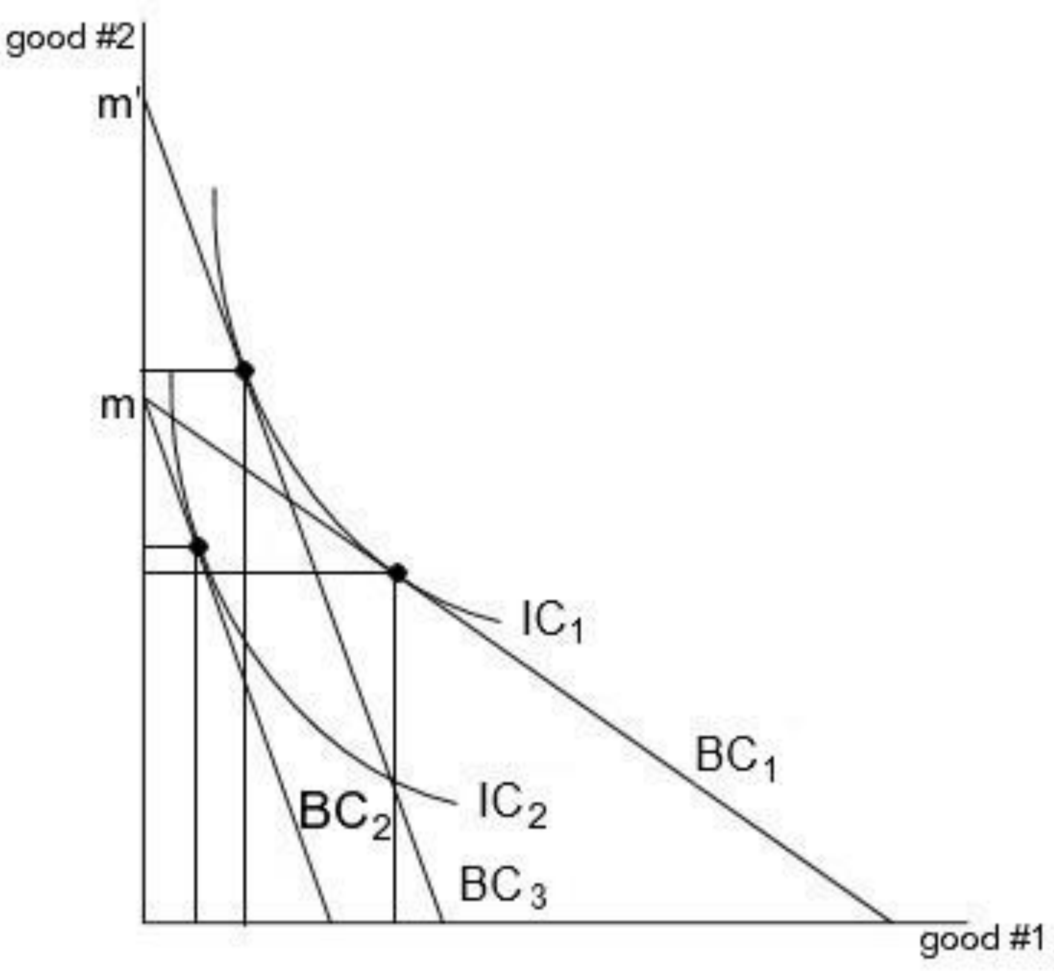
Money metric utility. Let n=2, and let the good i=2 represents aggregate consumption of everything except for the good i=1. This aggregate consumption is measured in money (and thus $p_2=1$). Then x_2 in utility $u(x_1,x_2)$ can be interpreted as the income left for purchasing all goods except for x_1 (once the good i=1 has been purchased), i.e. $x_2=m-p_1x_1$. If we assume that $p_2=1$, then consumer's choice depends on the price of good i=1 only. The price of this good will be denoted by p (without a subscript).

Compensating Variation (CV) implied by changing prices from $\mathbf{p^0}$ to $\mathbf{p^1}$ reflects the value of income change (decrease if $p^0>p^1$ or increase if $p^0<p^1$) necessary in order to keep the money metric utility at the original level despite the price change. CV is computed from the equation CV=m-m', where m (the old original income) and m' (the hypothetical new income) satisfy the equation:

$$u(x_1(p^0,m),m-p^0x_1(p^0,m)) = u(x_1(p^1,m'),m'-p^1x_1(p^1,m')).$$

In other words, CV informs to what extent the price change modifies the *de facto* income (-CV is the minimum amount the consumer should be paid in order to voluntarily accept such a price change; if CV>0, the consumer should be willing to pay for the change).

CV illustrated

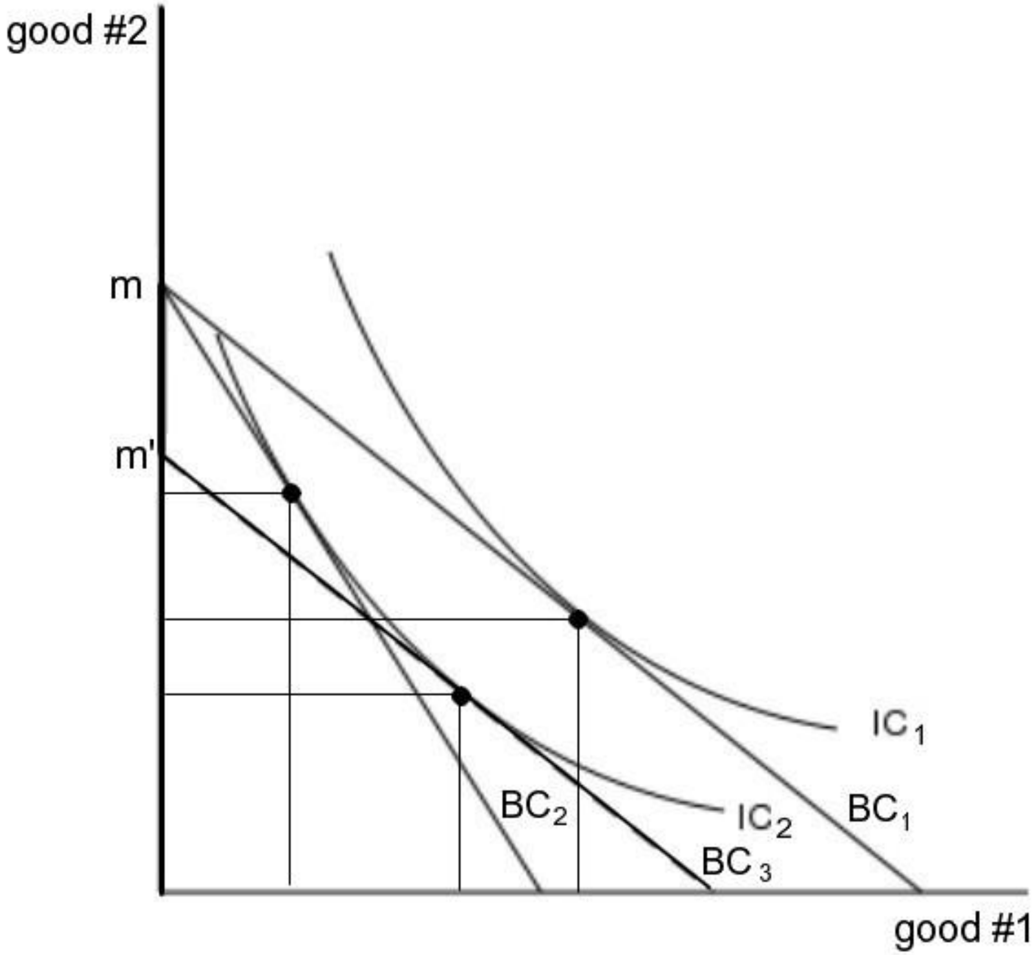


Equivalent Variation (EV) implied by changing prices from $\mathbf{p^0}$ to $\mathbf{p^1}$ reflects the value of income change (increase if $p^0 > p^1$ or decrease if $p^0 < p^1$) necessary in order to bring the money metric utility from the initial level to the final one without the price change. EV is computed from the equation EV=m'-m, where m (the old original income) and m' (the hypothetical new income) satisfy the equation:

$$u(x_1(p^1,m),m-p^1x_1(p^1,m)) = u(x_1(p^0,m'),m'-p^0x_1(p^0,m'))$$

In other words, EV informs to what extent the price change could be substituted by modifying the consumer's income (–EV is the maximum amount the consumer would be willing to pay to avoid the price change; if EV>0, the consumer should be paid to avoid this voluntarily).

EV illustrated

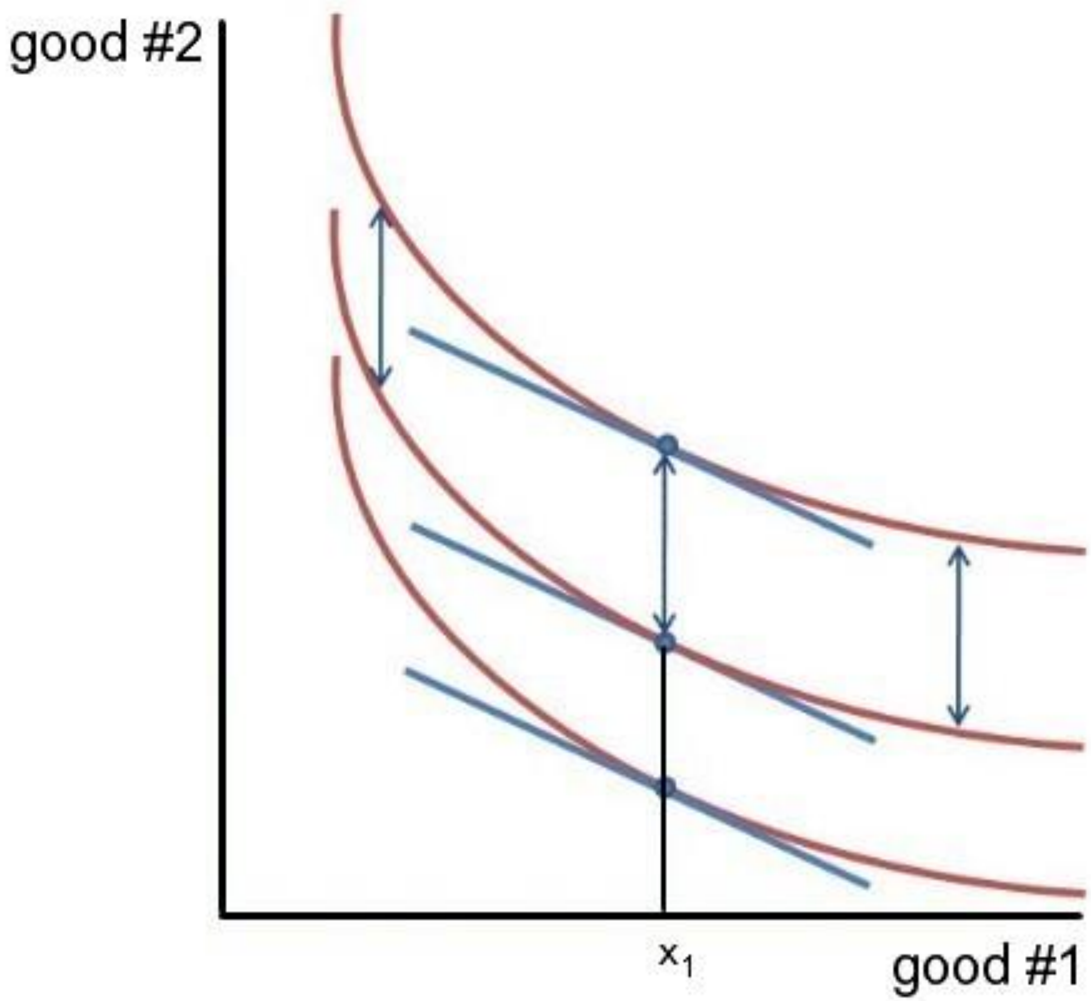


A money-metric utility function u is called <u>quasi-linear</u> with respect to the good number 1, if there exists a function v such that $u(x_1,x_2)=v(x_1)+x_2$

<u>Corollary</u>

Isoquants of a quasi-linear utility function have identical shape and they are simply translations of any given one along the line x_1=const.

Quasi-linear utilities



Note

For a quasi-linear utility, if the demand for the good number 1 is derived from the FOC, then it depends only on the price (not on the income)

Theorem

For a quasi-linear utility CV and EV are identical and they are equal to:

$$CV = EV =$$
$$v(x_1(p^1,m)) - v(x_1(p^0,m)) - (p^1 x_1(p^1,m) - p^0 x_1(p^0,m))$$

Questions:

Q-3 Money-metric utility

[a] assumes that both goods considered are perfect substitutes

[b] assumes that both goods considered are perfect complements

<u>[c] assumes that the utility from consuming good number 1 is measured in money</u>

[d] assumes that the utility from consuming good number 2 does not depend on money spent on this good

[e] none of the above

Exercises:

E-3 Please provide an example that in the statement "for a quasi-linear utility, if the demand for the good number 1 is derived from the FOC, then it depends only on the price (not on the income)" the assumption on FOC cannot be omitted.

Intertemporal choice

Discount rate (r) lets compare money amounts that belong to different time periods

$$X_t = X_0(1+r)^t, \text{ or } X_0 = X_t/(1+r)^t, \text{ where}$$

X_t is the present value (in year t) of the value X_0 observed in year 0; or X_0 is the present value (in year 0) of the value X_t observed in year t.

Intertemporal choice (interpretations)

A discount rate r can be understood as a measure to compare some amounts in two distinct time moments. This is not necessarily the same as an interest rate paid on deposits. Nevertheless interest rates accepted by depositors inform about discount rates used by them.

In many applications one is interested in "net present value" rather than simply "present value", where "net" refers to costs being subtracted. In what follows X_0, X_1, X_2, X_T, etc. are understood as "net" values (i.e. after the subtraction of whatever costs):

$$NPV = X_0/(1+r)^0+X_1/(1+r)^1+X_2/(1+r)^2+ \ldots +X_T/(1+r)^T,$$

where T is the last year that the decision (project) is expected to imply a cost or a benefit.

Continuous time

Interest rate contracts inform depositors about adding interest to the asset semiannually, monthly or even daily. If interest is added twice a year rather than once, it means that $X_1 = X_0(1+r/2)^2$, rather than $X_1 = X_0(1+r)^1$. If interest is added m times a year, the formula reads $X_1 = X_0(1+r/m)^m$,. In their elementary Calculus course students are requested to prove that $\lim_{m\to\infty} (1+r/m)^m = e^r$. Thus the 'continuous time' version of the formula $X_t = X_0(1+r)^t$ reads $X_t = X_0e^{rt}$, and $NPV=X(1+1/(1+r)+...+1/(1+r)^T)$ reads $NPV=Xe^{rT}$.

Long run

If $T \to \infty$, then (in the limit):

$$NPV = X/r$$

(easy proof referring to the high-school mathematics of geometric series)

The present value of the future amount of X_T=1,000,000

	T=1	*T*=5	*T*=10	*T*=20	*T*=50	*T*=100	*T*=200
r=0%	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
r=1%	990,099	951,466	905,287	819,544	608,039	369,711	136,686
r=4%	961,538	821,927	675,564	456,387	140,713	19,800	392
r=8%	925,926	680,583	463,193	214,548	21,321	455	0.21
r=12%	892,857	567,427	321,973	103,667	3,460	12	<0.01

The present value of a flow of a constant amount of X=100

	T=10	T=50	T=100	$T=\infty$
r=0%	1,000	5,000	10,000	∞
r=1%	947	3,920	6,303	10,000
r=4%	811	2,148	2,451	2,500
r=8%	671	1,223	1,249	1,250
r=12%	565	830	833	833

Internal Rate of Return (IRR)

IRR – a discount rate which makes the NPV=0. In other words, IRR is a discount rate r such that:

$$X_0/(1+r)^0 + X_1/(1+r)^1 + X_2/(1+r)^2 + \ldots + X_T/(1+r)^T = 0$$

If $X_0, X_1, \ldots, X_{\tau-1} < 0$, and $X_\tau, X_{\tau+1}, \ldots, X_T > 0$, then IRR is the only r which solves the equation above (IRR is unique).

For typical projects this condition is satisfied (i.e. one needs to bear some investment cost in the beginning, and then one benefits from it).

IRR is a useful indication of profitability:

- If IRR is higher than the interest rate available for investor (to borrow the money in order to finance it), the project is efficient (and it is worth financing)
- If IRR is lower than the interest rate available for investor (to borrow the money in order to finance it), the project is inefficient (and it should be abandoned)
- If IRR is negative then the project does not make sense irrespective of the terms of availability of the money (unless you do not have to pay for the credit; the bank pays you to borrow the money).

Wind-mill case study (1)

- 1 MW capacity
- works 2000 hours per annum
- costs 2 million euro
- sells electricity at 50 euro/MWh
- requires no maintenance costs for 30 years

Wind-mill case study (2)

- One-time investment cost of 2,000,000 € on January 1 (2,000 k€)
- The annual revenue of 100,000 € every December 31 (100 k€)

Wind-mill case study (2)

NPV =

$= -2{,}000k/(1+r)^0 + 100k/(1+r)^1 + 100k/(1+r)^2 + \dots + 100k/(1+r)^{30} =$

$= -2.000k + 100k(1/(1+r) + 1/(1+r)^2 + \dots +1/(1+r)^{30}) =$

$= -2{,}000k + 100k((1-q^{30})/(1-q)),$

where $q=1/(1+r)$, if $q\neq 1$, i.e. $r\neq 0$

Wind-mill case study (3)

- IRR=0.03, since NPV=0 if and only if q=0.97, i.e. r=0.03.
- Note: r needs to be calculated from the formula: q=1/(1+r). Hence r=1/q-1. Incidentally, it is 0.03.

- If the discount rate is higher than 0.03, then the investment will never pay back.
- If the discount rate is lower than 0.03 then the investment makes sense.

Time consistency

High-school mathematics:

$$(1+r)^{K+N}=(1+r)^{K}(1+r)^{N}$$

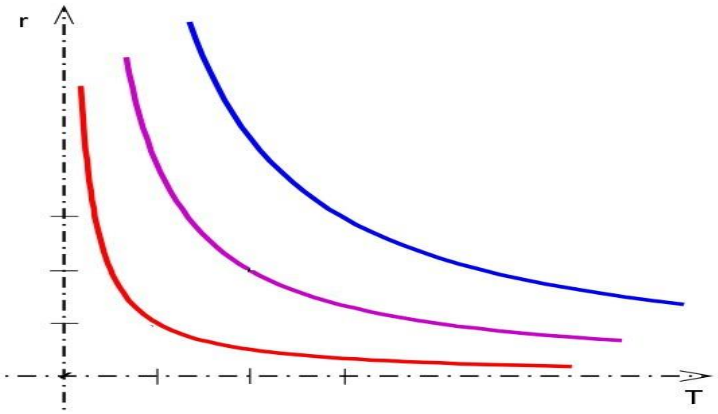
If X is to be observed in N+K years, then one can calculate its NPV in two steps. First in N (this will read $X/(1+r)^K$; because N+K-N=K), and then NPV of this amount now, i.e. N years earlier. The result will read:

$$(X/(1+r)^{K})/(1+r)^{N}=X/(1+r)^{K+N}).$$

This formula holds whenever r=const.

Hyperbolic discounting

Empirical research suggests that the longer the period to be discounted, people apply lower discount rates (if T<10, then people are used to, say, r=3%; if T=50 then they apply a much lower rate, perhaps r≈0. The relationship resembles the graph of f(x)=1/x, i.e. a hyperbole. Hence the name "hyperbolic discounting".



Economic rationale for hyperbolic discounting (or – in general – for applying a lower discount rate for a much longer period) is based on the following fact:

- If a decision can be taken in two steps (the first step at the beginning of a period, and then the second step after some time), then each of the steps (involving decisions of shorter – and thus implying more predictability – time horizons) must be justified more rigorously.

Ramsey equation

$$\mathbf{r = \rho+\eta g},$$

where:

r – discount rate,
ρ – "pure time preference"
η – income elasticity of consumption,
g – expected rate of wealth growth

- Derived as a solution to optimizing intertemporal allocation of effort
- Ethical consideration: $\rho = 0$

Questions:

Q-4 Discounting the future with a discount rate reflecting social time preference

[a] can be applied only in an economy where the rate of inflation does not exceed the rate of GDP growth.

[b] helps to justify investment projects characterized by low benefits which are spread over a long period of time.

[c] helps to justify investment projects characterized by high investment costs concentrated at the very beginning of the investment process.

<u>[d] allows to compare costs and benefits realized in different periods</u>.

[e] none of these.

Exercises:

E-4 Calculate the IRR of the following windmill project. Investment cost is 1.5 Meuro (to be spent once on December 31). Next year the windmill will start producing electricity of 2 GWh annually (we assume that revenues from these sales accrue to the windmill's account each year on December 31). It will operate for 30 years without any maintenance costs. We also assume that it will not require any repair expenditures. The wholesale price of electricity is 40 euro/MWh.

According to Frank Knight (and widely accepted by economists), <u>risk</u> means a kind of uncertainty that leads to an array of outcomes with probabilities attached.

<u>Definition</u>

Elements of the set X (where consumer's bundles are taken from) can be interpreted as lotteries $L=(p_1,...,p_n)$, where $p_1+...+p_n=1$. $\mathcal{L}$ – the set of such lotteries; their outcomes – numbered 1,...,n – are given.

Theorem
A convex combination of lotteries is also a lottery

Definition: Von Neumann-Morgenstern function of expected utility. U: $\mathcal{L} \rightarrow \Re$ such that:

$$\exists u_1,\ldots u_n \in \Re \, \forall L=(p_1,\ldots,p_n) \in \mathcal{L} \, [U(L)=u_1 p_1+\ldots+u_n p_n].$$

Numbers u_i can be interpreted as utilities of "degenerated" lotteries $L^1=(1,0,\ldots,0),\ldots,L^n=(0,\ldots,0,1)$.

Theorem

A function $U: \mathcal{L} \to \Re$ is a vNM function of expected utility if and only if $\forall K=1,2,\ldots \forall L_1,\ldots,L_K \in \mathcal{L}$ $\forall \alpha_1,\ldots,\alpha_K>0$

$$[U(\alpha_1 L_1+\ldots+\alpha_K L_K)=\alpha_1 U(L_1)+\ldots+\alpha_K U(L_K)]$$

Proof

$\Leftarrow$

Let $L=(p_1,\ldots,p_n)$. Let us define "degenerated" lotteries $L^1,\ldots,L^n$ such that $L^i=(0,\ldots,0,1,0,\ldots,0)$; there is a unit probability at the *i*th coordinate. Then $L=p_1L^1+\ldots+p_nL^n$ and $U(L)=U(p_1L^1+\ldots+p_nL^n)=p_1U(L^1)+\ldots+p_nU(L^n)=$ $=p_1u_1+\ldots+p_nu_n$, where middle equality is satisfied by the assumption.

$\Rightarrow$

Let us consider a combined lottery $(L_1,\ldots,L_K;\alpha_1,\ldots,\alpha_K)$, where $L^i=(p_1^i,\ldots,p_n^i)$. Let $L'=\alpha_1L_1+\ldots+\alpha_KL_K$. Then $U(L')=$
$=U(\alpha_1L_1+\ldots+\alpha_KL_K)=$
$=u_1(\alpha_1p_1^1+\ldots+\alpha_Kp_1^K)+\ldots+u_n(\alpha_1p_n^1+\ldots+\alpha_Kp_n^K)=$
$=\alpha_1(u_1p_1^1+\ldots+u_np_n^1)+\ldots+\alpha_K(u_1p_1^K+\ldots+u_np_n^K)=$
$=\alpha_1U(L_1)+\ldots+\alpha_KU(L_K)$. The second equality in this chain results from the assumption directly.

Definition

A lottery with money outcomes $x_1,\ldots x_n$. There is a function $u:\Re\rightarrow\Re$, and numbers $u(x_1),\ldots,u(x_n)$ are called Bernoulii's utilities. A certainty equivalent of the lottery $L=(p_1,\ldots,p_n)$ is the number $c(L,u)$ such that $u(c(L,u))=u(x_1)p_1+\ldots+u(x_n)p_n$.

Risk aversion and risk neutrality implied by Bernoulli's utility, respectively:

$\forall L=(p_1,\ldots,p_n)\in\mathcal{L}\ [u(x_1)p_1+\ldots+u(x_n)p_n\leq u(x_1p_1+\ldots+x_np_n)]$

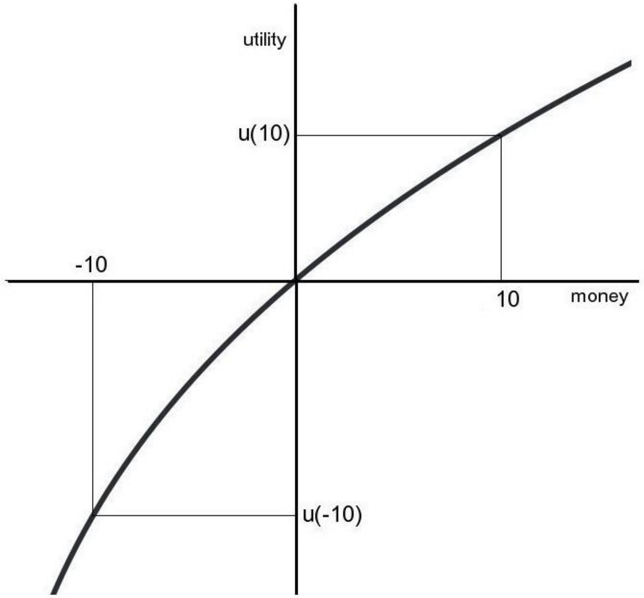
$\forall L=(p_1,\ldots,p_n)\in\mathcal{L}\ [u(x_1)p_1+\ldots+u(x_n)p_n=u(x_1p_1+\ldots+x_np_n)]$

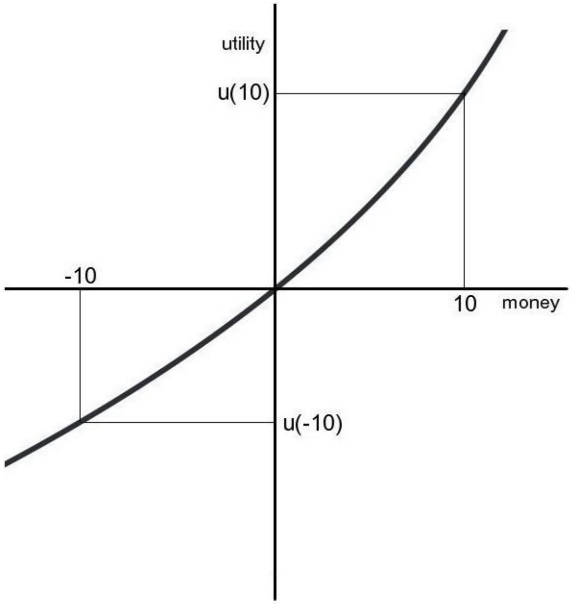
Theorem

The following conditions are equivalent:

1. The consumer has risk aversion
2. u is a concave function
3. $\forall L=(p_1,...,p_n)\in\mathcal{L}$ $[c(L,u)\leq x_1p_1+...+x_np_n]$

Attitudes towards risk





Note (the Allais paradox)
There are three outcomes of lotteries:

- x_1=0,
- x_2=1 million USD, and
- x_3=5 million USD

Experiment 1 (most people prefer L_1):
Choose between two lotteries:
L_1=(0,1,0) and L_2=(0.01,0.89,0.1)

Experiment 2 (most people prefer L_4):
Choose between two lotteries:
L_3=(0.89,0.11,0) and L_4=(0.9,0,0.1)

Allais paradox (cont.)

People who choose L_1 in the first experiment and L_4 in the second one do not comply with the vNM theory.

<u>Proof</u>:

If the vNM theory was followed, then Bernoulli's utilities would have been applied:

- $u(x_1)=u_1$,
- $u(x_2)=u_2$, and
- $u(x_3)=u_3$.

Proof (cont.)

The outcome of the first experiment implies that:

$$u_2 > 0.01u_1 + 0.89u_2 + 0.1u_3.$$

The outcome of the second experiment implies that:

$0.9u_1 + 0.1u_3 > 0.89u_1 + 0.11u_2.$

These two inequalities contradict each other, since the second one can rewritten as:

$$0.01u_1 + 0.1u_3 > 0.11u_2, \text{ and consequently}$$

$$0.01u_1 + 0.1u_3 > u_2 - 0.89u_2, \text{ or}$$

$$0.01u_1 + 0.89u_2 + 0.1u_3 > u_2$$

which contradicts the first one.

Questions:

Q-5 The "certainty equivalent" of a lottery

[a] depends on the lottery, not on the behaviour of a consumer

[b] measures whether the lottery is an honest one

[c] is equal to the lowest possible gain from the lottery

[d] is the weighted sum of outcomes of the lottery

<u>[e] none of the above</u>

Exercises:

E-5 The Allais paradox does not imply that people behave irrationally. It merely shows that vNM theory cannot explain how people actually behave. Please offer explanations why people's attitude towards risk may be not consistent with the vNM theory (even for people who understand probabilities and whose preferences for deterministic bundles are rational).

Partial equilibrium: demand and supply equate in a given (single) market

General equilibrium: demand and supply equate in all markets (simultaneously)

Notation:

k – the number of consumers
n – the number of markets (products); one of the commodities can be labour – i.e. a resource owned by every consumer
r – the number of firms

Numbering of consumers: i=1,...,k
Numbering of markets (products): j=1,...,n
Numbering of firms: h=1,...,r

Pure exchange economy (no production)

- *Endowment* (initial allocation) of the *i*th consumer: $\omega_{i1}, \omega_{i2},\ldots, \omega_{in}$
- *Total endowment* of the *j*th commodity: $\omega_j = \omega_{1j}+\omega_{2j}+\ldots+\omega_{kj}$
- *Gross demand* (final allocation) of the *i*th consumer: $x_{i1},x_{i2},\ldots,x_{in}$
- *Total demand* for the *j*th commodity: $x_j = x_{1j}+x_{2j}+\ldots+x_{kj}$
- *Excess demand* of the *i*th consumer: $x_{i1}-\omega_{i1},x_{i2}-\omega_{i2},\ldots,x_{in}-\omega_{in}$

General case: with production:

y_{hj} – net supply of the *j*th commodity from the *h*th firm

Note:

If $y_{hj}<0$, then the *h*th firm uses more of the *j*th commodity than it produces. The number $-y_{hj}=|y_{hj}|$ is then the value of the (secondary) demand of the *h*th firm for the *j*th commodity

Feasibility of production

Production schedules of the *h*th firm must be feasible, i.e. $(y_{h1},\ldots,y_{hn}) \in Y_h$, where Y_h is the firm's production set

Definitions of the total market supply and demand:

- $y_j = y_{1j}+\ldots+y_{rj}$ – total net supply of the *j*th commodity
- $z_j = x_j-\omega_j-y_j$ – total excess demand for the *j*th commodity

Profit of the *h*th firm: $\pi_h = p_1 y_{h1} + \ldots + p_n y_{hn}$

Share of the *i*th consumer in the *h*th firm's profit: θ_{ih}

For every firm *h*: $\theta_{1h} + \ldots + \theta_{kh} = 1$

Feasible allocation – the following system is satisfied:

$$x_{11}+x_{21}+\ldots+x_{k1} = \omega_{11}+\omega_{21}+\ldots+\omega_{k1}+y_1,$$
$$x_{12}+x_{22}+\ldots+x_{k2} = \omega_{12}+\omega_{22}+\ldots+\omega_{k2}+y_2,$$
$$\ldots$$
$$x_{1n}+x_{2n}+\ldots+x_{kn} = \omega_{1n}+\omega_{2n}+\ldots+\omega_{kn}+y_n,$$

Additional notation:

- **x** = $(x_1,\ldots,x_n)^T$ – demand vector (column)
- **p** = $(p_1,\ldots,p_n)$ – price vector (row)

Note:

The numbers ω_{ij} and θ_{ih} are parameters of the general equilibrium model. The numbers x_{ij}, y_{hj}, and therefore also z_j are functions of prices **p**.

Definition of the Walras equilibrium

Any pair (**p***,**x***) such that for every commodity *j* we have:

$x_j^{*} := x_{1j}(\mathbf{p^*})+\ldots+x_{kj}(\mathbf{p^*}) \leq \omega_{1j}+\ldots+\omega_{kj}+y_{1j}+\ldots+y_{hj}$,

i.e. $z_j(\mathbf{p^*})\leq 0$

Definition of the budget line (BL) of the *i*th consumer:

$p_1x_{i1}+\ldots+p_nx_{in} = p_1\omega_{i1}+\ldots+p_n\omega_{in}+\theta_{i1}\sum_j p_jy_{1j}+\ldots+\theta_{ir}\sum_j p_jy_{rj}$

Theorem (Walras Law)

If all the consumers keep their budget lines then the value of excess demand is 0, i.e.: $\sum_j p_jz_j = 0$

Proof of the Walras Law:

$$\sum_j p_j z_j$$

$$\overset{1}{=} \sum_j p_j \left(x_j - \omega_j - y_j\right) =$$

$$\overset{2}{=} \sum_j p_j \left(\sum_i x_{ij} - \sum_i \omega_{ij} - \sum_h y_{hj}\right) =$$

$$\overset{3}{=} \sum_j p_j \left(\sum_i x_{ij} - \sum_i \omega_{ij} - \sum_h \left(\sum_i \theta_{ih}\right) y_{hj}\right) =$$

$$\overset{4}{=} \sum_j \sum_i \left(p_j x_{ij} - p_j \omega_{ij} - \sum_h \theta_{ih} p_j y_{hj}\right) =$$

$$\overset{5}{=} \sum_j \sum_i \left(p_j x_{ij} - \sum_j p_j \omega_{ij} - \sum_h \theta_{ih} \sum_j p_j y_{hj}\right) =$$

$$\overset{6}{=} \sum_j 0 = 0.$$

Explanation of steps:

1 by definition of excess demand

2 definitions of x_j, ω_j and y_j were substituted

3 the factor $\theta_{1h}+...+\theta_{kh}=1$ was inserted for every h

4 multiplication by p_j was carried out

5 the order of summation was changed

6 the expression to be summed is the difference between the left-hand-side and the right-hand side of a budget line

Note

The Walras Law holds for any system prices **p**, not only for equilibrium prices

Corollaries from the Walras Law

1 If the value of excess demand in n-1 markets is equal to 0, then also in the remaining *n*th market the value of excess demand is 0

2 In order to find the Walras equilibrium it is sufficient to solve a system of n-1 equilibrium prices in corresponding n-1 markets

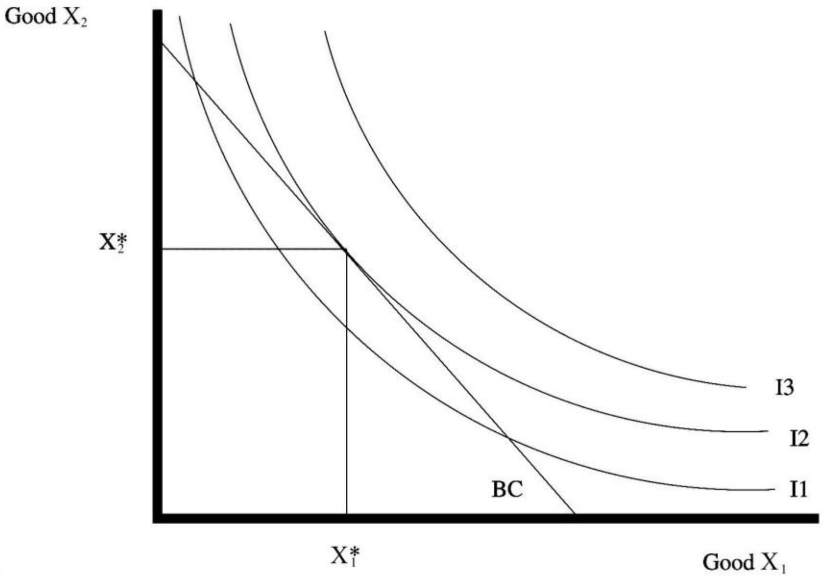
Corollaries from the Walras Law (cont.)

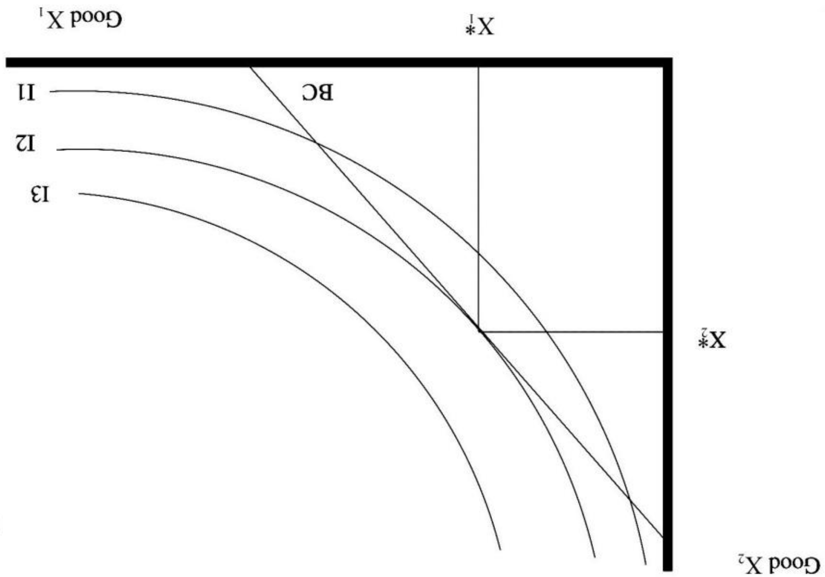
3 A *free good* is any good the excess demand for which is negative. In equilibrium the price of such a good is zero

4 A *desired good* is any good whose excess demand at the zero price is positive. If all goods 1,...,n are desired, then in equilibrium the excess demand in all markets is 0

Edgeworth box

A graphical analysis of feasible allocations in a pure exchange economy when k=n=2 (superposition of two coordinate systems for the analysis of a consumer's choice: the width of the rectangle = $\omega_{11}+\omega_{21}$, the height of the rectangle = $\omega_{12}+\omega_{22}$; the second system is rotated by 180^o)


Good X_2
X_2^*
I3
I2
I1
BC
X_1^*
Good X_1


Good X_1
X_1^*
BC
I1
I2
I3
X_2^*
Good X_2

Theorem

In a perfectly competitive market with two consumers characterized by convex indifference curves, equilibrium in an Edgeworth box will be achieved in the point where indifference curves of these consumers are tangent to each other. The slope coefficient of the tangent line is equal (in absolute value terms) to the proportion of prices p^*_1/p^*_2.

QF-6-17

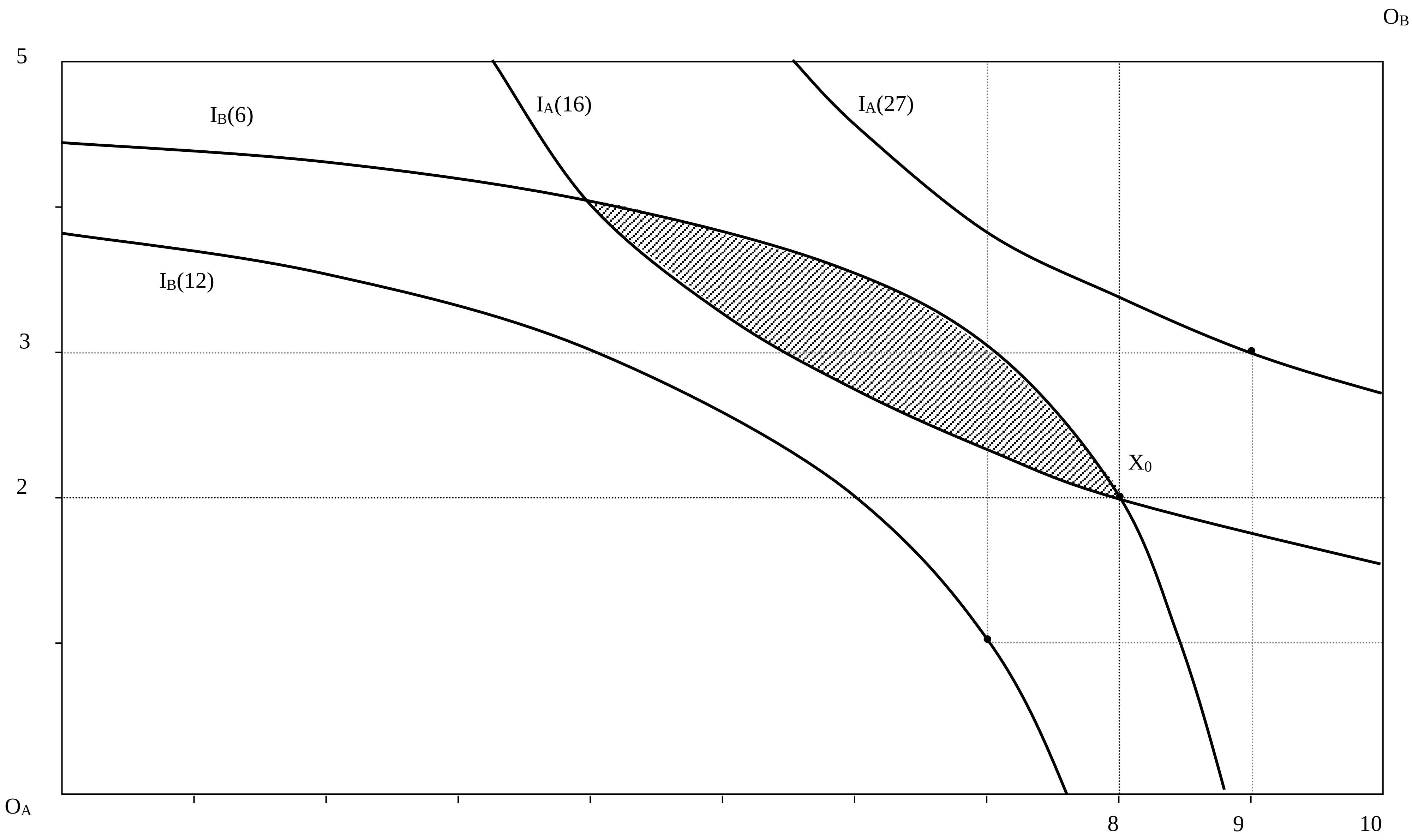


Fig. 1

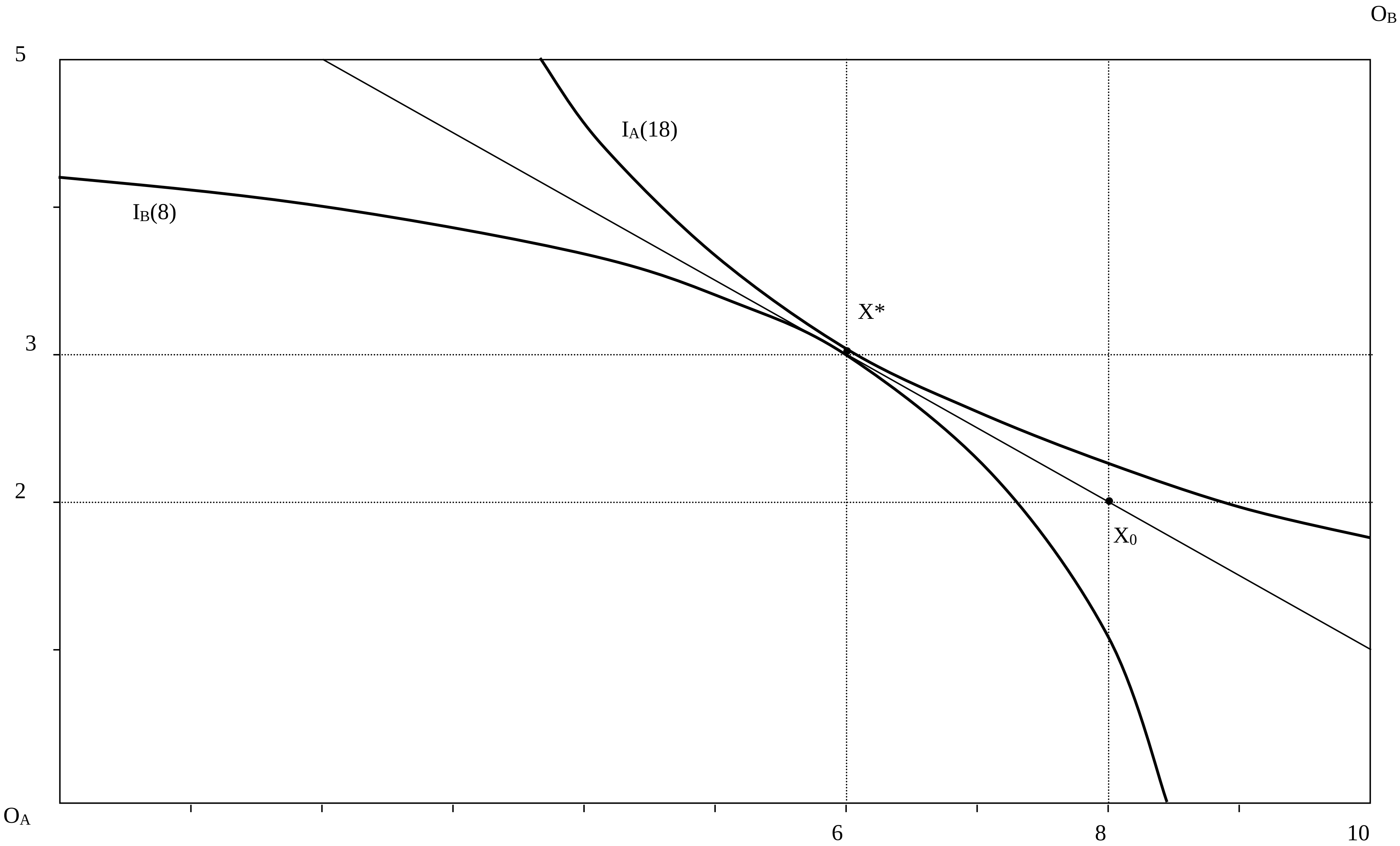


Fig.2

Pictures illustrate the following situation: Indifference curves of A, $I_A(\alpha)$ are given by the formula $x_{2A}=\alpha/x_{1A}$, while indifference curves of B, $I_B(\beta)$ are given by $x_{2B}=\beta/x_{1B}$ ($\alpha,\beta>0$ – parameters); additionally, we assume that the total quantity of the first good is 10, while that of the second – 5. Moreover, the diagram corresponds to the initial allocation of the first good 8:2, while of the second one – 2:3 (point X_0). There are two indifference curves containing this point: $x_{2A}=16/x_{1A}$ ($\alpha=16$) and $x_{2B}=6/x_{1B}$ ($\beta=6$). A would prefer to be on a higher indifference curve, say, in (9,3), i.e. on the curve $x_{2A}=27/x_{1A}$ ($\alpha=27$). At the same time, B would like to have more of everything, i.e. to be in, say, (3,4), i.e. on the curve $x_{2B}=12/x_{1B}$ (β=12). It is impossible to satisfy these expectations at the same time. One solution which can place both agents in a jointly preferred point (one should solve a system of simultaneous equations) is: $x_{1A}*=6$, $x_{2A}*=3$, $x_{1B}*=4$, $x_{2B}*=2$, $\alpha=18$, $\beta=8$, $p= p_1/p_2 =0,5$ (see Figure 2). Agents A and B are on $I_A(18)$ and $I_B(8)$, respectively, and they are better of than in X_0. One can see from the figure that they cannot improve their situations further simultaneously. Equilibrium prices which satisfy this solution are multiple, e.g. $p_1=1$, $p_2=2$, or $p_1=7$, $p_2=14$, or $p_1=0,5$, $p_2=1$ etc, as long as $p_1/p_2=p=0,5$.

Questions:

Q-6 Walras Law
[a] applies to an economy in market equilibrium
[b] implies that the equilibrium price of a free good is zero
[c] requires that all agents are price takers
[d] applies to a pure exchange economy only (no production)
[e] none of the above

Exercises:

E-6 Please design a price negotiation mechanism such that agents not located in a Pareto optimum (e.g. point X_0 in the Edgeworth box in Figure 1) can make a Pareto improvement, but do not achieve optimum (attainable in a single transaction, i.e. point X^* in Figure 2).

Two fundamental welfare economics theorems

They link Pareto optimality to market equilibrium

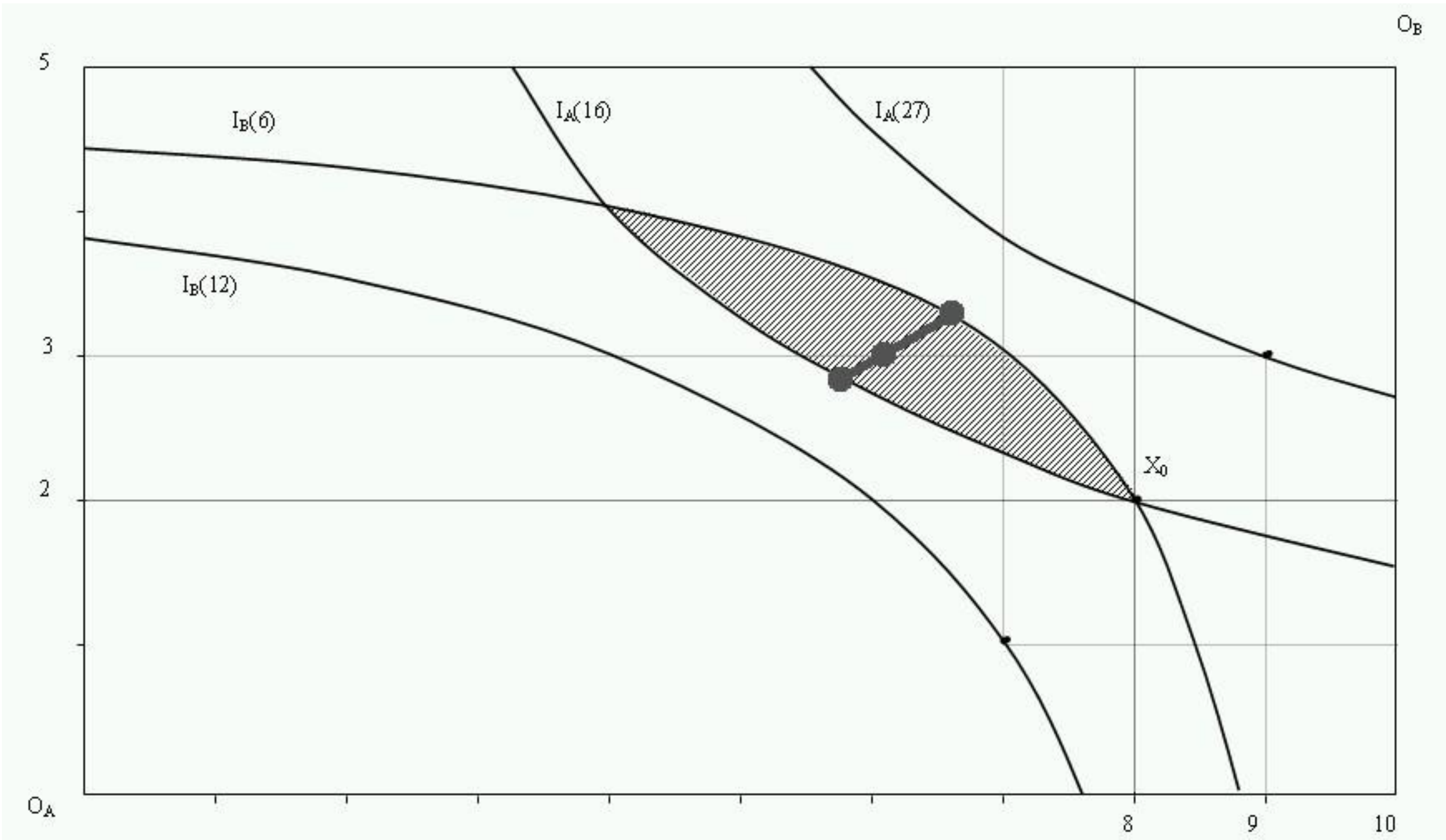
Definition

An allocation is called a Pareto optimum if no consumer can be made better off (moved to a higher indifference curve), unless another consumer is made worse off (moved to a lower indifference curve)

Definition

In Edgeworth box, a *contract curve* (or a *Pareto set*) is the set of all allocations that are Pareto optima

Contract curve



First Welfare Economics Theorem

If (**p***,**x***) is a Walras equilibrium in a market whose all participants are price-takers interacting with each other exclusively through voluntary transactions, then **x*** defines a Pareto optimum

Second Welfare Economics Theorem

Let us assume that

- all agents are price-takers,
- all firms have convex production sets,
- all consumers have convex indifference curves (surfaces).

Then for any Pareto optimum **x*** there exist a price vector **p*** and an initial allocation of endowments such that **x*** can be attained as a Walras equilibrium (**p***,**x***)

Overview of Welfare Economics Theorems

Let us assume that

Walras Equilibrium (WE)
Pareto Optimum (PO)

(1) WE$\Rightarrow$PO
(2) PO$\Rightarrow$WE

Not an equivalence (different assumptions)
Proof of the Second Theorem more complicated

Proof of the First Welfare Economics Theorem for a pure exchange economy

Let us assume that $(p^*_1,\ldots,p^*_n,\mathbf{x^*}_1,\ldots,\mathbf{x^*}_n)$ is a Walras equilibrium, and let us assume that it is not a Pareto optimum, i.e. there is an allocation $\mathbf{x'}=(\mathbf{x'}_1,\ldots,\mathbf{x'}_n)$ satisfying the same constraints yet preferred to $\mathbf{x^*}=(\mathbf{x^*}_1,\ldots,\mathbf{x^*}_n)$. But if $\mathbf{x'}$ was preferred to $\mathbf{x^*}$ while $\mathbf{x^*}$ was established as an equilibrium in a market with prices $\mathbf{p^*}$, then it means that $\mathbf{x'}$ must have been not affordable. In other words,

Proof of the First Welfare Economics Theorem for a pure exchange economy (cont.)

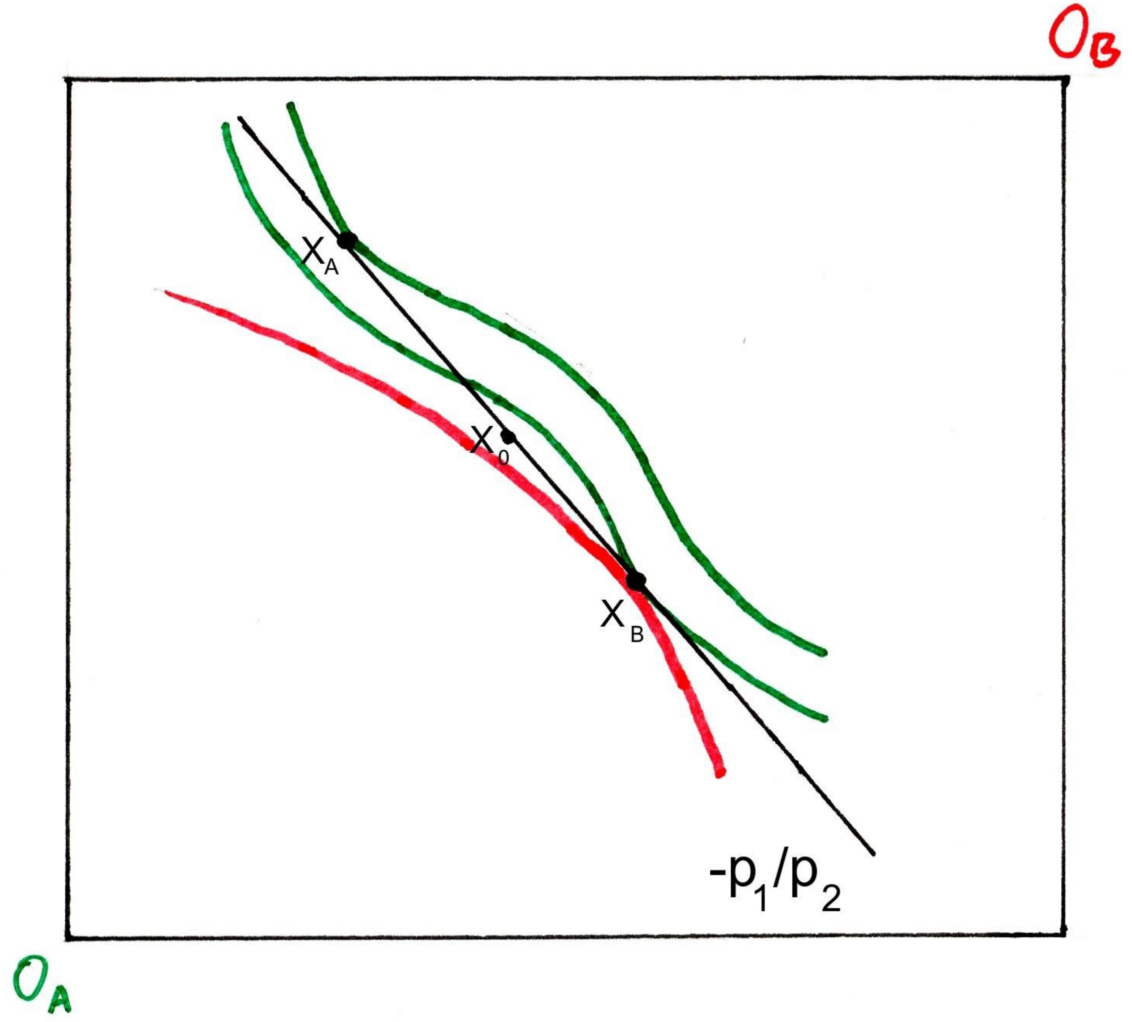
$$p^*_1(x'_{11}+...+x'_{k1}) +...+ p^*_n(x'_{1n}+...+x'_{kn}) >$$
$$p^*_1(x^*_{11}+...+x^*_{k1}) +...+ p^*_n(x^*_{1n}+...+x^*_{kn}).$$

However, in a pure exchange economy:

$$x'_{11}+...+x'_{k1} = x^*_{11}+...+x^*_{k1} = \omega_1, ... ,$$
$$x'_{1n}+...+x'_{kn} = x^*_{1n}+...+x^*_{kn} = \omega_n.$$

Hence $p^*_1\omega_1 +...+ p^*_n\omega_n > p^*_1\omega_1 +...+ p^*_n\omega_n$ which is a contradiction.

An Edgeworth box can be used to graphically prove the Second Welfare Economics Theorem in a special case of a pure exchange economy with two goods and two consumers



O_B
X_A
X_0
X_B
$-p_1/p_2$
O_A

Note:

- Walras equilibrium in a monopolistic market is not a Pareto optimum (neither of the welfare economics theorems applies).
- Violating convexity assumptions affects the second welfare economics theorem.

Welfare economics *versus* monopoly

- Walras equilibrium in a monopoly market may be not Pareto optimal
- A price-discriminating monopolist does not comply with the general equilibrium setting (different prices for different clients)

Tacit assumptions of welfare economic theorems

- No externalities
- No asymmetric information

Questions:

Q-7 Non-convexity of indifference curves in Edgeworth box
[a] does not allow to identify Pareto optima
<u>[b] makes Pareto optima impossible to attain as market equilibria</u>
[c] violates the consumers' rationality assumptions
[d] makes market equilibria non-optimal in Pareto sense
[e] none of the above

Exercises:

E-7 Please explain which step in the proof of the first welfare economics theorem uses the price-taking assumption.

Asymmetric information

The buyer has less information about the commodity than the seller or *vice versa*; acquiring information is possible, but costly.

Note

Asymmetric information implies a market failure

- Equilibrium does not have to be a Pareto optimum
- Equilibrium may not exist

Examples of asymmetric information

- Used cars
- Insurance policies
- Employment contracts

Adverse selection

A better product is driven out from the market by a worse one; hidden information implies suboptimal demand or supply

Examples of adverse selection

In insurance market with asymmetric information on damages:

- Efficiency can be improved by imposing obligatory insurance in order to bring in to the market low-risk groups
- Low-risk groups can be also brought in without government intervention (by decentralized measures)

Examples of adverse selection (cont.)

In second-hand car market with asymmetric information on the quality of a car:

- Efficiency can be improved spontaneously (e.g. by encouraging customers to buy cars offered with a guarantee)

Note

If hidden information leads to suboptimal demand resulting from a loss imposed on somebody else (e.g. caused by the supply of a low quality commodity) then the government intervention can correct for the market failure, if it lowers the average level of this loss (*externality* in economic language)

Moral hazard

The lack of *ex post* incentives to care for something that was *ex ante* assumed in the contract; hidden action leads to suboptimal supply

Notes

- There is no market failure if actions can be perfectly controlled (e.g. smoking habits)
- If hidden action leads to suboptimal supply (as higher supply would give buyers an incentive for a moral hazard), then government intervention is typically unpurposeful, since the problem results from the cost of information, rather than from losses imposed on others (*externalities*)
- If behaviour cannot be observed, then efficiency in the insurance market requires that the insurance coverage is not complete

<u>Signalling</u>
Making information credible

<u>Note</u>
Efficiency is improved as a result of a guarantee for the commodity sold (the guarantee does not imply unnecessary costs)

<u>Note</u>
Efficiency can be compromised as a result of signalling which implies additional (unnecessary) costs (so-called *separating equilibrium*)

Example

- Marginal productivity of workers (unobserved): dull a_1, smart a_2; $a_2 > a_1$
- Relative share in the population (observed): smart b, dull (1-b)
- If the employer cannot tell who is smart and who is dull (but $\partial Q/\partial L_1 = a_1$ and $\partial Q/\partial L_2 = a_2$), where Q – production, L_1 – the employment of the dull, L_2 – the employment of the smart, then – the wage offered should be the same for all and equal to $w = (1-b)a_1 + ba_2$

Example (cont.):

- The cost of acquiring education at the level e^* is c_1 for the dull and c_2 for the smart, and $c_2 < c_1$
- Let $(a_2 – a_1)/c_1 < e^* < (a_2 – a_1)/c_2$
- Then the dull choose $e_1 = 0$, and the smart ones $e_2 = e^*$, since for the smart ones: the gain $= a_2 – a_1 > c_2 e^* =$ cost, while for the dull ones: the gain $= a_2 – a_1 < c_1 e^* =$ cost
- An education certificate corresponding to e^* signals to the employer the quality of its prospective employee.

Questions:

Q-8 Asymmetric information implies that
[a] the boss knows more than the employee
[b] market does not clear in equilibrium
[c] seller always knows more than the buyer
[d] random effects cannot be predicted with certainty
<u>[e] none of the above</u>

Exercises:

E-8 Please comment on the Polish editor's note to the signalling game that – contrary to the assumptions made by Hal Varian – education may change marginal productivities. How this affects the signalling game?

Non-zero sum two-person game

A representation of a decision situation with a table of pairs of numbers (P_{ij},D_{ij}). Index i=1,...,m, where m is the number of strategies (decision variants) for the first player, and j=1,...,n, where n is the number of strategies (decision variants) for the second player. The numbers P_{ij} are payments to the first player, and D_{ij} – to the second player, if the first chose the ith strategy, and the second – the jth one.

Definition

The strategy i_0 of the first player is called *dominant*, if for any strategy i of the first player, and any strategy j of the second player, $P_{i_0 j} \geq P_{ij}$; likewise, strategy i_0 is *dominated*, if there exists strategy i such that for any strategy j, $P_{i_0 j} \leq P_{ij}$. (analogously for strategies of the second player).

Nash strategy

Any pair of strategies (i_0,j_0) such that $P_{i_0j_0}=\max_i\{P_{ij_0}\}$ and $D_{i_0j_0}=\max_j\{D_{i_0j}\}$.

Corollary

If players have dominant strategies, then their pair is a Nash equilibrium

Corollary

If players are in a Nash equilibrium, then – if they maximize their payoffs – neither has a motivation to unilaterally change his or her strategy

Note

There exist games with no Nash equilibrium

Example

		D	
		Y	N
P	Y	(4,2)	(-2,3)
	N	(2,1)	(-1,0)

Note

Strategies making a Nash equilibrium do not have to be dominant ones (nor unique)

Example

		D	
		Y	N
P	Y	**(4,2)**	**(-2,2)**
	N	(2,1)	(-3,0)

Note ("Prisoner's dilemma")

Nash equilibrium may include strategies that do not maximize the payoff sum for both players

		Second	
		C	D
First	C	**(-12,-12)**	(0,-18)
	D	(-18,0)	(-1,-1)

Note ("Prisoner's dilemma" continued)

Nash equilibrium is for (C,C) which is the very worst outcome for two players; i.e. Nash equilibrium does not necessarily 'optimize' the global outcome (which – in this case – would be (D,D))

Note (behavioural assumption)

Nash equilibrium explains market equilibrium in some circumstances (examples: Cournot and Bertrand models of duopoly)

Oligopolistic models

An oligopoly is when the supply comes mainly from a small number of firms each of which has some impact on the market price.

A duopoly is the simplest case of an oligopoly, when the supply comes from two firms only.

In an oligopolistic market, the price rule derived for a perfectly competitive market ($p=MC(y^*)=AC(y^*)$), may not hold. Prices depend on how oligopolists compete.

General notation:

- y_i – supply from the *i*th rival
- y – aggregate supply
- p_i – price charged by the *i*th rival
- p – price determined by the market

Key features of oligopolistic competition relate to two questions: (1) do rivals make decisions simultaneously, or is there one of them (known as the leader) who takes a decision first while others (known as followers) make decisions afterwards; and (2) do rivals compete with prices or quantities.

Accordingly we have four models of oligopolistic competition:

1. Cournot: rivals make quantity decisions simultaneously
2. Bertrand: rivals make price decisions simultaneously
3. Quantity leadership (Stackelberg): one makes a quantity decision first
4. Price leadership: one makes a price decision first

General idea of analyzing oligopolistic markets

- Determine a demand that the market will reveal: D(p)=y
- Let $p_i y_i$-TC(y_i) be the profit enjoyed by the *i*th firm (TC – total cost)
- Let p_i^* and y_i^* be profit maximizing decisions of the *i*th firm
- Firm *i* makes its decisions which maximize its profit, taking into account expected decisions of its rivals
- We look for an equilibrium in a sense that all firms make decisions exactly as they were expected by their rivals

In addition (for simplicity) we assume:

- The number of firms is two (i.e. i=1 or i=2)
- The demand curve is known and linear (i.e. p=a-by)
- Both firms have identical cost functions and $MC_1=MC_2=AC_1=AC_2=c=const$

Cournot duopoly model:

- The price is implied by the total supply: $p=a-b(y_1+y_2)$
- The first rival makes a quantity decision that maximizes its profit expecting that the second rival does the same
- Thus they solve two simultaneous maximization problems:
  - $\max_{y_1}\{(a-b(y_1+y_2))y_1-cy_1\}$
  - $\max_{y_2}\{(a-b(y_1+y_2))y_2-cy_2\}$

Cournot duopoly model (cont.)

- FOC for these problems are:
  - $a-2by_1-by_2-c=0$
  - $a-2by_2-by_1-c=0$

- Solving these equations yields
  - $y_1 = y_2 = (a-c)/3b$
  - $y = 2(a-c)/3b$
  - $p = (a+2c)/3$

Cournot duopoly model (cont.)

Note

Please note that p = (a+2c)/3 is greater than competitive price under 'typical conditions', i.e. than c. To see this, note that a>c (otherwise in the competitive market the demand and supply would not intersect). Thus (a+2c)/3>(c+2c)/3=c.

Corollary

The Cournot price is higher than the competitive one.

Bertrand duopoly model

- Each rival makes its price decision p_i
- The one whose price is higher loses all the buyers
- The one whose price is lower gets all the buyers
- The lower price becomes the market price p
- The winner enjoys profit (p-c)(a-p)/b
- If $p_1=p_2$ then $p= p_1=p_2$, $y=(a-p)/b$,
  and $y_1=y_2=(a-p)/(2b)$

Bertrand duopoly model (cont.)

In equilibrium $p=c$, and profits vanish

Proof:

- No price $p<c$ can be sustained since it implies losses
- No price $p>c$ can be sustained since at least one firm has a motivation to offer a price $\in(c,p)$ in order to undercut the rival's price, get all the clients and increase the profit

Questions:

Q-9 Nash equilibrium concept implies that
[a] players do not maximize their payoffs
[b] players seek strategies to minimize their rivals' payoffs
[c] it is always possible increase payoffs through cooperation between players
<u>[d] players are not pressed to co-operate with each other</u>
[e] none of the above

Exercises:

E-9 Please design a game such that Bertrand model can be interpreted as a Nash equilibrium.

Mixed strategies (Definition)

Strategies defined so far can be called *pure*. A game can be defined where pure strategies are selected by players randomly with certain probabilities $\boldsymbol{\pi}=(\pi_1,\ldots,\pi_m)$ and $\boldsymbol{\delta}=(\delta_1,\ldots,\delta_n)$, respectively (for the first and the second player), where $\pi_1,\ldots,\pi_m\geq 0$, $\pi_1+\ldots+\pi_m=1$ and $\delta_1,\ldots,\delta_n\geq 0$, $\delta_1+\ldots+\delta_n=1$. The pair $(\boldsymbol{\pi},\boldsymbol{\delta})$ is called a mixed strategy selection.

Payoffs in games with mixed strategies (Definition)

If the players select mixed strategies, then the payoffs are understood as expected payoffs from their pure strategies. In other words, the payoff for the first is $\sum_{ij}\pi_i\delta_j P_{ij}$, and for the second is $\sum_{ij}\pi_i\delta_j D_{ij}$.

Note

A 'traditional' game (with pure strategies) can be interpreted as a mixed-strategy game where probabilities can be either 0 or 1

Note

Nash equilibrium definition can be generalized for mixed strategies. In other words, a pair of strategies $(\boldsymbol{\pi^0},\boldsymbol{\delta^0})$ is a Nash equilibrium, if

- $\sum_{ij}\pi_i^0\delta_j^0P_{ij}=\max_\pi\{\sum_{ij}\pi_i\delta_j^0P_{ij}\}$, and
- $\sum_{ij}\pi_i^0\delta_j^0D_{ij}=\max_\delta\{\sum_{ij}\pi_i^0\delta_jD_{ij}\}$.

Theorem

For every non-zero two-person game there exists a Nash equilibrium for mixed strategies (proof can be derived from the Brouwer's fixed-point theorem).

Example

The following game has a Nash equilibrium in mixed strategies (even though it does not have a Nash equilibrium in pure strategies):

		M	
		Y	N
F	Y	(4,2)	(-2,3)
	N	(2,1)	(-1,0)

Calculating mixed-strategy Nash Equilibrium

If p is the probability of choosing Y for the player F, and q is the probability of choosing Y for the player M, then (1/2,1/2;1/3,2/3) is the Nash equilibrium in mixed strategies (when p=1/2, and q=1/3 the players do not have incentives to unilaterally change these probabilities).

In this example:

$$\pi_1 = p,\ \pi_2 = 1\text{-}p,\ \delta_1 = q, \text{ and } \delta_2 = 1\text{-}q$$

Calculating mixed-strategy Nash Equilibrium (cont.)

$E(P)=4pq-2p(1-q)+2(1-p)q-(1-p)(1-q)=$
$=4pq-2p+2pq+2q-2pq-1+q+p-pq=-p+3q+3pq+1.$

$E(D)=2pq+3p(1-q)+(1-p)q=$
$=2pq+3p-3pq+q-pq=3p+q-2pq$

$\partial E(P)/\partial p=0$
$\partial E(D)/\partial q=0$

$-1+3q=0;\ q=1/3$
$1-2p=0;\ p=1/2.$

Be careful!

Calculating mixed-strategy Nash Equilibrium (cont.)

In general one needs to write the expected value of the payoff for the first player: $\sum_{ij}\pi_i\delta_j P_{ij}$; and for the second player: $\sum_{ij}\pi_i\delta_j D_{ij}$; and to find maxima of these expressions (with respect to π of the former, and with respect to δ of the latter).

<u>Two-stage game (Definition)</u>

A decision situation where in the second stage players react to their moves disclosed in the first stage

<u>Sequential game (Definition)</u>

A series of decision situations where players react to their moves disclosed in earlier stages

Example of a two-stage game

- Stage I:
  Firm F (the 'entrant') decides whether (E) or not (D) to enter a market; firm I (the 'incumbent') enjoys a monopolistic position and earns profit of 4; in this stage plans of the firms what to do in the second stage are not binding
- Stage II:
  Both firms decide whether to cooperate (C) or to attack (A)

Example of a two-stage game (cont.)

The game has the following payoff matrix:

		Incumbent (I)	
		C if F enters	A if F enters
Entering Firm (F)	D&C	0,4	**0,4**
	D&A	0,4	**0,4**
	E&C	**3,1**	-2,-1
	E&A	1,-2	-3,-1

Example of a two-stage game (cont.)

This two-stage game has three Nash equilibria (D&C,A), (D&A,A), and (E&C,C) with payoffs (0,4), (0,4), and (3,1), respectively. The first two are not likely to be observed in practice (I's threat that it will attack if F enters is not credible).

Example of a two-stage game (cont.)

The second stage has the following payoff matrix:

		Incumbent (I)	
		C	A
Entering Firm (F)	C	**3,1**	-2,-1
	A	1,-2	-3,-1

The only Nash equilibrium in this game is (C,C).

Definition

A subgame is a part of a sequential game formed by eliminating its initial stages (this can be mathematically defined more precisely). Strategies available in the original game have components related to any of its stages. It is obvious how to understand a strategy induced in a subgame (consisting of components related to respective stages).

Subgame-perfect Nash equilibrium (SPNE)
SPNE is a Nash equilibrium of a sequential game whose components that are induced in all relevant subgames are their respective Nash equilibria

Example

Let us consider the sequential game of entry (studied before). In the first stage Entrant makes a decision whether to enter or not. In the latter case the payoffs are: 0 for the Entrant, and 4 for the Incumbent. In the former (i.e. in the second stage) the rivals play a simultaneous game with the 2x2 payoff matrix as in the example).

Example (cont.)

The game has two subgames: the original game, and the second stage (assuming that the entrant chooses to enter). By screening all possible scenarios, one concludes that the original game has three Nash equilibria.

Example (cont.)

In the subgame consisting of the second stage alone there is only one Nash equilibrium composed of strategies "cooperate" for both players. This Nash equilibrium is induced by one of the three Nash equilibria of the original game. Therefore there is only one SPNE (the other Nash equilibria do not induce strategies that prove to be rewarding in the second stage).

Questions:

Q-10 The SPNE concept eliminates Nash equilibria

[a] that do not induce Nash equilibria in all stages of the original game

[b] that are not Pareto optima

[c] which assume the players to make decisions simultaneously

[d] which assume the players to make decisions sequentially

[e] none of the above

Exercises:

E-10 Please design a sequential game with a Nash equilibrium that is not an SPNE.

Industrial Organisation

- Does Nash equilibrium explain economic behaviour in any circumstances?

- <u>*Network markets*</u> – as an example of an 'atypical' behaviour
  - Compatibility and product standards
  - External effects of consumption
  - *Switching costs* and loyalty programmes
  - Production scale effects

Discrete Hotelling model

1. Firms a and b produce a discernible product. There are n_a consumers who prefer a, and n_b who prefer b. Production costs are zero.
2. Consumer demand functions are unitary, and the disutility from consuming the non-preferred product is $\delta>0$.
3. The utility of a type a consumer is:
   - $U_a = -p_a$ when buying from the supplier a
   - $U_a = -p_b-\delta$ when buying from the supplier b
4. The utility of a type b consumer is:
   - $U_b = -p_b$ when buying from the supplier b
   - $U_b = -p_a-\delta$ when buying from the supplier a

Discrete Hotelling model (cont.)

5. Therefore the numbers n_a and n_b of consumers who buy from a and b are, respectively:
   - $q_a = 0$, if $p_a > p_b + \delta$,
   - $q_a = n_a$, if $p_b - \delta \le p_a \le p_b + \delta$,
   - $q_a = n_a + n_b$, if $p_a < p_b - \delta$,
   - $q_b = 0$, if $p_b > p_a + \delta$,
   - $q_b = n_b$, if $p_a - \delta \le p_b \le p_a + \delta$,
   - $q_b = n_a + n_b$, if $p_b < p_a - \delta$

Theorem

There is no Nash equilibrium in the discrete Hotelling model

Proof:

On the contrary, let us assume that (p_a,p_b) is a Nash equilibrium. Then one of the following three conditions holds:

1. $|p_a-p_b| > \delta$,
2. $|p_a-p_b| < \delta$,
3. $|p_a-p_b| = \delta$.

It turns out that in each of these cases, one of the suppliers has a motivation to undercut the rival's price. Thus this is not a Nash equilibrium.

<u>Definition</u> (*Undercut-proof prices*)
Supplier a undercuts b if $p_a<p_b-\delta$ (i.e. deprives it of potential buyers)

Prices (p_a^U,p_b^U) are undercut-proof, if:

1. p_a^U is the maximum price that for given p_b^U and q_b^U satisfies:
$$\Pi_b^U=p_b^U q_b^U \geq (p_b^U-\delta)(n_a+n_b)$$
2. p_b^U is the maximum price that for given p_a^U and q_a^U satisfies:
$$\Pi_a^U=p_a^U q_a^U \geq (p_a^U-\delta)(n_a+n_b)$$

Theorem (*Undercut-proof equilibrium*)

The following pair makes an undercut-proof equilibrium:

- $p_a = \delta(n_a+n_b)(n_a+2n_b)/((n_a)^2+n_a n_b+(n_b)^2)$,
- $p_b = \delta(n_a+n_b)(n_b+2n_a)/((n_a)^2+n_a n_b+(n_b)^2)$

Corollary

If $n_a=n_b$, then $p_a^U=p_b^U=2\delta$

Questions:

Q-11 Industrial organization analyses predicted that introducing mobile phone number portability

[a] would inflate SMS prices as a result of increasing technical sophistication of mobile phone operations

[b] would equal SMS and MMS prices as a result of increased competition

[c] would deflate average service prices as a result of decreased switching costs

[d] would deflate average service prices as a result of decreasing technical sophistication of mobile phone operations

[e] none of the above

Exercises:

E-11 Please explain why the undercut-proof equilibrium in the discrete Hotelling model is not a Nash equilibrium.

Portfolio selection

- The problem of a consumer who contemplates what to do with the money.
- The rate of return on alternative investment projects are considered random variables with known means and variances.
- It is assumed that – as a rule – if one wants to achieve a higher rate of return, one needs to face a higher risk (measured e.g. by the variance of the rate)

Problem:

- How to take into account the rate of return and risk simultaneously?
- Risk to be measured with standard deviation (variance) of the rate of return of an asset

Mean-variance model

Assumptions:

Two assets with different mean rates of return (r), and variances of the respective rate of return (σ^2)

- Risk-free: $r_F > 0$, and $(\sigma_F)^2 = 0$ ($\sigma_F = 0$)
- Risky: $r_R > r_F$, and $(\sigma_R)^2 > 0$ ($\sigma_R > 0$)

Mean-variance model (cont.)

Optimization problem:

- For a given level of expected rate of return find a portfolio with the least variance; or
- For a given level of variance (riskiness) find a portfolio with the highest expected rate of return

Mean-variance model (cont.)

Solution:

Let α be a portion of the money spent on the risk-free asset ($1-\alpha$ is the portion of money spent on the risky asset). If $\alpha=1$ then only the risk-free asset is in the portion of the portfolio

- Expected rate of return on both assets combined is:

$$r = \alpha r_F + (1-\alpha) r_R, \text{ and}$$

- Variance of rate of return is (for uncorrelated assets):

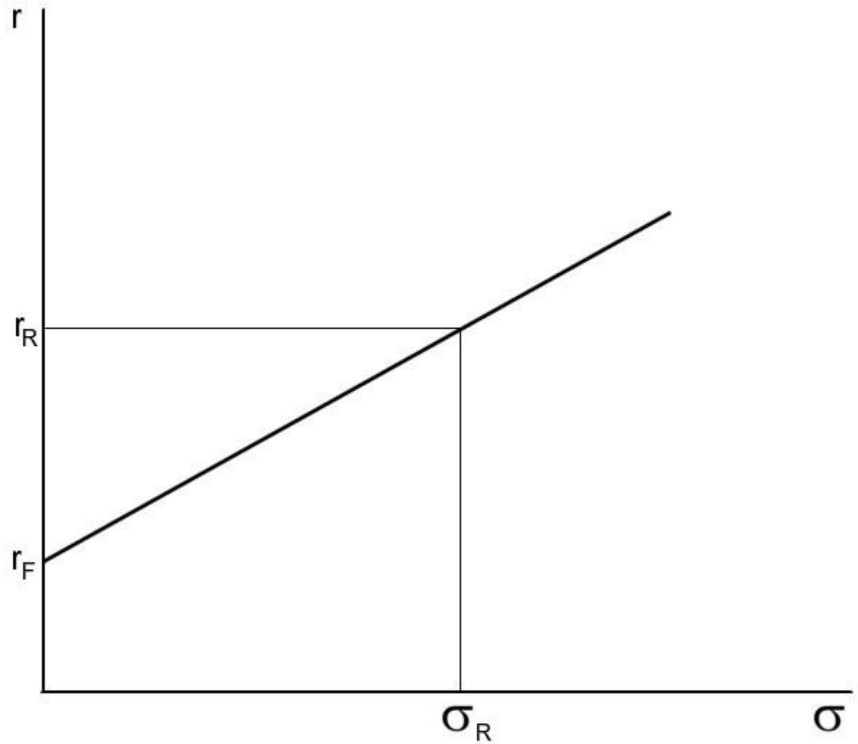
$$\sigma^2 = \alpha^2(\sigma_F)^2+(1-\alpha)^2(\sigma_R)^2 = (1-\alpha)^2(\sigma_R)^2$$

$$(\sigma=(1-\alpha)\sigma_R)$$

Mean-variance model (cont.)

Corollary: Any solution (with rate of return between r_F and r_R, and risk measured by standard deviation not higher than σ_R) can be achieved by taking a convex combination of the risk-free and the risky asset.

Mean-variance model (cont.)



Mean-variance model (cont.)

Extension: It can be assumed that $\alpha<0$ corresponds to the fact that the risk-free asset can be sold rather than bought (money can be borrowed). Then any expected rate of return $r>r_R$ can be achieved (at the cost of variance increased beyond $(\sigma_R)^2$). This paradoxical finding is based on the fact that the rate of return is calculated with respect to the money owned, not with respect to the money borrowed.

Mean-variance model (cont.)

There are two risky assets A_i with respective rates of return and variances: r_i, and $(\sigma_i)^2$ for i=1,2. They are combined with weights α_1 and α_2, where $\alpha_2=1-\alpha_1$. The rate of return of this combination is given by the formula:

$$r(A_1,A_2) = \alpha_1 r_1+\alpha_2 r_2,$$

and the variance is given by the formula:

$$var(A_1,A_2) = (\alpha_1)^2(\sigma_1)^2+2\alpha_1\alpha_2 cov(A_1,A_2)+(\alpha_2)^2(\sigma_2)^2.$$

Mean-variance model (cont.)

Corollary: if α_1, α_2>0 then

$$\mathrm{var}(A_1,A_2) = (\alpha_1)^2(\sigma_1)^2+(\alpha_2)^2(\sigma_2)^2, \text{ if } \mathrm{cov}(A_1,A_2)=0,$$
$$\mathrm{var}(A_1,A_2) > (\alpha_1)^2(\sigma_1)^2+(\alpha_2)^2(\sigma_2)^2, \text{ if } \mathrm{cov}(A_1,A_2)>0,$$
$$\mathrm{var}(A_1,A_2) < (\alpha_1)^2(\sigma_1)^2+(\alpha_2)^2(\sigma_2)^2, \text{ if } \mathrm{cov}(A_1,A_2)<0.$$

- Variance (riskiness) can be lowered by investing in assets that are negatively correlated

Example:

Consumer contemplates investing into:
(1) umbrella; or
(2) ice-cream
Businesses, but profitability depends on weather:
(A) sunny; or
(B) rainy
The weather is a random variable with 50%-50% probabilities (impossible to predict at the time of investment).

Example (cont.):

Expected rate of return on investment

	Rainy	Sunny
Umbrella	6%	2%
Ice-cream	2%	6%

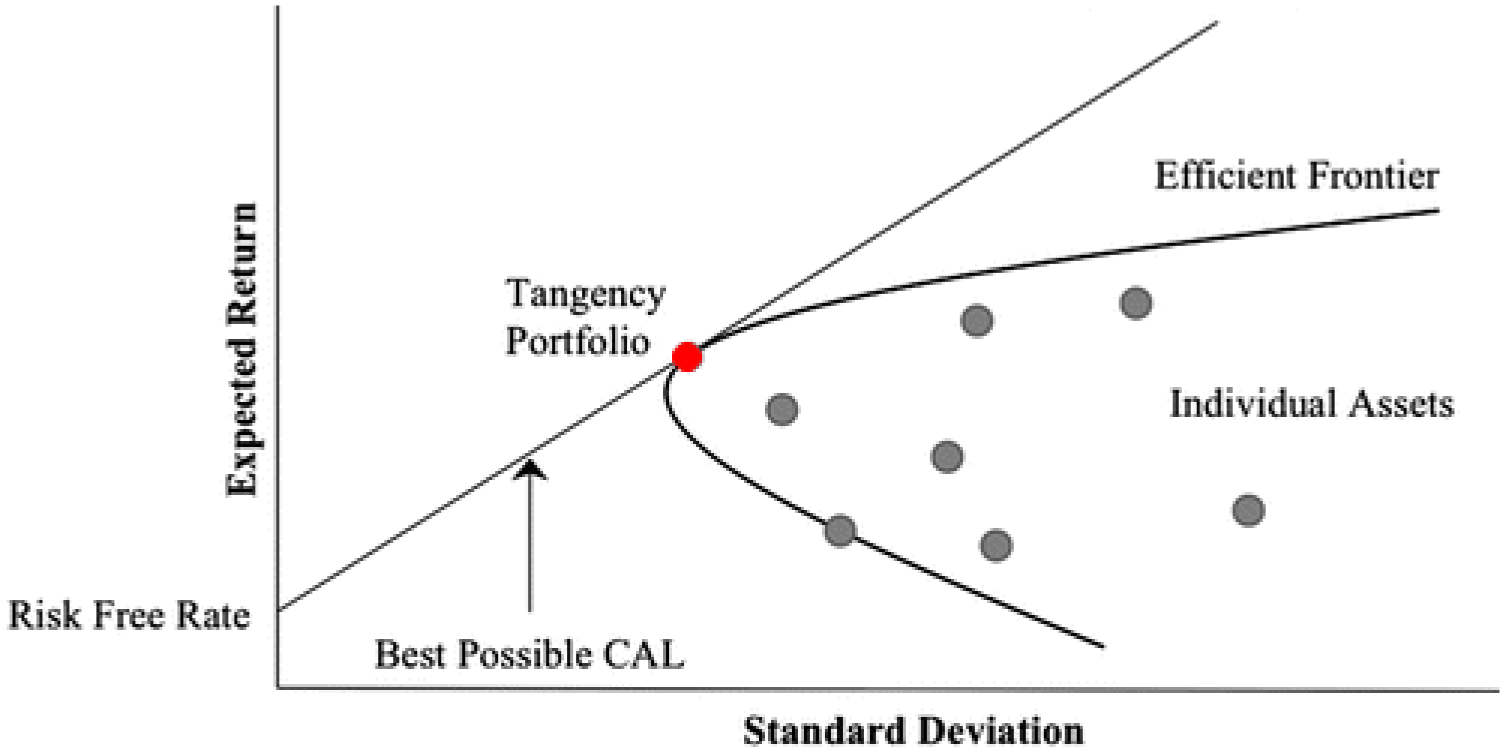
If the consumer invests 50% of money into the umbrella and 50% into the ice-cream business, the expected rate of return will be 4%, and variance will be 0.

Markowitz model

Mean-variance model for two assets can be generalized for any number of assets

When plotted in the r-σ plane, formulae for the expected rate of return and the variance of the combination of assets define a parabola (the "efficiency frontier")

Markowitz model (cont.)



where:

- "tangency portfolio" represents the best combination of assets (given an opportunity to invest in a risk-free asset)
- CAL – Capital Allocation Line

Markowitz model (cont.)

- Explains how to optimally choose a combination of risky assets to invest.
- If a risk-free asset is available, and money can be borrowed, CAL indicates the highest rate of return possible given the variance of a combination of assets
- The rate of return (and variance) indicated by the red dot can be achieved by investing all the money owned in the "tangency portfolio" (an optimum combination of assets)
- A higher rate of return (at the cost of a higher variance) can be achieved by borrowing money (investing a "negative" amount in the risk-free asset)

Questions:

Q-12 If one invests in two risky assets then

[a] the total risk must increase

[b] the total risk will always be lower

<u>[c] the total risk can be higher if the returns on these assets are positively correlated</u>

[d] the total risk is higher than for the more risky asset

[e] none of the above

Exercises:

E-12 Please recommend an investment strategy if there is a possibility to earn a risk-free rate of 2% and a possibility to earn 6% subject to some risk, measured by a standard deviation σ of the rate of return.

Introduction to financial economics

There are a number of formal assumptions listed in financial economics. Many of them are so obvious that they do not need to be listed. The most important (and a controversial) one is the no-arbitrage assumption. It can be seen as equivalent to the one-price law or to an assumption that there is no possibility of enjoying profits without risk.

Note

The *no-arbitrage* assumption can be reconciled with the assumption that there are *arbitrageurs*, i.e. economic agents who try to make money on arbitrage (and they succeed occasionally). In other words, the reason that – as a rule – there is no arbitrage, is because there are arbitrageurs who try to take advantage of asymmetric information, buy at a lower price, and sell at a higher one immediately. This cannot continue for ever. Hence whenever there is such an opportunity, it disappears quickly.

Definition

A derivative instrument ("derivative") is a contract with a payoff which depends on an (*ex ante*) unknown realization of a random variable.

- The random variable does not have to be a purely economic phenomenon (e.g. weather)
- A typical random variable is economic (e.g. exchange rate, performance index, asset price, etc.). In the Black-Scholes model it is the price of an underlying stock

Definition

- An option is a contract which gives its party the right (but not the obligation) to buy or sell a commodity at a predetermined price. A "call option" ("call") gives the right to buy, and a "put option" ("put") – the right to sell.
- In a European option, the right is to be executed at a given date (typically on the third Friday of a given month). In an American option, the right is to be executed anytime prior to the given date. In a Bermudan option, the right is subject to specific conditions.
- Options can be seen as derivatives, since the execution of the right depends on the market price of the commodity

Definition

The risk-free asset is called money. Its value increases over time at the pace of the discount rate. If B_t is the value of money at time t then its value at time τ is:

$$B_\tau = e^{r(\tau - t)} B_t$$

where r is a discount rate. Please note that if $\tau = t$, then – according to the formula – $B_\tau = B_t$. If $r \geq 0$, then $B_\tau > B_t$ for $\tau > t$, and $B_\tau < B_t$ for $\tau < t$. This is not risk-free "profiting" since – by definition of the discount rate (see lecture 4) – economic agents are indifferent between B_τ at time τ and B_t at time t.

Definition

Short selling – selling an asset that is not owned by the seller. This can be interpreted as borrowing from its owner, and selling the right to use the asset. If the owner asks for the asset, then it needs to be bought in the market and given back to the owner. By short selling, economic agents expect that the market price of the asset will decrease and thus they will make a profit (having returned the asset to where they borrowed from, but keeping the difference between the original higher and final lower price).

Definition

Long position – buying an asset whose value is expected to rise. Investors in 'long position' expect to make a profit by selling the asset when its price is higher than the original one. With respect to options:

- Long position on call options – investor expects the price of the underlying asset to increase
- Long position on put options – investor expects the price of the underlying asset to decrease

Definitions

- Strike price (exercise price) – the price indicated in a European option (for call options: the holder has the right to buy the underlying asset at this price; for put options: the holder has the right to sell at this price)
- Spot price – the price established by the market
- The call option is Out-of-The Money (OTM) if the spot price of the underlying asset is lower than the strike price. It is In-The-Money (ITM) if the spot price of the underlying asset is higher than the strike price
- For put options it is the other way around

Notation

- S – the price of the underlying stock,
- V(S,t) – the price of a derivative,
- C(S,t) – the price of a European call option,
- P(S,t) – the price of a European put option,
- K – the strike price of the option,
- σ – the standard deviation of the stock's returns,
- T – expiry date ("maturity"; established in the option).

Black-Scholes equation

$$\partial V/\partial t + (1/2)\sigma^2 S^2 \partial^2 V/\partial S^2 + rS\partial V/\partial S = rV$$

Partial differential equation established as a result of analysing how to optimally manage investment risk (accomplished through "hedging", i.e. combining options and shares in the underlying stock)

Black-Scholes formula

For European call options:

$$C(S_t,t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)},$$

For European put options:

$$P(S_t,t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t$$

where:

$$d_1 = 1/(\sigma(T-t)^{1/2})(\ln(S_t/K)+(r+\sigma^2/2)(T-t))$$

$$d_2 = d_1 - \sigma(T-t)^{1/2}$$

N is the normal distribution function, i.e.

$$N(x) = 1/(2\pi)^{1/2} {}_{-\infty}\int^{x}(\exp(-z^2/2))dz$$

Black-Scholes formula (cont.)

- Derived from the Black-Scholes equation
- Based on the assumption of the normal distribution of random effects
- Extensions take into account:
  - Non-normal distributions
  - Relaxing classical assumptions (such as zero transaction costs, no taxes, etc.)

Black-Scholes formula (cont.)

Establishes European call and put option unique prices based on:

- S_t – the current price of the underlying stock, and
- K – the strike price of the option,

It is not based on S_T (i.e. the stock price at the time of maturity of the option)!

Practical applications risky because of rigorous probabilistic assumptions required

Questions:

Q-13 The Black-Scholes model assumes that
[a] there are no arbitrageurs
<u>[b] there is no possibility of arbitrage</u>
[c] stock prices are not random
[d] trading derivatives is banned
[e] none of the above

Exercises:

E-13 A so-called Zero Coupon Bond with maturity T is an asset that pays 1 at time T (in the meantime it does not pay anything). Z(t) is its price at time t, and the discount rate is r (money can be borrowed at the rate r). Please prove that $Z(t)=e^{-r(T-t)}$.

Econophysics

Econophysics is an approach to microeconomic problems by assuming that economic objects behave like molecules (studied by physics), and therefore the results of their behaviour can be modelled by equations analyzed in physics.

Hit diffusion equation

In particular, it is assumed that certain economic phenomena can be modelled by the well-known in mathematics *Hit Diffusion Equation*:

$$\partial u/\partial t = \alpha(\partial^2 u/\partial x^2+\partial^2 u/\partial y^2+\partial^2 u/\partial z^2)$$

Hit diffusion equation (cont.)

Sometimes the equation is written in the form:

$$\partial u/\partial t = \alpha\nabla^2 u,$$

where

∇ is a so-called Laplace operator. In a 3-dimensional space: $\nabla^2 u = \partial^2 u/\partial x^2+\partial^2 u/\partial y^2+\partial^2 u/\partial z^2$

Hit diffusion equation (cont.)

The equation models temperature changes u(x,y,z,t) of a spatial object if there is no external source of heat. The temperature varies in time t, and in all spatial dimensions x, y, and z. If there is an external source of heat, an additional term is added at the right-hand-side of the equation.

Hit diffusion equation (cont.)

The model is often simplified by assuming that there is only one spatial dimension:

$$\partial u/\partial t = \alpha\partial^2 u/\partial x^2$$

In this case, the equation predicts the distribution of temperature in a one-dimensional object (a rod) which is perfectly insulated (there is no exchange of heat with its environment).

Hit diffusion equation (cont.)

- Financial economic models apply this approach by assuming that assets are priced in processes similar to those which determine distribution of the temperature in physical objects.
- The no-arbitrage assumption corresponds to the First Law of Thermodynamics.

Brownian motion

- Random movement of molecules in fluids
- Used in economics in order to model the movement of economic variables considered random

Brownian motion (cont.)

- Random walk
- Wiener process (continuous version of a random walk)

Wiener process (definition)

- $W_0 = 0$ almost surely (with probability 1 $W_0=0$),
- W has independent increments: for every t>0, the future increments $W_{t+u}-W_t$, $u \geq 0$, are independent of the past values W_s, $s<t$,
- W has Gaussian increments: $W_{t+u}-W_t$ is normally distributed with mean 0 and variance u, i.e. $W_{t+u}-W_t \sim N(0,u)$; in other words, the variance of an increment grows with its length, but its mean is zero,
- W has continuous paths almost surely (with probability 1 W_t is continuous in t).

Wiener process (with or without drift)

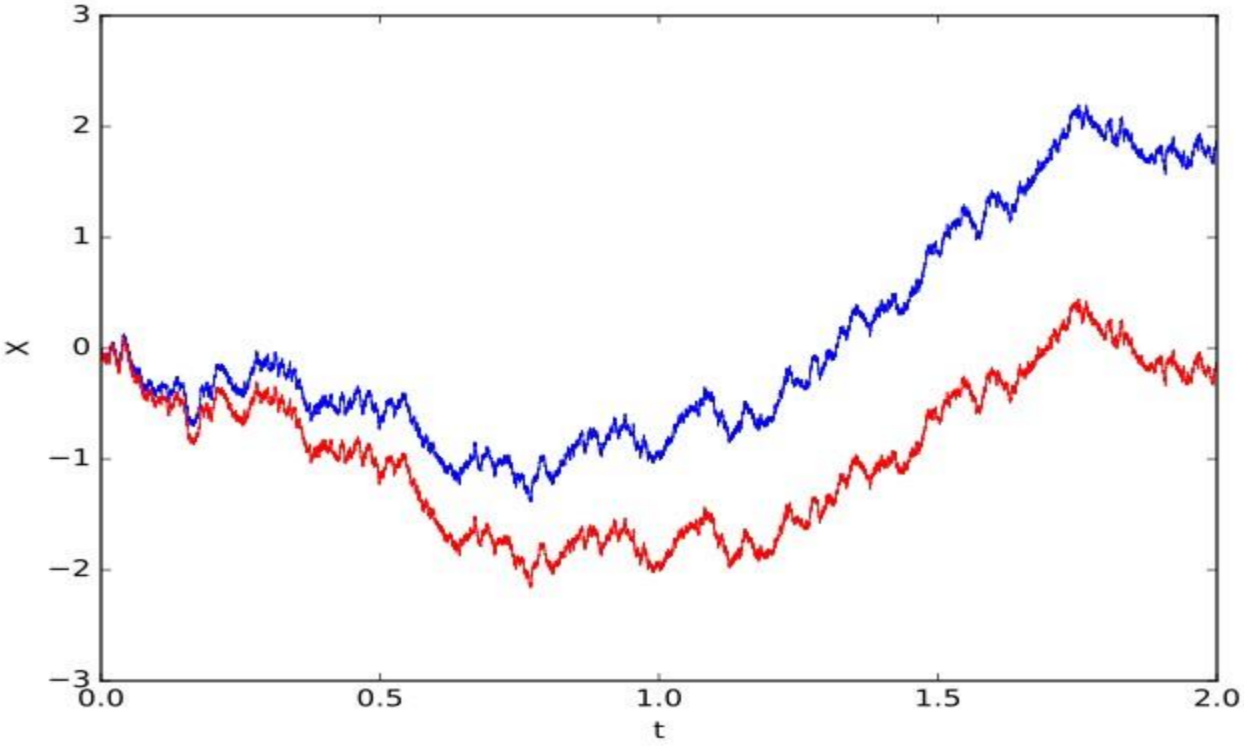
Stochastic process defined as:

$$X_t = \mu t + \sigma W_t,$$

where:

- W – Wiener process
- μ – drift (long-term trend)
- σ^2 – infinitesimal variance (terminology of stochastic processes; interpreted as "volatility" – the standard deviation is proportional to time t elapsed since start: $t\sigma$)

Wiener process (illustration)



- Red line – without drift
- Blue line – with drift

Wiener processes in financial economics

In order to establish the Black-Scholes model, the following Stochastic Differential Equation (so-called Geometric Brownian Motion) was used:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

whose solution is:

$$S_t = S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t)$$

Wiener processes in financial economics (cont.)

- W can be interpreted as the underlying economic variable (e.g. the return on an asset); and
- S can be interpreted as an index characterising a derivative instrument (e.g. a European option)

Questions:

Q-14 The econophysics approach to microeconomic modelling
[a] refers to the materials balance paradigm as advocated for by natural scientists
[b] assumes that some consumers spend their financial resources irrationally
<u>[c] considers stock price movements as resembling Brownian motion</u>
[d] assumes that population growth follows Brownian motion
[e] none of the above

Exercises:

E-14 Discuss pros and cons for assuming that one-dimensional heat diffusion equation can be used for modelling the supply of a product.

The Kuhn-Tucker theorem

It extends the 18th and 19th century findings on constrained maxima to account for non-strict inequalities (the original analyses dealt either with strict inequalities or equalities).

<u>Definition</u> (conditional maximum or minimum)
A function f: $\Re^n \to \Re$ has a conditional maximum (minimum) subject to $g_1(\mathbf{x})=b_1,\ldots,g_m(\mathbf{x})=b_m$ for $\mathbf{x}^* \in \Re^n$, if for any $\mathbf{x} \in \Re^n$ satisfying $g_1(\mathbf{x})=b_1,\ldots,g_m(\mathbf{x})=b_m$ there is: $f(\mathbf{x}) \leq (\geq) f(\mathbf{x}^*)$. The problem of finding a conditional maximum (minimum) is coded as:

$$\mathrm{Max(Min)}_{\mathbf{x}}\{f(\mathbf{x}): g_1(\mathbf{x})=b_1, \ldots, g_m(\mathbf{x})=b_m\}.$$

Definition (Lagrange function)

For the conditional maximum problem the Lagrange function L: $\Re^{n+m} \rightarrow \Re$ is defined as:

$$L(\mathbf{x},\lambda_1,\ldots\lambda_m)=f(\mathbf{x})+\lambda_1(b_1-g_1(\mathbf{x}))+\ldots+\lambda_m(b_m-g_m(\mathbf{x}))$$

Variables $\lambda_1,\ldots,\lambda_m$ are called "Lagrange multipliers".

Theorem (Lagrange)

Let us assume that functions f, $g_1,\ldots,g_m$ are continuously differentiable and the matrix of derivatives of $g_1,\ldots,g_m$ (the Jacobi matrix) has the rank m. Then:

1. If $\mathbf{x}^*$ is the solution of the conditional maximum problem, then there exist $\lambda^*_1,\ldots,\lambda^*_m$, such that $(\mathbf{x}^*,\lambda^*_1,\ldots,\lambda^*_m)$ solve system of equations $\partial L/\partial \mathbf{x}=\mathbf{0}$, $\partial L/\partial \boldsymbol{\lambda}=\mathbf{0}$.
2. If functions f, $g_1,\ldots g_m$ are twice continuously differentiable, and matrix of second derivatives $\partial^2 L/\partial \mathbf{x}^2$ (the Hesse matrix of L) is negatively definite, then the point $\mathbf{x}^*$ from the solution $(\mathbf{x}^*,\lambda^*_1,\ldots,\lambda^*_m)$ of the system $\partial L/\partial \mathbf{x}=\mathbf{0}$, $\partial L/\partial \boldsymbol{\lambda}=\mathbf{0}$ solves the original problem.

Note

Equations $\partial L(\mathbf{x}^*,\lambda^*_1,\ldots,\lambda^*_m)/\partial\mathbf{x}=\mathbf{0}$ are equivalent to $\partial f(\mathbf{x}^*)/\partial\mathbf{x}=\lambda^*_1\partial g_1(\mathbf{x}^*)/\partial\mathbf{x}+\ldots+\lambda^*_m\partial g_m(\mathbf{x}^*)/\partial\mathbf{x}$, where the equality applies to each of the coordinates, and $\partial g_j(\mathbf{x}^*)/\partial\mathbf{x}$ are vectors of partial derivatives with respect to $x_1,\ldots,x_n$. Equations $\partial L(\mathbf{x}^*,\lambda^*_1,\ldots\lambda^*_m)/\partial\boldsymbol{\lambda}=\mathbf{0}$ are equivalent to $b_j-g_j(\mathbf{x}^*)=0$ (for $j=1,\ldots,m$).

Note

In the Lagrange theorem (2) for a conditional minimum, the Hesse matrix should be positively definite

Note (interpretation of Lagrange multipliers)

Taking into account that $L(\mathbf{x}^*, \lambda^*_1, \ldots, \lambda^*_m) = f(\mathbf{x}^*)$, then $F(.) = f(\mathbf{x}^*)$ understood as a function of right hand sides of constraints $(b_1, \ldots, b_m)$ yields as derivatives:

$\lambda^*_j = \partial F / \partial b_j$.

Definition (mathematical programming; optimization) $\text{Max}_{\mathbf{x}}\{f(\mathbf{x}): g_1(\mathbf{x})\leq b_1, \ldots, g_m(\mathbf{x})\leq b_m, \mathbf{x}\geq\mathbf{0}\}$ (likewise for the minimum). Compared to the 'classical' conditional maximum, in this definition inequalities substitute for equalities, and variables must be non-negative.

Note

In the optimization problem an equality $\mathbf{g}(\mathbf{x})=\mathbf{b}$ can be represented by a pair of inequalities: $\mathbf{g}(\mathbf{x})\leq\mathbf{b}$ and $-\mathbf{g}(\mathbf{x})\leq-\mathbf{b}$

Note

In an optimization problem, if a variable x_i may be a negative one, then it can be substituted with a pair of non-negative variables x^-_i and x^+_i, such that $x_i=x^+_i-x^-_i$.

Corollary

A conditional maximum (minimum) problem is a special case of an optimization problem

Lemma

If f is twice continuously differentiable, then a solution $\mathbf{x}^*$ to the optimization problem $\text{Max}_{\mathbf{x}}\{f(\mathbf{x}): \mathbf{x}\geq\mathbf{0}\}$ must satisfy the following 2n+1 conditions:

$$\partial f(\mathbf{x}^*)/\partial\mathbf{x}\leq\mathbf{0},\ \mathbf{x}^*\geq\mathbf{0} \text{ and } \partial f(\mathbf{x}^*)/\partial\mathbf{x}\cdot\mathbf{x}^* = 0.$$

The last condition ("complementarity") can be written as:

$$\partial f(\mathbf{x}^*)/\partial x_1\cdot x^*_1+\ldots+\partial f(\mathbf{x}^*)/\partial x_n\cdot x^*_n = 0$$

and – given the other 2n conditions – is equivalent to the following series of implications:

- if $x^*_i>0$, then $\partial f(\mathbf{x}^*)/\partial x_i=0$ and
- if $\partial f(\mathbf{x}^*)/\partial x_i<0$, then $x^*_i=0$.

Note

Inequalities in the optimization problem ($\mathbf{g}(\mathbf{x})\leq\mathbf{b}$) can be substituted with equalities, as in the conditional maximum problem, once m auxiliary variables $s_i\geq0$ are introduced and inserted into the constraints: $\mathbf{g}(\mathbf{x})+\mathbf{s}=\mathbf{b}$. The Jacobi matrix of this new constraints systems has the rank m, since it includes a unitary matrix of that rank). Therefore an appropriate assumption from the Lagrange theorem is satisfied.

Definition (differential Kuhn-Tucker conditions for an optimization problem)

- $\partial f(\mathbf{x})/\partial \mathbf{x} \leq \lambda_1 \partial g_1(\mathbf{x})/\partial \mathbf{x} + \ldots + \lambda_m \partial g_m(\mathbf{x})/\partial \mathbf{x}$,
- $(\partial f(\mathbf{x})/\partial \mathbf{x} - \lambda_1 \partial g_1(\mathbf{x})/\partial \mathbf{x} + \ldots - \lambda_m \partial g_m(\mathbf{x})/\partial \mathbf{x}) \cdot \mathbf{x} = 0$ ("complementarity" conditions),
- $\mathbf{x} \geq \mathbf{0}$,
- $b_1 - g_1(\mathbf{x}) \geq 0, \ldots, b_m - g_m(\mathbf{x}) \geq 0$,
- $\lambda_1(b_1 - g_1(\mathbf{x})) + \ldots + \lambda_m(b_m - g_m(\mathbf{x})) = 0$ ("complementarity" conditions),
- $\lambda_1, \ldots, \lambda_m \geq 0$.

Corollary

The differential Kuhn-Tucker conditions as necessary for a solution to an optimization problem can be derived from the Lagrange theorem using the lemma and substituting auxiliary non-negative variables $s_1=b_1-g_1(\mathbf{x}), \ldots, s_m=b_m-g_m(\mathbf{x})$. (The problem to be solved can be then written as $\mathrm{Max}_{\mathbf{x},\mathbf{s}}\{f(\mathbf{x})\colon g(\mathbf{x})+\mathbf{s}=\mathbf{b},\ \mathbf{x}\geq\mathbf{0},\ \mathbf{s}\geq\mathbf{0}\}$.)

Corollary

If the definition of the Lagrange function is generalized for optimization problems, then differential Kuhn-Tucker conditions can be written as:

- $\partial L/\partial \mathbf{x} \leq \mathbf{0}$,
- $(\partial L/\partial \mathbf{x})\cdot\mathbf{x}=0$ ("complementarity"),
- $\mathbf{x} \geq \mathbf{0}$,
- $\partial L/\partial\lambda_1 \geq 0,\ldots,\partial L/\partial\lambda_m \geq 0$,
- $\lambda_1 \partial L/\partial\lambda_1+\ldots+\lambda_m \partial L/\partial\lambda_m=0$ ("complementarity"),
- $\lambda_1 \geq 0,\ldots,\lambda_m \geq 0$.

Definition (saddle point)

A function h: $\Re^{n+m} \rightarrow \Re$ has a non-negative saddle point $(\mathbf{y}^*, \mathbf{z}^*)$, if for every $\mathbf{y} \in \Re^n$, $\mathbf{y} \geq \mathbf{0}$ and every $\mathbf{z} \in \Re^m$, $\mathbf{z} \geq \mathbf{0}$ there is: $h(\mathbf{y}, \mathbf{z}^*) \leq h(\mathbf{y}^*, \mathbf{z}^*) \leq h(\mathbf{y}^*, \mathbf{z})$.

Theorem (Kuhn-Tucker)

Solutions to $\mathrm{Max}_{\mathbf{x}}\{f(\mathbf{x}): g_1(\mathbf{x})\leq b_1, \ldots, g_m(\mathbf{x})\leq b_m, \mathbf{x}\geq\mathbf{0}\}$ can be characterized in the following way:

1. If $(\mathbf{x}^*,\lambda^*_1,\ldots\lambda^*_m)$ is a non-negative saddle point of the Lagrange function, then $\mathbf{x}^*$ is a solution to the original optimization problem;
2. If f is a concave function, and $g_1,\ldots,g_m$ are convex functions, and if there is $\mathbf{x}^0\geq 0$ such that $g_1(\mathbf{x}^0)<b_1, \ldots, g_m(\mathbf{x}^0)<b_m$ (so-called Slater "constraint qualification"), then for every solution $\mathbf{x}^*$ to the original optimization problem, there exist $\lambda^*_1,\ldots\lambda^*_m$ such that $(\mathbf{x}^*,\lambda^*_1,\ldots\lambda^*_m)$ is a non-negative saddle point of the Lagrange function.

Proof:

(1) Let $(\mathbf{x}^*, \lambda^*_1, \ldots \lambda^*_m)$ be a non-negative saddle point of the Lagrange function. Hence for all $\mathbf{x} \geq \mathbf{0}$:

$$f(\mathbf{x}) + \lambda^*_1(b_1 - g_1(\mathbf{x})) + \ldots + \lambda^*_m(b_m - g_m(\mathbf{x})) \leq$$
$$f(\mathbf{x}^*) + \lambda^*_1(b_1 - g_1(\mathbf{x}^*)) + \ldots + \lambda^*_m(b_m - g_m(\mathbf{x}^*))$$

and for all $\lambda^*_1, \ldots \lambda^*_m \geq 0$

$$f(\mathbf{x}^*) + \lambda^*_1(b_1 - g_1(\mathbf{x}^*)) + \ldots + \lambda^*_m(b_m - g_m(\mathbf{x}^*)) \leq$$
$$f(\mathbf{x}^*) + \lambda_1(b_1 - g_1(\mathbf{x}^*)) + \ldots + \lambda_m(b_m - g_m(\mathbf{x}^*))$$

The second inequality can be written as

$(\lambda_1 - \lambda^*_1)(b_1 - g_1(\mathbf{x}^*)) + \ldots + (\lambda_1 - \lambda^*_m)(b_m - g_m(\mathbf{x}^*)) \geq 0$ for $\lambda_1, \ldots, \lambda_m \geq 0$. Since $\lambda_1, \ldots, \lambda_m$ can be arbitrarily large, the following inequalities should hold:

$g_1(\mathbf{x}^*) \leq b_1, \ldots, g_m(\mathbf{x}^*) \leq b_m$.

Proof (cont.):

On the other hand, by substituting $\lambda_1=...=\lambda_m=0$ we will get $\lambda^*_1(b_1-g_1(\mathbf{x}^*))+...+\lambda^*_m(b_m-g_m(\mathbf{x}^*)) \leq 0$. However, since every number in the parentheses is non-negative, the inequality is simply an equality:

$\lambda^*_1(b_1-g_1(\mathbf{x}^*))+...+\lambda^*_m(b_m-g_m(\mathbf{x}^*)) = 0$.

By substituting this equality to the first inequality of the saddle point definition we get:

$f(\mathbf{x}^*)\geq f(\mathbf{x})+\lambda^*_1(b_1-g_1(\mathbf{x}))+...+\lambda^*_m(b_m-g_m(\mathbf{x}))$

for all $\mathbf{x}\geq\mathbf{0}$. If additionally $\mathbf{x}$ satisfies the constraints, then $f(\mathbf{x}^*)\geq f(\mathbf{x})$ (since $\lambda^*_1,...\lambda^*_m\geq 0$), which completes this part of the proof.

Proof (cont.):

(2) Let $\mathbf{x}^*$ be a solution to the optimization problem, so $\mathbf{x}^* \geq \mathbf{0}$, $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{b}$ and $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x}$ satisfying $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{g}(\mathbf{x}) \leq \mathbf{b}$ ($\mathbf{g}=[g_1,...,g_m]'$, $\mathbf{b}=[b_1,...,b_m]'$). Let us define two sets in $\Re^{m+1}$: $A=\{(a_0,\mathbf{a}): \exists \mathbf{x} \geq \mathbf{0}\ [a_0 \leq f(\mathbf{x}) \text{ and } \mathbf{a} \leq \mathbf{b} - \mathbf{g}(\mathbf{x})]\}$ and $C=\{(c_0,\mathbf{c}): c_0 > f(\mathbf{x}^*) \text{ and } \mathbf{c} > \mathbf{0}\}$, where $a_0, c_0 \in \Re$, $\mathbf{a},\mathbf{c} \in \Re^m$ (column vectors). The set C is convex as an interior of a convex set. From the concavity of f and convexity of **g** we know that the set A is also convex (this is fairly easy to verify). As $\mathbf{x}^*$ is a solution, then the sets are disjoint.

<u>Proof</u> (cont.):

Thus by the separation theorem there exists a non-zero vector $(y_0,\mathbf{y})\in \Re^{m+1}$ ($\mathbf{y}$ is a row vector) such that for all $(a_0,\mathbf{a})\in A$ and all $(c_0,\mathbf{c})\in C$ we have: $y_0a_0\leq y_0c_0$ and $\mathbf{y}\cdot\mathbf{a}\leq\mathbf{y}\cdot\mathbf{c}$. From the definition of sets A and C $(y_0,\mathbf{y})\geq\mathbf{0}$ (otherwise, by substituting negative $\mathbf{a}$ and a_0 and positive $\mathbf{c}$ and c_0 we get a contradiction with the separation theorem). The point $(f(\mathbf{x}^*),\mathbf{0})$ belongs to the border of the set C, so all (non-strict) inequalities that hold for C, hold for this point too. Therefore: $y_0a_0\leq y_0f(\mathbf{x}^*)$ and $\mathbf{y}\cdot\mathbf{a}\leq\mathbf{y}\cdot\mathbf{0}$ (i.e. $\mathbf{y}\cdot\mathbf{a}\leq 0$).

Proof (cont.):

Summing up these two inequalities yields: $y_0 a_0 + \mathbf{y} \cdot \mathbf{a} \leq y_0 f(\mathbf{x}^*)$. In particular, by substituting $a_0 = f(\mathbf{x})$ and $\mathbf{a} = \mathbf{b} - \mathbf{g}(\mathbf{x})$, we get $y_0 f(\mathbf{x}) + \mathbf{y} \cdot (\mathbf{b} - \mathbf{g}(\mathbf{x})) \leq y_0 f(\mathbf{x}^*)$. Now we will prove that $y_0 > 0$. From the constraint qualification (there exists $\mathbf{x^0} \geq \mathbf{0}$, for which $\mathbf{b} - \mathbf{g}(\mathbf{x^0}) > \mathbf{0}$ for all coordinates) and from the non-negativity of the vector $\mathbf{y} \neq \mathbf{0}$ we get $\mathbf{y} \cdot (\mathbf{b} - \mathbf{g}(\mathbf{x^0})) > 0$. If $y_0 = 0$, then the inequality $y_0 f(\mathbf{x}) + \mathbf{y} \cdot (\mathbf{b} - \mathbf{g}(\mathbf{x})) \leq y_0 f(\mathbf{x}^*)$ could not have been satisfied.

Proof (cont.):
Knowing that $y_0>0$, then both sides of this inequality can be divided by y_0. Then one gets $f(\mathbf{x})+\mathbf{y}^*\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x}))\leq f(\mathbf{x}^*)$, where $\mathbf{y}^*=\mathbf{y}/y_0$. By substituting $\mathbf{x}=\mathbf{x}^*$ one gets $\mathbf{y}^*\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x}^*))\leq 0$. Given the fact that all elements of this sum are non-negative (as products of non-negative numbers), this in fact implies that $\mathbf{y}^*\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x}^*))=0$. The pair $(\mathbf{x}^*,\mathbf{y}^*)$ is the non-negative saddle point of the Lagrange function (one needs to substitute $\lambda_1=y_1, \ldots, \lambda_m=y_m$), since it is easy to check that $L(\mathbf{x},\mathbf{y}^*)\leq L(\mathbf{x}^*,\mathbf{y}^*)\leq L(\mathbf{x}^*,\mathbf{y})$, i.e. $f(\mathbf{x})+\mathbf{y}^*\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x})) \leq f(\mathbf{x}^*)+\mathbf{y}^*\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x}^*)) \leq f(\mathbf{x}^*)+\mathbf{y}\cdot(\mathbf{b}-\mathbf{g}(\mathbf{x}^*))$.

Note

In the Kuhn-Tucker theorem the concavity/convexity assumptions for f and **g**, as well as constraint qualifications for **g** can be weakened by assuming, respectively, quasi-concavity, quasi-convexity and constraint qualifications for those constraints only whose left-hand side is non-linear (this modified constraint qualification is called Martos condition).

Questions:

Q-15 Thanks to the convexity/concavity assumptions

[a] the solution of the mathematical programming studied must be in the saddle point identified in the Kuhn-Tucker theorem

[b] in order to find the optimum one does not have to solve a system of simultaneous equations

[c] equations to be solved are linear

[d] the number of equations is the same as the number of unknowns

[e] none of the above

Exercises:

E-15 An internal solution $\mathbf{x}^*$ to the constrained maximization problem analysed in this lecture satisfies the following condition: for all i=1,…,n, $x^*_i>0$. Please prove that if $\mathbf{x}^*$ is an internal solution, then "complementarity" in the Kuhn-Tucker differential theorem reads: for any i=1,…,n: $\partial f(\mathbf{x})/\partial x_i-\lambda_1\partial g_1(\mathbf{x})/\partial x_i+...-\lambda_m\partial g_m(\mathbf{x})/\partial x_i = 0$ (thus it is a system of equations).

Outline solutions to exercises

E-1 First we can prove that a concave function does not have two local maxima with different values. In other words, if $x_1$and x_2 are local maxima of such a function then $f(x_1)=f(x_2)$. To see this, let us assume, for instance, that $f(x_1)<f(x_2)$. Then the segment linking $f(x_1)$ and $f(x_2)$ cannot lie entirely below the graph of f, since it must lie above it in some small neighbourhood of x_1. Thus if a concave function has two local maxima they need to be equal, and – moreover – the entire segment linking $f(x_1)$ and $f(x_2)$ must coincide with the graph of f. Therefore what needs to be proved is that for a strictly concave function no portion of the graph can be flat. But this is precisely what the definition of strict concavity includes ($f(\lambda y+(1-\lambda)z)>\lambda f(y)+(1-\lambda)f(z)$).

E-2 Yes. The question remains whether there can be single-bundle indifference curves in non-trivial cases (trivial cases include, for instance, a two-element set of bundles where one is preferred over the other). A two-dimensional lexicographic preference relationship serves as an example. Unless two bundles are exactly identical, one is always strictly preferred to the other one.

E-3 An example referred to in the exercise should demonstrate that if the demand for the good number 1 is not derived from the FOC, then it may depend on income. Indeed, for quasi-linear preferences, i.e. preferences represented by a utility function $u(x_1,x_2)=v(x_1)+x_2$ the demand for x_1 is derived by maximizing $v(x_1)+x_2$ subject to $px_1+x_2 = m$. One can rephrase this as a one-dimensional problem by substituting $x_2=m-px_1$. The problem is then to maximize the function $v(x_1)+m-px_1$. If the maximum is where the derivative vanishes, the FOC condition reads $MU_1=p_1$, i.e. it is independent of m. However, the maximum of a one-dimensional function can be attained at the border of its domain (not where its derivative vanishes). In such a case, the demand may depend on m.

E-4 The windmill project has the following cash flow: X_1 = -1,500,000, and $X_2 = X_3 = \ldots = X_{31}$ = 80,000, since 2,000x40=80,000. By the way, if discounting is ignored, it will pay off after 18 years (in the year 19 – for the first time – the net accumulated revenue of 1,520,000 will be higher than the investment cost of 1,500,000). If a discount rate is taken into account, the analysis becomes more complex. One needs to calculate r^* such that

$$1{,}500{,}000=80{,}000/(1+r^*)+80{,}000/(1+r^*)^2+80{,}000/(1+r^*)^3+\ldots+80{,}000/(1+r^*)^{30}.$$

This is an algebraic equation of the 30th degree. There are no formulae to find its solutions analytically. However there are software packages that find solutions numerically. The problem can be solved using, for instance, Microsoft Excel. The result is 3.3%. If money is available to the investor at the price of, say, 3%, then the project makes sense. If the cost of money is, say, 4%, the project is economically inefficient.

E-5 Let us assume that people understand probabilities and their preferences for deterministic bundles are rational, but they choose as in the Allais paradox, i.e. they define preferences with respect to lotteries not like in the vNM theory. A popular explanation for this phenomenon refers to the so-called Prospect Theory developed by Daniel Kahneman and Amos Tversky. It states that people's decisions are affected by "prospects", that is by what we consider a "starting point", and whether we are to face a gain or a loss. The same amount of money is considered more valuable if we face a prospect of losing it than if we face a prospect of gaining it. Thus if our starting point is getting 1 million dollars then the prospect of having zero instead is felt as an acute loss. If – on the contrary – our starting point is getting zero then the prospect of having 1 million dollars is felt as a nice gain (although not so strongly as the loss of being deprived of 1 million dollars).

E-6 In figures 1 and 2 one can see how consumers can move from X_0 to X^* in one transaction which brings them to a Pareto optimum (if they choose the trading ratio 'one-orange-for-two-apples', i.e. the Walras equilibrium price, and A buys one orange and sells two apples to B). But they can improve their situations simultaneously by doing other exchanges too. For instance, they can agree to evaluate 2 apples for 1 orange (the price ratio is ½, like in the Walras equilibrium), but to swap one apple only (for a half of an orange). If they do so, then they can move from X_0 to (7,2½,3,2½). This is a Pareto improvement, as it provides both agents with higher utilities (the first with 7x2½=16½ instead of 8x2=16, and the second with 3x2½=7½ instead of 3x2=6). But in the second transaction they may agree to swap, say, 4 apples for 1 orange (the price ratio is 1/4). If they do so, they can move from (7,2½,3,2½) to, say, (6½,2⅝,3½,2⅜) if they swap ½ of an apple for ⅛ of an orange. This is also a Pareto improvement, as it provides the first consumer with utility 6½x2⅝≈17.37, and the second consumer with utility 3½x2⅜≈8.31. By doing subsequent transactions (each at a perhaps different price ratio) they can move further.

E-7 We use the price-taking assumption, when we argue that $p^*_1(x'_{11}+...+x'_{k1}) +...+ p^*_n(x'_{1n}+...+x'_{kn}) > p^*_1(x^*_{11}+...+x^*_{k1}) +...+ p^*_n(x^*_{1n}+...+x^*_{kn})$, i.e. when we assume that **x**' was preferred to **x***, but nevertheless **x*** was chosen when the price was **p***. If the prices were not exogenous (i.e. if somebody could manipulate them, prices at the left-hand-side and at the right-hand-side could be different).

E-8 Indeed the education may change marginal productivities. In such a case, the game becomes more complicated. There exist four categories of workers; (1) uneducated low-productive, (2) educated low-productive (whose productivity is higher than in the previous category, as a result of education), (3) uneducated high-productive, (4) educated high-productive (whose productivity is higher than in the previous category, as a result of education). But the employer cannot distinguish between (1) and (3), and between (2) and (4). He can only offer two levels of salaries: low for (1) and (3), and high for (2) and (4).

E-9 Let there be two firms who sell at prices $p_1=MC+\Delta_1$, and $p_2=MC+\Delta_2$, respectively (thus they add a mark up to their marginal costs assumed to be identical). If they add a monopolistic mark up μ, they will enjoy the maximum (monopolistic) profit jointly (π^M). We can assume that they will share 50% of this amount each ($\pi^M/2$). However, the Bertrand rivalry means that whoever offers a lower price will gain the entire market. Thus they have incentives to cheat, i.e. to offer a lower price perhaps despite that they promised not to do so. If one firm lowers its price by a little bit, then the second is left with no buyers, while the first one enjoys the profit π, lower than π^M but higher than $\pi^M/2$. If both of them cheat, then the total profit will be shared by both of them (π/2). Their rivalry can be portrayed as the following game:

	$MC+\mu-\varepsilon$	$MC+\mu$
$MC+\mu-\varepsilon$	**$(\pi/2,\pi/2)$**	$(\pi^M,0)$
$MC+\mu$	$(0,\pi^M)$	$(\pi^M/2,\pi^M/2)$

There is the only Nash equilibrium (π/2,π/2) in this game realized when both firms cheat (i.e. lower the price by the same amount). If the game is played repeatedly, the firms are likely to take earlier market prices as starting points for further cheating. They can continue lowering prices until they reach the marginal cost. Lowering the price below the marginal cost does not make sense and the MC+μ-μ is the lowest Nash equilibrium price predicted by the model. This is the Bertrand (MC) price.

E-10 The game from the lecture example serves as the case. It has three Nash equilibria: (D&C,A), (D&A,A), and (E&C,C). The first two imply for the incumbent attack in the second stage. However, such a strategy for the incumbent in the second stage is not a Nash equilibrium. Hence only the last one "survives" the second stage. Therefore the first two are not SPNE.

E-11 In the definition of Nash equilibrium, one needs to check whether players have a motivation to unilaterally change their strategies. The undercut-proof equilibrium in the discrete Hotelling model assumes that players may switch to other strategies simultaneously (not unilaterally). It may turn out that there are some attractive strategies that rivals can choose, but they do not posses characteristics applied in the definition of the Nash equilibrium. Consequently, a pair of prices can be found undercut-proof, but it can fail to be a Nash equilibrium. In other words, in the case of the Nash equilibrium we check that there is no motivation for a unilateral change; in the case of undercut-proof equilibrium we check that there is no possibility for changing one's price in a way which leaves the rival with no chance to reciprocate.

E-12 The recommendation depends on two circumstances: (1) whether the investor is ready to bear any risk; and (2) whether the investor can borrow the money. It is understood that investor has one unit of his/her own money either to buy a unit of the risk-free asset, or a unit of the risky asset (they cost the same). If investor is not ready to bear any risk then – irrespective of whether there is a possibility to borrow additional money – everything should be invested in the risk-free asset. The rate of return will be 2% for sure. Otherwise, for every level of the allowable risk (not exceeding σ), the investor should invest a proportion of $\alpha \in [0,1]$ in the risk-free asset (and $1-\alpha$ in the risky one) in order to achieve the expected rate of return $\alpha r_F+(1-\alpha)r_R$ and to face a risk measured by the standard deviation of $(1-\alpha)\sigma$. For instance, if the investor is ready to face the risk measured by 0.5σ, then 50% of the money should be invested in the risk-free asset, and the rest, i.e. 50%, in the risky asset. The expected rate of return will then be 4%. If the investor is ready to face risk exceeding σ, and money can be borrowed, then if the allowable level of risk measured by standard deviation is $n\sigma$ ($n>1$), the investor should borrow ($n-1$) units of money, add this to his/her own money and buy n units of the risky asset.

E-13 Please consider two portfolios at time t: Portfolio A – one Zero Coupon Bond with maturity T; and Portfolio B – $e^{-r(T-t)}$ of cash that we deposit (without any risk) at the rate r. The value of either portfolio at T is 1 (in the case of A – by definition; by applying the continuous formula for NPV, the value of portfolio B is $e^{-r(T-t)}e^{r(T-t)}=1$). Thus – by the no-arbitrage assumption – the value of both portfolios has to be the same at t as well (otherwise it would be possible to make a risk-free profit by trading these portfolios).

E-14 Let $x \in [0,1]$ denote a continuum of potential suppliers of a newly developed product. Let us assume that there is one supplier, say, $x=0$ who invented the product at time $t=0$. The supply of the product is thus $S(0,0)>0$, and $S(x,0)=0$ for any $x \in (0,1]$. We can assume further that – because of spill-over effects – the supply from other agents $x \in (0,1]$ will grow, i.e. it will be $S(x,t)>0$ for $t>0$. It is not clear how to interpret the proximity of other suppliers to the one who made the innovation, but perhaps their geographical location could explain how quickly they start to produce. These are arguments for using the econophysics approach. However, there are important reasons that justify a somewhat more sceptical view. First of all, it may be difficult to order potential suppliers along the [0,1] segment. Second, technology diffusion process may not be purely random (it is not like Brownian motion). Third, production costs may determine the level of supply from a given firm, while heat diffusion equation does not account for such additional aspects. Finally, with more realistic assumptions on transaction costs, tax incentives etc. the model will lose its tractability.

E-15 "Complementarity" means that $(\partial f(\mathbf{x})/\partial \mathbf{x} - \lambda_1 \partial g_1(\mathbf{x})/\partial \mathbf{x} + \ldots - \lambda_m \partial g_m(\mathbf{x})/\partial \mathbf{x}) \cdot \mathbf{x} = 0$. The expression at the left hand side is a scalar product of two vectors: $[\partial f(\mathbf{x})/\partial x_1 - \lambda_1 \partial g_1(\mathbf{x})/\partial x_1 + \ldots - \lambda_m \partial g_m(\mathbf{x})/\partial x_1,\ \partial f(\mathbf{x})/\partial x_2 - \lambda_1 \partial g_1(\mathbf{x})/\partial x_2 + \ldots - \lambda_m \partial g_m(\mathbf{x})/\partial x_2, \ldots,\ \partial f(\mathbf{x})/\partial x_n - \lambda_1 \partial g_1(\mathbf{x})/\partial x_n + \ldots - \lambda_m \partial g_m(\mathbf{x})/\partial x_n]$ and $\mathbf{x}$. If all coordinates of the vector $\mathbf{x}$ are positive, then for the product to be zero all coordinates of the left vector must be zeroes (since they are non-positive according to the definition from lecture 15).